

Optical Shop Testing

Third Edition

Edited by
DANIEL MALACARA



Wiley-Interscience
A John Wiley & Sons, Inc., Publication

Optical Shop Testing



THE WILEY BICENTENNIAL—KNOWLEDGE FOR GENERATIONS

Each generation has its unique needs and aspirations. When Charles Wiley first opened his small printing shop in lower Manhattan in 1807, it was a generation of boundless potential searching for an identity. And we were there, helping to define a new American literary tradition. Over half a century later, in the midst of the Second Industrial Revolution, it was a generation focused on building the future. Once again, we were there, supplying the critical scientific, technical, and engineering knowledge that helped frame the world. Throughout the 20th Century, and into the new millennium, nations began to reach out beyond their own borders and a new international community was born. Wiley was there, expanding its operations around the world to enable a global exchange of ideas, opinions, and know-how.

For 200 years, Wiley has been an integral part of each generation's journey, enabling the flow of information and understanding necessary to meet their needs and fulfill their aspirations. Today, bold new technologies are changing the way we live and learn. Wiley will be there, providing you the must-have knowledge you need to imagine new worlds, new possibilities, and new opportunities.

Generations come and go, but you can always count on Wiley to provide you the knowledge you need, when and where you need it!

WILLIAM J. PESCE
PRESIDENT AND CHIEF EXECUTIVE OFFICER

PETER BOOTH WILEY
CHAIRMAN OF THE BOARD

Optical Shop Testing

Third Edition

Edited by
DANIEL MALACARA



Wiley-Interscience
A John Wiley & Sons, Inc., Publication

Contents

Preface	xvii
Contributors	xix
Chapter 1. Newton, Fizeau, and Haidinger Interferometers	1
<i>M. V. Mantravadi and D. Malacara</i>	
1.1. Introduction	1
1.2. Newton Interferometer	1
1.2.1. Source and Observer's Pupil Size Considerations	9
1.2.2. Some Suitable Light Sources	11
1.2.3. Materials for the Optical Flats	12
1.2.4. Simple Procedure for Estimating Peak Error	12
1.2.5. Measurement of Spherical Surfaces	13
1.2.6. Measurement of Aspheric Surfaces	15
1.2.7. Measurement of Flatness of Opaque Surfaces	17
1.3. Fizeau Interferometer	17
1.3.1. The Basic Fizeau Interferometer	18
1.3.2. Coherence Requirements for the Light Source	20
1.3.3. Quality of Collimation Lens Required	22
1.3.4. Liquid Reference Flats	23
1.3.5. Fizeau Interferometer with Laser Source	23
1.3.6. Multiple-Beam Fizeau Setup	24
1.3.7. Testing Nearly Parallel Plates	26
1.3.8. Testing the Inhomogeneity of Large Glass or Fused Quartz Samples	27
1.3.9. Testing the Parallelism and Flatness of the Faces of Rods, Bars and Plates	28
1.3.10. Testing Cube Corner and Right-Angle Prisms	28
1.3.11. Fizeau Interferometer for Curved Surfaces	30
1.3.12. Testing Concave and Convex Surfaces	32
1.4. Haidinger Interferometer	33
1.4.1. Applications of Haidinger Fringes	35
1.4.2. Use of Laser Source for Haidinger Interferometer	36
1.4.3. Other Applications of Haidinger Fringes	39
1.5. Absolute Testing of Flats	40

Chapter 2. Twyman–Green Interferometer	46
<i>D. Malacara</i>	
2.1. Introduction	46
2.2. Beam-Splitter	48
2.2.1. Optical Path Difference Introduced by the Beam Splitter Plate	49
2.2.2. Required Accuracy in the Beam Splitter Plate	51
2.2.3. Polarizing Cube Beam Splitter	53
2.2.4. Nonpolarizing Cube Beam Splitter	55
2.3. Coherence Requirements	56
2.3.1. Spatial Coherence	56
2.3.2. Temporal Coherence	60
2.4. Uses of a Twyman–Green Interferometer	62
2.4.1. Testing of Prisms and Diffraction Rulings	64
2.4.2. Testing of Lenses	69
2.4.3. Testing of Microscope Objectives	71
2.5. Compensation of Intrinsic Aberrations in the Interferometer	72
2.6. Unequal-Path Interferometer	73
2.6.1. Some Special Designs	75
2.6.2. Improving the Fringe Stability	76
2.7. Open Path Interferometers	77
2.7.1. Mach-Zehnder Interferometers	77
2.7.2. Oblique Incidence Interferometers	78
2.8. Variations from the Twyman–Green Configuration	80
2.8.1. Multiple Image Interferometers	80
2.8.2. Interferometers with Diffractive Beam Splitters	80
2.8.3. Phase Conjugating Interferometer	81
2.9. Twyman–Green Interferograms and their Analysis	83
2.9.1. Analysis of Interferograms of Arbitrary Wavefronts	91
Chapter 3. Common-Path Interferometers	97
<i>S. Mallick and D. Malacara</i>	
3.1. Introduction	97
3.2. Burch’s Interferometer Employing Two Matched Scatter Plates	98
3.2.1. Fresnel Zone Plate Interferometer	102
3.2.2. Burch and Fresnel Zone Plate Interferometers for Aspheric Surfaces	102
3.2.3. Burch and Fresnel Zone Plate Interferometers for Phase Shifting	102
3.3. Birefringent Beam Splitters	104
3.3.1. Savart Polariscopes	104
3.3.2. Wollaston Prism	106
3.3.3. Double-Focus Systems	107
3.4. Lateral Shearing Interferometers	108
3.4.1. Use of a Savart Polariscopes	108
3.4.2. Use of a Wollaston Prism	111

3.5. Double-Focus Interferometer	112
3.6. Saunders's Prism Interferometer	114
3.7. Point Diffraction Interferometer	116
3.8. Zernike Tests with Common-Path Interferometers	118
Chapter 4. Lateral Shear Interferometers	122
<i>M. Strojnik, G. Paez, and M. Mantravadi</i>	
4.1. Introduction	122
4.2. Coherence Properties of the Light Source	123
4.3. Brief Theory of Lateral Shearing Interferometry	124
4.3.1. Interferograms of Spherical and Flat Wavefronts	126
4.3.2. Interferograms of Primary Aberrations upon Lateral Shear	128
4.4. Evaluation of an Unknown Wavefront	134
4.5. Lateral Shearing Interferometers in Collimated Light (White Light Compensated)	137
4.5.1. Arrangements Based on the Jamin Interferometer	137
4.5.2. Arrangements Based on the Michelson Interferometer	139
4.5.3. Arrangements Based on a Cyclic Interferometer	140
4.5.4. Arrangements Based on the Mach–Zehnder Interferometer	142
4.6. Lateral Shearing Interferometers in Convergent Light (White Light Compensated)	143
4.6.1. Arrangements Based on the Michelson Interferometer	143
4.6.2. Arrangements Based on the Mach–Zehnder Interferometer	146
4.7. Lateral Shearing Interferometers Using Lasers	149
4.7.1. Other Applications of the Plane Parallel Plate Interferometer	152
4.8. Other Types of Lateral Shearing Interferometers	157
4.8.1. Lateral Shearing Interferometers Based on Diffraction	158
4.8.2. Lateral Shearing Interferometers Based on Polarization	162
4.9. Vectorial Shearing Interferometer	164
4.9.1. Shearing Interferometry	165
4.9.2. Directional Shearing Interferometer	166
4.9.3. Simulated Interferometric Patterns	168
4.9.4. Experimental Results	173
4.9.5. Similarities and Differences With Other Interferometers	176
Chapter 5. Radial, Rotational, and Reversal Shear Interferometer	185
<i>D. Malacara</i>	
5.1. Introduction	185
5.2. Radial Shear Interferometers	187
5.2.1. Wavefront Evaluation from Radial Shear Interferograms	189
5.2.2. Single-Pass Radial Shear Interferometers	190
5.2.3. Double-Pass Radial Shear Interferometers	195
5.2.4. Laser Radial Shear Interferometers	197
5.2.5. Thick-Lens Radial Shear Interferometers	202

5.3. Rotational Shear Interferometers	204
5.3.1. Source Size Uncompensated Rotational Shear Interferometers	207
5.3.2. Source Size Compensated Rotational Shear Interferometers	211
5.4. Reversal Shear Interferometers	211
5.4.1. Some Reversal Shear Interferometers	213
Chapter 6. Multiple-Beam Interferometers	219
<i>C. Roychoudhuri</i>	
6.1. Brief Historical Introduction	219
6.2. Precision in Multiple-Beam Interferometry	221
6.3. Multiple-Beam Fizeau Interferometer	224
6.3.1. Conditions for Fringe Formation	224
6.3.2. Fizeau Interferometry	229
6.4. Fringes of Equal Chromatic Order	232
6.5. Reduction of Fringe Interval in Multiple-Beam Interferometry	235
6.6. Plane Parallel Fabry–Perot Interferometer	236
6.6.1. Measurement of Thin-Film Thickness	236
6.6.2. Surface Deviation from Planeness	237
6.7. Tolansky Fringes with Fabry–Perot Interferometer	241
6.8. Multiple-Beam Interferometer for Curved Surfaces	243
6.9. Coupled and Series Interferometers	244
6.9.1. Coupled Interferometer	245
6.9.2. Series Interferometer	246
6.10. Holographic Multiple-Beam Interferometers	247
6.11. Temporal Evolution of FP Fringes and Its Modern Applications	247
6.12. Final Comments	250
Chapter 7. Multiple-Pass Interferometers	259
<i>P. Hariharan</i>	
7.1. Double-Pass Interferometers	259
7.1.1. Separation of Aberrations	259
7.1.2. Reduction of Coherence Requirements	262
7.1.3. Double Passing for Increased Accuracy	264
7.2. Multipass Interferometry	266
Chapter 8. Foucault, Wire, and Phase Modulation Tests	275
<i>J. Ojeda-Castañeda</i>	
8.1. Introduction	275
8.2. Foucault or Knife-Edge Test	275
8.2.1. Description	275
8.2.2. Geometrical Theory	280
8.2.3. Physical Theory	289
8.3. Wire Test	293
8.3.1. Geometrical Theory	297

8.4.	Platzeck–Gaviola Test	298
8.4.1.	Geometrical Theory	299
8.5.	Phase Modulation Tests	302
8.5.1.	Zernike Test and its Relation to the Smart Interferometer	302
8.5.2.	Lyot Test	305
8.5.3.	Wolter Test	307
8.6.	Ritchey–Common Test	310
8.7.	Conclusions	313

Chapter 9. Ronchi Test 317

A. Cornejo-Rodriguez

9.1.	Introduction	317
9.1.1.	Historical Introduction	317
9.2.	Geometrical Theory	318
9.2.1.	Ronchi Patterns for Primary Aberrations	320
9.2.2.	Ronchi Patterns for Aspherical Surfaces	327
9.2.3.	Null Ronchi Rulings	328
9.3.	Wavefront Shape Determination	331
9.3.1.	General Case	333
9.3.2.	Surfaces with Rotational Symmetry	335
9.4.	Physical Theory	337
9.4.1.	Mathematical Treatment	337
9.4.2.	Fringe Contrast and Sharpness	340
9.4.3.	Physical versus Geometrical Theory	343
9.5.	Practical Aspects of the Ronchi Test	344
9.6.	Some Related Tests	347
9.6.1.	Concentric Circular Grid	347
9.6.2.	Phase Shifting Ronchi Test	348
9.6.3.	Sideband Ronchi Test	348
9.6.4.	Lower Test	349
9.6.5.	Ronchi–Hartmann and Null Hartmann Tests	350

Chapter 10. Hartmann, Hartmann–Shack, and Other Screen Tests 361

D. Malacara-Doblado and I. Ghozeil

10.1.	Introduction	361
10.2.	Some Practical Aspects	363
10.3.	Hartmann Test Using a Rectangular Screen	366
10.4.	Wavefront Retrieval	368
10.4.1.	Tilt and Defocus Removal	368
10.4.2.	Trapezoidal Integration	370
10.4.3.	Southwell Algorithm	373
10.4.4.	Polynomial Fitting	374
10.4.5.	Other Methods	375

10.5.	Hartmann Test Using a Screen with Four Holes	376
10.5.1.	Four Holes in Cross	377
10.5.2.	Four Holes in X	378
10.6.	Hartmann Test of Ophthalmic Lenses	379
10.7.	Hartmann Test Using Nonrectangular Screens	379
10.7.1.	Radial Screen	380
10.7.2.	Helical Screen	382
10.8.	Hartmann–Shack Test	383
10.9.	Crossed Cylinder Test	386
10.10.	Testing with an Array of Light Sources or Printed Screens	387
10.10.1.	Testing Convergent Lenses	388
10.10.2.	Testing Concave and Convex Surfaces	389
10.11.	Michelson–Gardner–Bennett Tests	393
10.12.	Other Developments	394
Chapter 11. Star Tests		398
<i>D. Malacara and W. T. Welford</i>		
11.1.	Introduction	398
11.2.	Star Test with Small Aberrations	399
11.2.1.	The Aberration Free Airy Pattern	400
11.2.2.	The Defocused Airy Pattern	403
11.2.3.	Polychromatic Light	405
11.2.4.	Systems with Central Obstructions	407
11.2.5.	Effects of Small Aberrations	408
11.2.6.	Gaussian Beams	409
11.2.7.	Very Small Convergence Angles (Low Fresnel Numbers)	409
11.3.	Practical Aspects with Small Aberrations	410
11.3.1.	Effects of Visual Star Testing	410
11.3.2.	The Light Source for Star Testing	412
11.3.3.	The Arrangement of the Optical System for Star Testing	413
11.3.4.	Microscope Objectives	415
11.4.	The Star Test with Large Aberrations	416
11.4.1.	Spherical Aberration	417
11.4.2.	Longitudinal Chromatic Aberration	418
11.4.3.	Axial Symmetry	418
11.4.4.	Astigmatism and Coma	419
11.4.5.	Distortion	419
11.4.6.	Non-Null Tests	420
11.5.	Wavefront Retrieval with Slope and Curvature Measurements	421
11.5.1.	The Laplacian and Local Average Curvatures	421
11.5.2.	Wavefront Determination with Iterative Fourier Transforms	422
11.5.3.	Irradiance Transport Equation	425

11.6.	Wavefront Determination with Two Images Using the Irradiance Transport Equation	426
11.7.	Wavefront Determination with a Single Defocused Image Using Fourier Transform Iterations	429
11.8.	Wavefront Determination with Two or Three Defocused Images Using Fresnel Transform Iterations	430
	Chapter 12. Testing of Aspheric Wavefronts and Surfaces	435
	<i>D. Malacara, K. Creath, J. Schmit and J. C. Wyant</i>	
12.1.	Introduction	435
12.2.	Some Methods to Test Aspheric Wavefronts	437
12.3.	Imaging of the Interference Pattern in Non-Null Tests	439
12.4.	Some Null Testing Configurations	442
	12.4.1. Flat and Concave Spherical Surfaces	442
	12.4.2. Telescope Refracting Objectives	442
	12.4.3. Concave Paraboloidal Surfaces	443
	12.4.4. Concave Ellipsoidal or Spheroidal Surfaces	444
12.5.	Testing of Convex Hyperboloidal Surfaces	445
	12.5.1. Hindle Type Tests	445
	12.5.2. Testing by Refraction	449
12.6.	Testing of Cylindrical Surfaces	453
12.7.	Early Compensators	454
	12.7.1. Couder, Burch, and Ross Compensators	456
	12.7.2. Dall Compensator	458
12.8.	Refractive Compensators	461
	12.8.1. Refractive Offner Compensator	462
	12.8.2. Shafer Compensator	464
	12.8.3. General Comments about Refracting Compensators	465
12.9.	Reflecting Compensators	466
	12.9.1. Reflective Offner Compensators	468
	12.9.2. Reflective Adaptive Compensator	471
12.10.	Other Compensators for Concave Conicoids	471
12.11.	Interferometers Using Real Holograms	474
	12.11.1. Holographic Wavefront Storage	476
	12.11.2. Holographic Test Plate	476
12.12.	Interferometers Using Synthetic Holograms	477
	12.12.1. Fabrication of Computer-Generated Holograms (CGHs)	478
	12.12.2. Using a CGH in an Interferometer	480
	12.12.3. Off-Axis CGH Aspheric Compensator	483
	12.12.4. In-Line CGH Aspheric Compensator	485
	12.12.5. Combination of CGH with Null Optics	486
12.13.	Aspheric Testing with Two-Wavelength Holography	488
12.14.	Wavefront Stitching	491

12.14.1. Annular Zones	491
12.14.2. Circular Zones	493
12.14.3. Dynamic Tilt Switching	493
Chapter 13. Zernike Polynomial and Wavefront Fitting	498
<i>Virendra N. Mahajan</i>	
13.1. Introduction	498
13.2. Aberrations of a Rotationally Symmetric System with a Circular Pupil	499
13.2.1. Power Series Expansion	499
13.2.2. Primary or Seidel Aberration Function	501
13.2.3. Secondary or Schwarzschild Aberration Function	504
13.2.4. Zernike Circle Polynomial Expansion	505
13.2.5. Zernike Circle Polynomials as Balanced Aberrations for Minimum Wave Aberration Variance	508
13.2.6. Relationships Between Coefficients of Power-Series and Zernike-Polynomial Expansions	510
13.2.7. Conversion of Seidel Aberrations into Zernike Aberrations	513
13.2.8. Conversion of Zernike Aberrations into Seidel Aberrations	515
13.3. Aberration Function of a System with a Circular Pupil, but Without an Axis of Rotational Symmetry	516
13.3.1. Zernike Circle Polynomial Expansion	516
13.3.2. Relationships Among the Indices n , m , and j	518
13.3.3. Isometric, Interferometric, and PSF Plots for a Zernike Circle Polynomial Aberration	520
13.3.4. Primary Zernike Aberrations and Their Relationships with Seidel Aberrations	521
13.4. Zernike Annular Polynomials as Balanced Aberrations for Systems with Annular Pupils	525
13.4.1. Balanced Aberrations	525
13.4.2. Zernike Annular Polynomials	525
13.4.3. Isometric, Interferometric, and PSF Plots for a Zernike Annular Polynomial Aberration	529
13.5. Determination of Zernike Coefficients From Discrete Wavefront Error Data	530
13.5.1. Introduction	530
13.5.2. Orthonormal Coefficients and Aberration Variance	535
13.5.3. Orthonormal Polynomials	537
13.5.4. Zernike Coefficients	538
13.5.5. Numerical Example	539
13.6. Summary	543

Chapter 14. Phase Shifting Interferometry	547
<i>Horst Schreiber and John H. Bruning</i>	
14.1. Introduction	547
14.2. Fundamental Concepts	548
14.3. Advantages of PSI	550
14.4. Methods of Phase Shifting	552
14.5. Detecting the Wavefront Phase	557
14.6. Data Collection	560
14.6.1. Temporal Methods	560
14.6.2. Spatial Methods	564
14.7. PSI Algorithms	568
14.7.1. Three Step Algorithms	569
14.7.2. Least-Squares Algorithms	571
14.7.3. Carré Algorithm	574
14.7.4. Family of Averaging Algorithms	576
14.7.5. Hariharan Algorithm	577
14.7.6. $2 + 1$ Algorithm	580
14.7.7. Methods to Generate Algorithms	582
14.7.8. Methods to Evaluate Algorithms	586
14.7.9. Summary of Algorithms	591
14.8. Phase Shift Calibration	596
14.9. Error Sources	599
14.9.1. Phase Shift Errors	600
14.9.2. Detector Nonlinearities	602
14.9.3. Source Stability	605
14.9.4. Quantization Errors	606
14.9.5. Vibration Errors	607
14.9.6. Air Turbulence	610
14.9.7. Extraneous Fringes and Other Coherent Effects	610
14.9.8. Interferometer Optical Errors	611
14.10. Detectors and Spatial Sampling	613
14.10.1. Solid State Sensors	613
14.10.2. Spatial Sampling	614
14.11. Quality Functions	617
14.11.1. Modulation	618
14.11.2. Residues	619
14.11.3. Filtering	622
14.12. Phase Unwrapping	623
14.12.1. Unwrapping in One Dimension	623
14.12.2. 2-D Phase Unwrapping	625
14.12.3. Path-Following Algorithms	626
14.12.4. Path Independent Methods	628
14.13. Aspheres and Extended Range PSI Techniques	629
14.13.1. Aliasing	630

14.13.2. Sub-Nyquist Interferometry	631
14.13.3. Two Wavelength PSI	635
14.13.4. Subaperture Stitching	637
14.14. Other Analysis Methods	638
14.14.1. Zero Crossing Analysis	638
14.14.2. Synchronous Detection	639
14.14.3. Heterodyne Interferometry	640
14.14.4. Phase Lock Interferometry	641
14.14.5. Spatial Synchronous and Fourier Methods	642
14.15. Computer Processing and Output	644
14.16. Implementation and Applications	647
14.16.1. Commercial Instrumentation	647
14.16.2. Interferometer Configurations	650
14.16.3. Absolute Calibration	651
14.16.4. Sources	654
14.16.5. Alignment Fiducials	655
14.17. Future Trends for PSI	655
Chapter 15. Surface Profilers, Multiple Wavelength, and White Light Interferometry	667
<i>J. Schmit, K. Creath, and J. C. Wyant</i>	
15.1. Introduction to Surface Profilers	667
15.1.1. Contact Profilometers	668
15.1.2. Optical Profilometers	668
15.1.3. Interferometric Optical Profilometers	668
15.1.4. Terms and Issues in Determining System Performance	669
15.2. Contact Profilometers	670
15.2.1. Stylus Profilers	670
15.2.2. Scanning Probe Microscopes	674
15.2.3. Comparison of AFM and Stylus Profiler	683
15.3. Optical Profilers	685
15.3.1. Optical Focus Sensors	687
15.3.2. Confocal Microscopy	689
15.4. Interferometric Optical Profilers	695
15.4.1. Common Features	696
15.5. Two Wavelength and Multiple Wavelength Techniques	702
15.5.1. Two-wavelengths Phase Measurement	704
15.5.2. Multiple-wavelength Phase Measurement	707
15.5.3. Reducing Measurement Time	710
15.6. White Light Interference Optical Profilers	711
15.6.1. White Light Interference	711
15.6.2. Image Buildup	712
15.6.3. Signal Processing of White Light Interferograms	713

15.6.4.	Light Sources	716
15.6.5.	Dispersion in White Light Fringes	716
15.6.6.	Other Names for Interferometric Optical Profilers	723
15.7.	Wavelength Scanning Interferometer	724
15.7.1.	Wavelength Tunable Light Sources	724
15.7.2.	Image Buildup	725
15.7.3.	Signal Analysis	728
15.7.4.	Film and Plate Thickness Measurement	729
15.8.	Spectrally Resolved White Light Interferometry (SRWLI)	731
15.8.1.	Image Buildup	731
15.8.2.	Signal Analysis	732
15.8.3.	Other Names for Spectral Interferometry	735
15.9.	Polarization Interferometers	735
15.9.1.	Differential Interference Contrast Microscope (Nomarski)	736
15.9.2.	Geometric Phase Shifting	738
15.10.	Optical Ranging Methods	741
15.10.1.	Interferometric Ranging	741
15.10.2.	Optical Triangulation	742
15.10.3.	Time of Flight (TOF)	742
15.11.	Summary	742
Chapter 16. Optical Metrology of Diffuse Surfaces		756
<i>K. Creath, J. Schmit, and J. C Wyant</i>		
16.1.	Moiré and Fringe Projection Techniques	756
16.1.1.	Introduction	756
16.1.2.	What is Moiré?	757
16.1.3.	Moiré and Interferograms	762
16.1.4.	Historical Review	768
16.1.5.	Fringe Projection	769
16.1.6.	Shadow Moiré	773
16.1.7.	Projection Moiré	777
16.1.8.	Two-angle Holography	778
16.1.9.	Common Features	779
16.1.10.	Comparison to Conventional Interferometry	779
16.1.11.	Coded and Structured Light Projection	780
16.1.12.	Applications	781
16.1.13.	Summary	783
16.2.	Holographic and Speckle Tests	783
16.2.1.	Introduction	783
16.2.2.	Holographic Interferometry for Nondestructive Testing	784
16.2.3.	Speckle Interferometry and Digital Holography	791

Chapter 17. Angle, Prisms, Curvature, and Focal Length Measurements	808
<i>Z. Malacara</i>	
17.1. Introduction	808
17.2. Angle Measurements	808
17.2.1. Divided Circles and Goniometers	808
17.2.2. Autocollimator	810
17.2.3. Interferometric Measurements of Angles	812
17.3. Testing of Prisms	812
17.4. Radius of Curvature Measurements	817
17.4.1. Mechanical Measurement of Radius of Curvature	817
17.4.2. Optical Measurement of Radius of Curvature	820
17.5. Focal Length Measurements	823
17.5.1. Nodal Slide Bench	823
17.5.2. Focimeters	824
17.5.3. Other Focal Length Measurements	825
Chapter 18. Mathematical Representation of an Optical Surface and Its Characteristics	832
<i>D. Malacara</i>	
18.1. Definition of an Optical Surface	832
18.1.1. Parameters for Conic Surfaces	835
18.1.2. Some Useful Expansions of z	835
18.1.3. Aberration of the Normals to the Surface	836
18.2. Caustic Produced by an Aspheric Surface	837
18.3. Primary Aberrations of Spherical Surfaces	839
18.3.1. Spherical Aberration of and Aspherical Surface	839
18.3.2. Coma of a Concave Mirror	840
18.3.3. Astigmatism of a Concave Mirror	841
18.4. Astigmatic Surfaces	841
18.4.1. Toroidal Surface	842
18.4.2. Astigmatic Ellipsoidal and Oblate Spheroidal Surfaces	842
18.4.3. Sphero-Cylindrical Surface	844
18.4.4. Testing Astigmatic Surfaces and Reference Astigmatic Surface	846
18.4.5. Comparison Between Astigmatic Surfaces	847
18.5. Off-Axis Conicoids	849
18.5.1. Off-Axis Paraboloids	850
Appendix. Optical Testing Programs	852
Index	855

Preface

Since the publication of the second edition of this book, many important advances have taken place in the field of optical testing. On one hand, the requirements for faster and more precise tests are stronger than ever; on the other hand, the new technological tools permit us to do these tasks much better than before. The need to describe these advances in this book would lead us to a thicker and hence more expensive book. This was not compatible with our desire to keep the price as low as possible, and therefore several new things had to be done. One of them was to reduce the description of some of the most mathematical sections in the book so as leaving space for some more applied subjects. Another modification was to reduce as much as possible the number of references at the end of each chapter, leaving only the most relevant ones. To compensate it, a CD with the complete and almost exhaustive list of references is included in the book. Another advantage of this is that the full list of references is properly classified by topics or its possible applications. Since many publications may have two, three or more subjects, it is included in each of these sections. For example, a publication may describe a test that is useful for testing flats, spheres, and prisms. In that case, this publication is present in all of these sections. A reader with a particular optical tests need, may find some help by using this reference list in PDF format. The list of publications in optical testing is so large that it is impossible to expect that no important reference is missing. If so, the Editor apologizes for overlooking any important reference. Of course the list may be updated every one or two years.

In the CD, which is included in the book, the reader will also find two programs for Windows, which may be useful when teaching or working in optical testing. One of these programs displays on the screen of the computer some of the most common interferogram types, test patterns, or transverse or wave aberration functions. These images as well as their associated numerical results can be saved in computer files. The other program helps in the design of phase shifting algorithms with the desired properties by using its Fourier mathematical representation.

Some classic chapters where no important recent advances have taken place remain almost the same, but most chapters are substantially modified, updated and enlarged, describing the most important new developments.

In the process of revising the book many important people have contributed, for example, the highly important work of authors of each chapter. A book like this would have been absolutely impossible without their fundamental contributions. The Editor is deeply thankful to all of them. Also, many other people, colleagues, and

friends contributed with many constructive criticisms and suggestions and some times with hard work, preparing figures, or collecting references. The help of my secretary Marisa and my student Armando Gómez has been extremely useful. Finally, I cannot conclude without acknowledging the support and encouragement of my wife Isabel and all my family with whom I am indebted and grateful.

DANIEL MALACARA

Contributors

John H. Bruning

GCA Tropel Div.
60 O'Connor Road
Fairport, NY 14450
U.S.A.
Phone: (585) 388-3500
Fax: (585) 377-6332
E-mail: bruningjh@corning.com

Alejandro Cornejo-Rodríguez

Instituto Nacional de Astrofísica, Óptica
y Electrónica
Luis Enrique Erro # 1
P.O. Box 51, 216
Postal Code 72840
Tonantzintla, Puebla
Mexico
Phone: 52 (222) 266 31 00
Fax: 52 (222) 247 22 31
E-mail: acornejo@inaoep.mx

Katherine Creath

Optineering
Tucson Arizona
U.S.A.
Phone: (520) 882-2950
Fax: (520) 882-6976
E-mail: kcreath@ieee.org

Parmeswaran Hariharan

School of Physics, University of Sydney
Australia
Phone: (612) 9413 7159
Fax: (612) 9413 7200
E-mail: hariharan_optics@hotmail.com

Virendra N. Mahajan

The Aerospace Corporation
2350 E. El Segundo Blvd
El Segundo, CA 90245
U.S.A.

Daniel Malacara

Centro de Investigaciones en Óptica, A. C.
Loma del Bosque # 115 Col. Lomas del
Campestre
P.O. Box 1-948
Postal Code 7 150
León, Gto.
México
Phone: 52 (477) 441 42 00
Fax: 52 (447) 441 42
E-mail: dmalacara@cio.mx

Daniel Malacara-Doblado

Centro de Investigaciones en Óptica, A. C.
Loma del Bosque # 115 Col. Lomas del
Campestre
P. O. Box. 1-948
Postal Code 7150
León, Gto.
México
Phone: 52 (477) 441 42 00
Fax: 52 (447) 441 42
E-mail: zmalacar@cio.mx

Zacariás Malacara

Centro de Investigaciones en Óptica, A. C.
Loma del Bosque # 115 Col. Lomas del
Campestre
P. O. Box. 1-948

Postal Code 7150
León, Gto.
México
Phone: 52 (477) 441 42 00
Fax: 52 (447) 441 42
E-mail: zmalacar@cio.mx

Murty V. Mantravadi
2301 West 120th Street
P. O. Box 5032
Hawthorne, CA 90251-5032
U.S.A.
Phone: (213) 606-6050
Fax: (213) 606-6495

Jorge Ojeda-Castañeda
Universidad de las Américas Puebla
Sta. Catarina Martir.
Postal Code 72820
Cholula Puebla
México
Phone: 52 (222) 229 202 00
E-mail: jorge.ojeda@uslap.mx

Gonzalo Páez Padilla
Centro de Investigaciones en Optica, A. C.
Loma del Bosque # 115 Col. Lomas del
Campestre
P.O. Box 1-948
Postal Code 7 150
León, Gto.
México
Phone: 52 (477) 441 42 00
Fax: 52 (447) 441 42
E-mail: gpaez@cio.mx

Chandrasekhar Roychoudhuri
Photonics Lab, Physics Department
University of Connecticut
54 Ahern Lane
Storrs, CT 06269-5192
U.S.A.
Voice: 860-486-2587

Fax: 860-486-2886
E-mail: chandra@phys.uconn.edu

Horst Schreiber
GCA Tropel Div.
60 O'Connor Road
Fairport, NY 14450
U.S.A.
Phone: (585) 388-3500
Fax: (585) 377-6332
E-mail: schreibeh@corning.com

Joanna Schmit
Veeco Instruments Inc.
2650 E. Elvira Rd.
Tucson, AZ 85706
U.S.A.
Phone: (520) 741-1044
Fax: (520) 294 1799
E-mail: jschmit@veeco.com

Marija Strojnik
Centro de Investigaciones en Optica, A. C.
Loma del Bosque # 115 Col. Lomas del
Campestre
P. O. Box 1-948
Postal Code 7 150
León, Gto.
México
Tel: 52 (477) 441 42 00
Fax: 52 (477) 441 42
E-mail: marias@cio.mx

James C. Wyant
College of Optical Sciences
The University of Arizona
1630 East University Boulevard
Tucson, AZ 85721-0094
U.S.A.
Phone: (520) 621-6997
E-mail: jcwyant@optics.arizona.edu

W. T. Welford, deceased

1

Newton, Fizeau, and Haidinger Interferometers

M. V. Mantravadi and D. Malacara

1.1. INTRODUCTION

This chapter has been updated by the second author; it includes much of the material from the previous version of the book. Newton, Fizeau, and Haidinger interferometers are among the simplest and most powerful tools available to a working optician. With very little effort, these interferometers can be set up in an optical workshop for routine testing of optical components to an accuracy of a fraction of the wavelength of light. Even though these instruments are simple in application and interpretation, the physical principles underlying them involve a certain appreciation and application of physical optics. In this chapter, we examine the various aspects of these interferometers and also consider the recent application of laser sources to them. The absolute testing of flats will also be considered in this chapter.

1.2. NEWTON INTERFEROMETER

We will take the liberty of calling any arrangement of two surfaces in contact illuminated by a monochromatic source of light a Newton interferometer. Thus, the familiar setup to obtain Newton rings in the college physical optics experiment is also a Newton interferometer; the only difference being the large air gap as one moves away from the point of contact, as seen in Figure 1.1. Because of this, it is sometimes necessary to view these Newton rings through a magnifier or even a low-power microscope. In the optical workshop, we are generally concerned that an optical flat, one being fabricated, is matching the accurate surface of another reference flat or that a curved spherical surface is matching the correspondingly opposite curved spherical master surface. Under these conditions, the air gap is seldom more than a few wavelengths of light in thickness. In the various forms of the

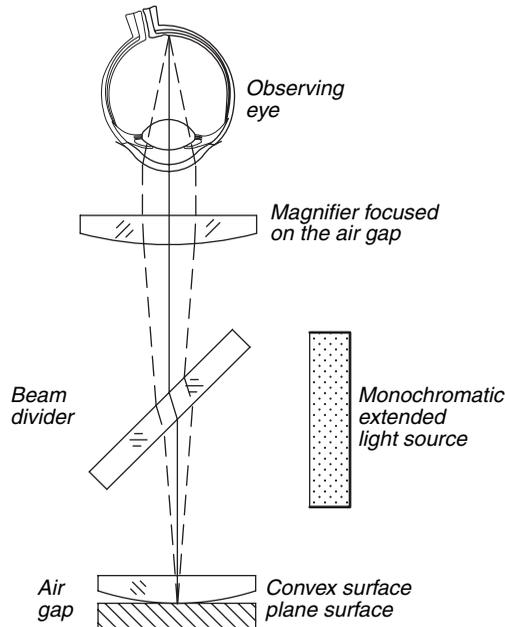


FIGURE 1.1. Illustration of the setup for Newton rings. A plano-convex lens of about 1 or 2 m in focal length is placed with its convex surface in contact with the plano surface of an optical flat and illuminated by monochromatic light.

Newton interferometer, we are mainly interested in determining the nonuniformity of this air gap thickness by observing and interpreting Newton fringes. A simple way to observe these Newton fringes is illustrated in Figure 1.2. Any light source such as a sodium vapor lamp, low-pressure mercury vapor lamp, or helium discharge lamp can be used in the setup. Under certain situations, even an ordinary tungsten lamp can serve this purpose.

Let us first see what happens when two perfect optical flats are placed one over the other with only a thin air gap between them as illustrated in Figure 1.3. The surfaces are not exactly parallel, so that the air gap is thinner on the left than on the right. Generally this separation is not zero at any place, unless the surfaces are extremely clean, and one presses very hard to get them in close contact. Hence, we may imagine that the two planes are projected backward, as shown in Figure 1.3, and they meet at a line of intersection. Let the monochromatic light of wavelength λ be incident on the optical flat combination having an angle α between them, almost normally. If the air gap is x at a given point, the two reflected rays will have an optical path difference (OPD) equal to $2x$. One of the reflected rays is reflected internally on one of the surfaces, while the other is reflected externally. We know that in dielectrics, like glass, one of these two reflected rays, and only one of them, has a phase change by 180° . In this case it is the reflected ray on the bottom surface which will have this phase change. Thus, the phase difference between the two reflected rays will produce a dark fringe when the optical path difference is an

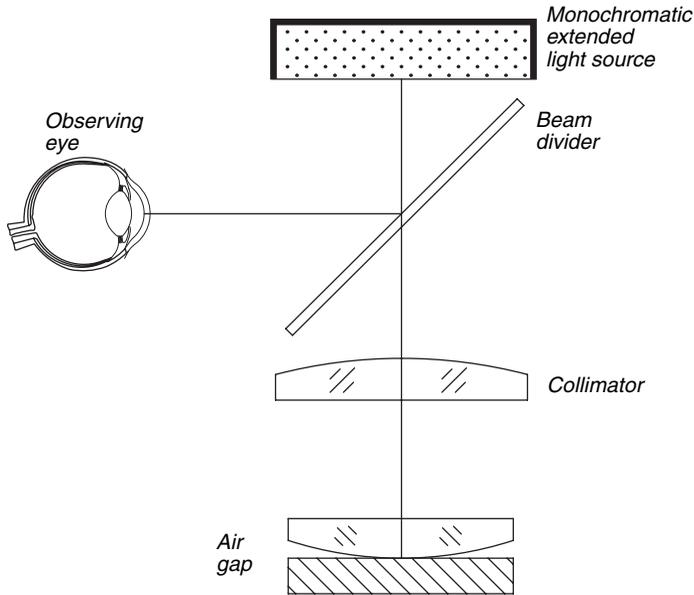


FIGURE 1.2. A simple arrangement to observe the Newton fringes in the optical workshop. With this arrangement plane and long radius spherical surfaces can be tested.

integer multiple of the wavelength. We may easily conclude that if the separation x is zero, there is a dark fringe.

Hence the dark fringes may be represented by

$$2\alpha x = n\lambda, \quad (1.1)$$

where n is an integer, and the bright fringes may be represented by

$$2\alpha x + \frac{\lambda}{2} = n\lambda. \quad (1.2)$$

Each of these equations represents a system of equally spaced straight fringes, and the distance d between two consecutive bright or dark fringes is

$$d = \frac{\lambda}{2\alpha}. \quad (1.3)$$

Thus the appearance of the fringes is as shown in Figure 1.3, when two good optical flats are put in contact with each other, forming a small air wedge, and are viewed in monochromatic light.

Now let us see what the appearance of Newton fringes is when one surface is optically flat while the other surface is not. Several situations are possible and in fact occur in actual practice. When one starts making a surface a plane, it does not turn out to be a plane on the first try; probably it becomes spherical with a long radius of curvature. It is necessary to test the surface from time to time with a reference flat to

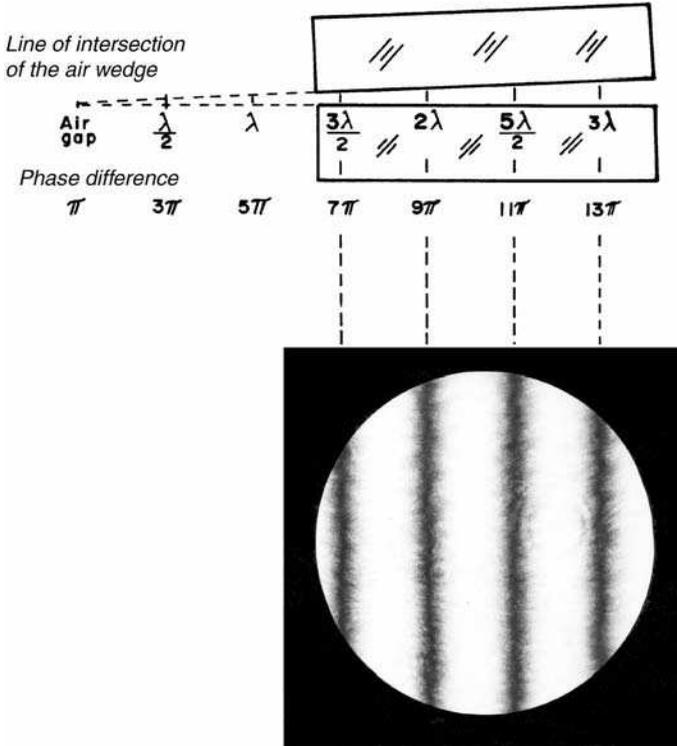


FIGURE 1.3. The principle of the formation of straight, equally spaced fringes between two optically plane surfaces when the air gap is in the form of a wedge. The fringes are parallel to the line of intersection of the two plane surfaces.

ascertain its deviation from flatness. Let us consider a spherical surface of large radius of curvature R in contact with the optical flat.

Then the sag of the surface is given by $x^2/2R$, where x is the distance measured from the center of symmetry. Hence the OPD is given by $x^2/R + \lambda/2$, and the positions of the dark fringes are expressed by

$$\frac{x^2}{R} = n\lambda. \quad (1.4)$$

Hence the distance of the n th dark fringe from the center is given by

$$x_n = \sqrt{nR\lambda}. \quad (1.5)$$

From this, it is easy to show that the distance between the $(n + 1)$ th and the n th fringe is given by

$$x_{n+1} - x_n = \sqrt{R\lambda}(\sqrt{n+1} - \sqrt{n}), \quad (1.6)$$

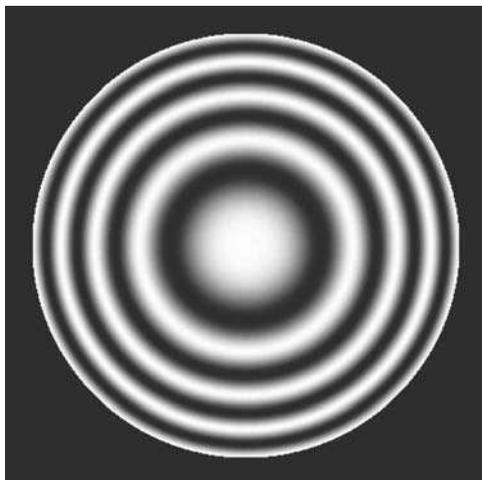


FIGURE 1.4. Appearance of the Newton fringes when a long radius of curvature is kept on a good optical flat. This situation is for a surface deviating 2λ from the plane at its maximum.

and similarly the distance between the $(n + 2)$ th and the $(n + 1)$ th fringe is given by

$$x_{n+2} - x_{n+1} = \sqrt{R\lambda}(\sqrt{n+2} - \sqrt{n+1}). \quad (1.7)$$

From Eqs. (1.6) and (1.7) we can form the ratio

$$\frac{x_{n+1} - x_n}{x_{n+2} - x_{n+1}} \approx 1 + \frac{1}{2n}. \quad (1.8)$$

Thus, it is seen that when we look at fringes with large values of n , they appear to be almost equally spaced. Hence, when we are testing for the presence of curvature in the surface, it is desirable to manipulate the plates in such a way that we see the fringes with lower order n . In Figure 1.4, the appearance of Newton fringes is shown when the maximum value of $x^2/2R$ is 2λ . Thus, there will be four circular fringes in this situation. If the maximum value of $x^2/2R$ is $\lambda/2$, we have just one circular fringe. Thus, by observation of full circular fringes, we can detect a maximum error of $\lambda/2$ in the flatness of the surface. If the maximum error is less than $\lambda/2$, we have to adopt a different procedure. In this case, the center of the symmetry of the circular fringes is displaced sideways by suitable manipulation of the two components. Thus, we obtain fringes in the aperture of the two surfaces in contact with a larger value of n ; these fringes are arcs of circles, and their separations are almost, but not exactly, equal. Let us take as examples of maximum value $x^2/2R = \lambda/4$ and $\lambda/8$. Figures 1.5 and 1.6, respectively, illustrate the appearance of the fringes in these two cases. As can be inferred, the fringes become straighter and straighter as the value of R increases.

In the optical workshop, it is also necessary to know whether the surface that is being tested is concave or convex with respect to the reference optical flat. This can

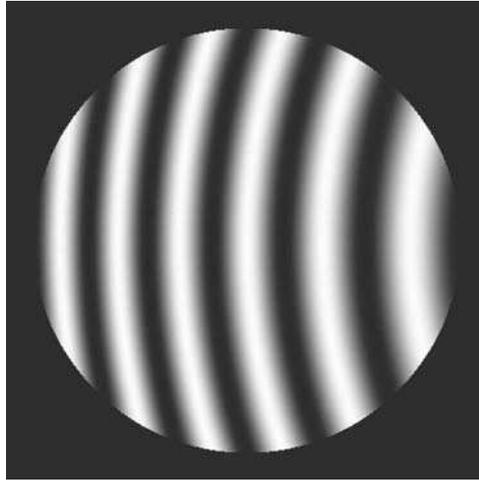


FIGURE 1.5. Appearance of the Newton fringes when a surface of long radius of curvature is kept on a good optical flat. This situation is for a surface deviating by $\lambda/4$ from the plane at its maximum. The center of symmetry of the fringes is outside the aperture of the surfaces, and hence only arcs of circles are seen.

be easily judged by several procedures. One simple method involves pressing near the edge of the top flat gently by means of a wooden stick or pencil. If the surface is convex, the center of the fringe system moves toward the point of the application of pressure. If the surface is concave, the center of the fringe system moves away from the point of the application of pressure, as shown in Figure 1.7 (a).

A second very simple method is to press near the center of the ring system on the top flat, as shown in Figure 1.7 (b). If the surface is convex, the center of the fringe is not displaced but the diameter of the circular fringes is increased.

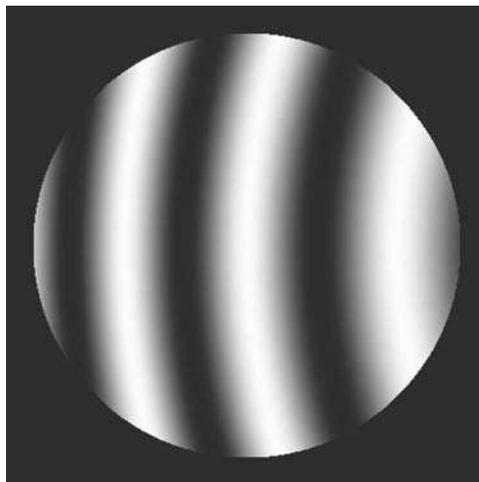


FIGURE 1.6. Same as Figure 1.5 except that the maximum error is $\lambda/8$ and some tilt is introduced.

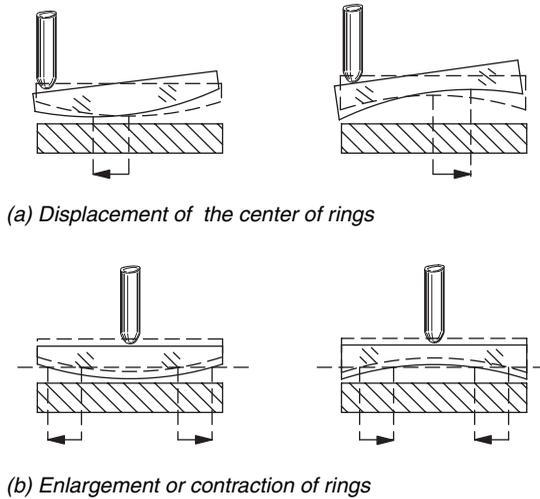


FIGURE 1.7. Two methods to determine whether the surface under test is convex or concave with respect to the surface: (a) by pressing near the edge and (b) by pressing near the center of the top plate.

Another method of deciding whether the surface is convex or concave involves the use of a source of white light. If slight pressure is applied at the center of the surfaces, the air gap at this point tends to become almost zero when the surface is convex. Hence the fringe at this point is dark, and the first bright fringe will be almost colorless or white. The next bright fringe is tinged bluish on the inside and reddish on the outside. On the contrary, if the surface is concave, the contact is not a point contact but occurs along a circle, and the air gap thickness tends to become zero along this circle. The dark fringe will be along this circle, and the sequence of colored fringes will be the same as before as one proceeds from the black fringe. This situation is illustrated in Figures 1.8 and 1.9. This procedure is not very easy to perform unless the surfaces are clean and is not generally recommended.

A fourth and simpler procedure is based on the movement of the fringe pattern as one moves the eye from a normal to an oblique viewing position. Before explaining this procedure, it is necessary to find a simple expression for the optical path difference between the two reflected rays at an air gap of thickness t and an angle of incidence θ . This is illustrated in Figure 1.10, where it can be seen that

$$OPD = \frac{2t}{\cos \theta} - 2t \tan \theta \sin \theta = 2t \cos \theta. \tag{1.9}$$

Thus, the OPD at the normal of incidence, namely $2t$, is always greater than the OPD at an angle θ for the same value of air gap thickness t . Using this fact, let us see what happens when we have a convex contact between the two surfaces. The air gap increases as we go away from the point of contact. When we view the fringes obliquely, the OPD at a particular point is decreased, and consequently the fringes appear to move away from the center as we move our eye from the normal to oblique

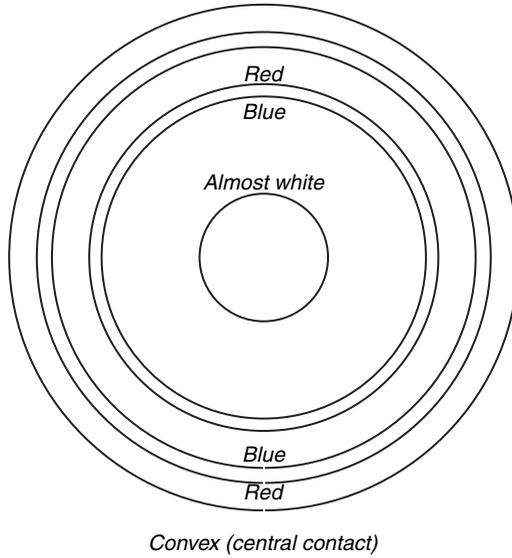


FIGURE 1.8. Convex contact and appearance of the colored fringes with white light illumination. Pressure is applied at the center.

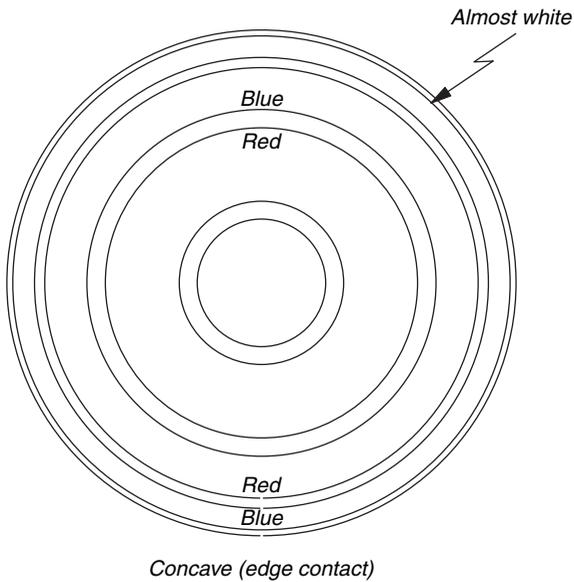
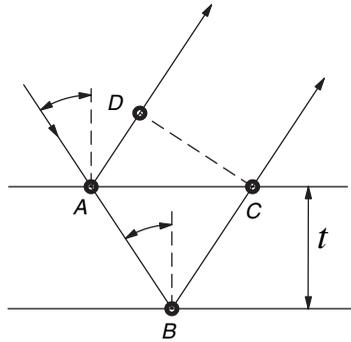


FIGURE 1.9. Convex contact and appearance of the colored fringes with white light illumination. Pressure is applied at the center.



$$OPD = AB + BC - AD = 2t \cos \theta$$

FIGURE 1.10. Ray diagram for calculation of the optical path difference between two reflected rays from an air gap of thickness t and angle of incidence θ .

position. The reverse of this situation occurs for a concave surface in contact with a plane surface.

We may consider many other situations where the surfaces are not plane or spherical. The nature and the appearance of such fringes when viewed are given in the usual manner in Table 1.1.

We have mentioned that the reference surface is a flat surface against which a nearly plane surface that is being made is tested. By the same procedure, spherical or cylindrical surfaces having long radii of curvature can be tested. However, when such surfaces have very short radii of curvature, it is necessary to use special illumination, which will be discussed in Section 1.2 on the Fizeau interferometer.

1.2.1. Source and Observer's Pupil Size Considerations

The OPD given in Eq. (1.9) shows that this value depends on the angle of the reflected rays being observed, which for small angles θ can be approximated by

$$OPD = 2t \cos \theta \approx 2t - t\theta^2. \tag{1.10}$$

Now, in the Newton interferometer we are interested in measuring glasses where t is not constant and thus θ is not constant either. Hence, to reduce the influence of θ , as much as possible, we should have

$$t\theta^2 \leq \frac{\lambda}{k}, \tag{1.11}$$

where λ/k is the maximum allowed error due to variations in θ . Typically, to have a reasonably small error, we at least require that

$$t\theta^2 \leq \frac{\lambda}{4}. \tag{1.12}$$

TABLE 1.1. Nature of Newton fringes for different surfaces with reference to a standard flat.

S. No.	Surface type	Appearance of the Newton fringes	
		Without tilt	With tilt
1	Plane		
2	Almost plane		
3	Spherical		
4	Conical		
5	Cylindrical		
6	Astigmatic (curvatures of same sign)		
7	Astigmatic (curvatures of opposite sign)		
8	Highly irregular		

Thus, to ensure a small error, both t and θ should be small. Regarding the value of t , we may safely assume that the value of t should never exceed a few wavelengths in the gap. If the surfaces are clean, then flat t should not exceed about 6λ . With this maximum value of t , the maximum allowed value of θ is such that $\theta^2 \leq 1/24$ or $\theta \leq 0.2$. For example, let us set the accuracy, to which the thickness t is to be assessed, to be equal to $\lambda/20$, thus, writing Eq. (1.12) as

$$t\theta^2 \leq \frac{\lambda}{20} \text{ or } 2\theta \leq 0.2. \quad (1.13)$$

From the foregoing analysis, it can be seen that the illumination angle on the two flats in contact should never exceed 0.2 rad or 12° approximately.

The size of the light source becomes irrelevant if the angular diameter of the entrance pupil of the observer, as seen from the flats, is smaller than this value. The light source can thus be extended to any size. It is only necessary that the observation, visual or photographic, is made nearly perpendicular to the flats and from a minimum distance, such that it is roughly five times the diameter of the optical flats in contact. To obtain higher accuracy, the distance from which the observation is made has to be larger. Alternatively, a collimating lens can be used and the entrance pupil of the observing eye or camera is then placed at the focus of the collimator.

If the observing distance is not large enough, equal thickness fringes will not be observed. Instead, localized fringes will appear. These fringes are called localized because they seem to be located either above or below the air gap. The fringes are localized in the region where corresponding rays from the two virtual images of the light source intersect each other. It has been shown that this condition may be derived from the van Cittert–Zernike theorem (Wyant, 1978; Simon and Comatri, 1987; Hariharan and Steel, 1989).

1.2.2. Some Suitable Light Sources

For setting up a Newton interferometer, we require a suitable monochromatic source. Several sources are available and are convenient. One source is, of course, a sodium vapor lamp, which does not require any filter. Another source is a low-pressure mercury vapor lamp with a glass envelope to absorb the ultraviolet light. A third possible source is a helium discharge lamp in the form of a zigzag discharge tube and with a ground glass to diffuse the light. Table 1.2 gives the various wavelengths that can be

TABLE 1.2. Characteristics, such as wavelength, of various lamps suitable as light sources in Newton’s interferometer.

Serial number	Lamp type	Wavelength(s) normally used (nm)	Remarks
1	Sodium vapor	589.3	The wavelength is the average of the doublet 589.0 and 589.6 nm. Warm-up time is about 10 min.
2	Low-pressure mercury vapor	546.1	Because of other wavelengths of mercury vapor present, the fringes must be viewed through the green filter, isolating the 546.1 nm line. There is no warm-up time. Tube lights without fluorescent coating can be used.
3	Low-pressure helium discharge	587.6	Because of other wavelengths of helium discharge present, a yellow filter must be used to view the fringes. There is no warm-up time.
4	Thallium vapor	535.0	Characteristics are similar to those of the sodium vapor lamp. Warm-up time is about 10 min.
5	Cadmium vapor	643.8	Red filter to view the fringes is required. Warm-up time is about 10 min.

used in these different spectral lamps. Even an ordinary fluorescent lamp with a plastic or glass green filter in front of the lamp works, but the fringe visibility is not high.

1.2.3. Materials for the Optical Flats

The optical flats are generally made of glass, fused silica, or more recently developed zero expansion materials such as CerVit and ULE glass. Small optical flats of less than 5 cm in diameter can be made of glass; they reach homogeneous temperature conditions reasonably quickly after some handling. It is preferable to make optical flats of larger sizes from fused silica or zero expansion materials. Table 1.3 gives relevant information regarding the materials commonly used for making optical flats.

When making a reference optical flat, it is necessary to consider carefully not only the material to be used but also the weight, size, testing methods, and many other important parameters (Primak 1984, 1989a, 1989b; Schulz and Schwider 1987).

1.2.4. Simple Procedure for Estimating Peak Error

Generally, optical surfaces are made to an accuracy ranging from a peak error of 2λ on the lower accuracy side to $\lambda/100$ on the higher side. It is possible by means of the

TABLE 1.3. Materials used for making optical flats and their properties.

Serial number	Material	Coefficient of linear expansion (per °C)	Remarks
1	BK7, BSC	$75-80 \times 10^{-7}$	These are borosilicate glasses that can be obtained with a high degree of homogeneity.
2	Pyrex	$25-30 \times 10^{-7}$	This is also a borosilicate glass but has higher silica content. Several manufacturers make similar type of glass under different brand names. This is a good material for making general quality optical flats and test plates.
3	Fused silica or quartz	6×10^{-7}	This is generally the best quartz material for making optical flats. Different grades of the material are available, based mainly on the degree of homogeneity.
4	CerVit, Zerodur	$0-1 \times 10^{-7}$	This material and similar ones made by different companies under different trade names have practically zero expansion at normal ambient temperatures.
5	ULE fused silica	$0-1 \times 10^{-7}$	This is a mixture of silica with about 7% titania.

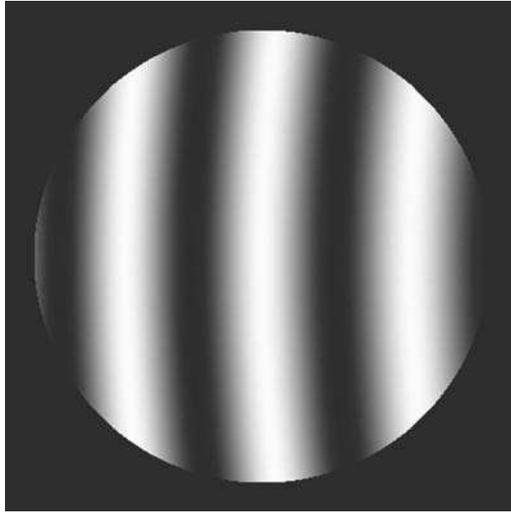


FIGURE 1.11. Newton fringes for an optical flat showing peak error of $\lambda/20$.

Newton interferometer to estimate peak errors up to about $\lambda/10$ by visual observation alone. Beyond that, it is advisable to obtain a photograph of the fringe system and to make measurements on this photograph. Figure 1.11 shows a typical interferogram as viewed in a Newton interferometer. Here, we have a peak error much less than $\lambda/4$. Consequently, the top plate is tilted slightly to obtain the almost straight fringes. The central diametral fringe is observed against a straight reference line such as the reference grid kept in the Newton interferometer in Figure 1.2. By means of this grid of straight lines, it is possible to estimate the deviation of the fringe from its straightness and also from the fringe spacing. The optical path difference is $2t$, so the separation between two consecutive fringes implies a change in the value of t equal to $\lambda/2$. Thus, if the maximum fringe deviation from the straightness of the fringes is d/k with d being the fringe separation, the peak error is given by

$$\text{Peak error} = \left(\frac{k}{d}\right) \left(\frac{\lambda}{2}\right) \quad (1.14)$$

In Figure 1.11 $k = 2.5$ mm and $d = 25$ mm; hence, we can say that the peak error is $\lambda/20$. Even in this case, it is desirable to know whether the surface is convex or concave, and for this purpose we can use the procedure described earlier. The only difference is that we have to imagine the center of the fringe system to be outside the aperture of the two flats in contact.

1.2.5. Measurement of Spherical Surfaces

Probably one of the most common applications of the Newton interferometer is the testing of the faces of small lenses while they are being polished. A small test plate

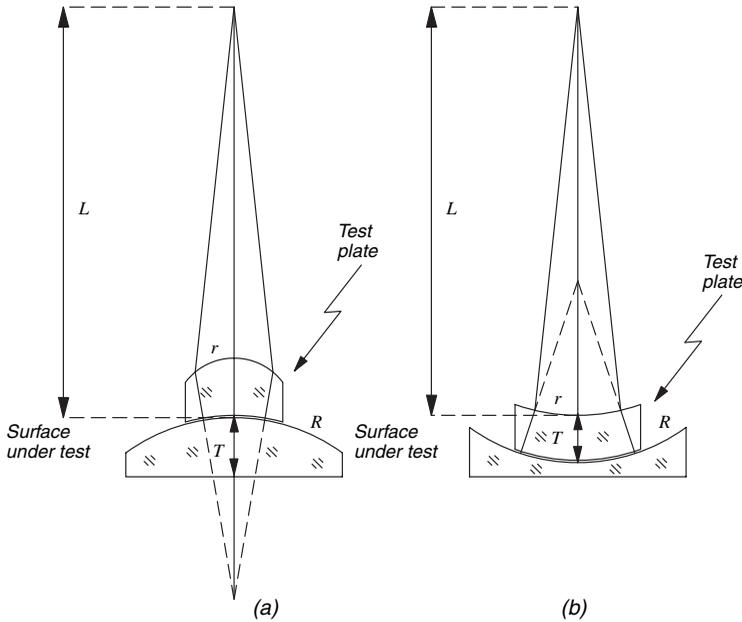


FIGURE 1.12. Test plates to test spherical surfaces with Newton fringes.

with the opposite radius of curvature is made according to the required accuracy and then placed over the surface under test. A test plate is useful not only to detect surface irregularities but also to check the deviation of the radius of curvature from the desired value (Karow 1979).

The observation should be made in such a way that the light is reflected almost perpendicular to the interferometer surfaces. Convex surfaces can be tested with the test plate shown in Figure 1.12(a), with a radius of curvature r in the upper surface given by

$$r = \frac{(N - 1)(R + T)L}{NL + R + T}, \quad (1.15)$$

where N is the refractive index of the test plate glass. Concave surfaces can be tested as in Figure 1.12 (b). In this case, the radius of curvature r of the upper surface is

$$r = \frac{(N - 1)(R - T)L}{NL - R + T} \quad (1.16)$$

It is important to remember that the fringes are localized very near to the interferometer surfaces, and therefore the eye should be focused at that plane.

The radius of curvature is checked by counting the number of circular fringes. The relation between the deviation in the radius of curvature and the number of rings can be derived with the help of Figure 1.13, where it can be shown that the distance

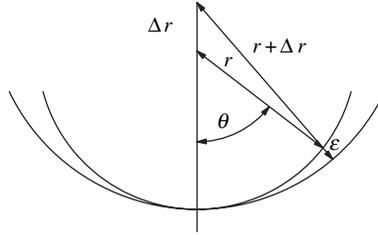


FIGURE 1.13. Geometry to find the separation between two spherical surfaces with different radii of curvature measured along the radius of one of them.

ε between the two surfaces, measured perpendicularly to one of the surfaces, is given by

$$\varepsilon = (r + \Delta r) \left\{ 1 - \left[1 - \frac{2(1 - \cos \theta)r\Delta r}{(r + \Delta r)^2} \right]^{1/2} \right\}. \quad (1.17)$$

If either Δr or the angle θ is small, this expression may be accurately represented by

$$\varepsilon = (1 - \cos \theta)\Delta r. \quad (1.18)$$

Since the number of fringes n is given by $n = 2\varepsilon/\lambda$, we can also write

$$\frac{n}{\Delta r} = \frac{2(1 - \cos \theta)}{\lambda} \quad (1.19)$$

If D is the diameter of the surface, the angle θ is defined as $\sin \theta = D/2r$. Therefore, a relation can be established between the increment per ring in the radius of curvature and the surface ratio r/D , as shown in Table 1.4.

1.2.6. Measurement of Aspheric Surfaces

Malacara and Cornejo (1970) used the method of Newton fringes to determine the aspheric profile of a surface that deviates markedly from a spherical surface. This method is useful if the aspheric deviates from the nearest spherical by a few wavelengths of light (say, $10\text{--}20\lambda$). The method consists in using a spherical test plate in contact with the aspherical surface and finding the position of the fringes by means of a measuring microscope. From these position values, one can then obtain the actual air gap as a function of the distance, and a plot can be made and compared with the required aspheric plot. Figure 1.14 shows a typical schematic arrangement for this method.

It is important to consider that the surface under test probably does not have rotational symmetry. Therefore, the measurements must be made along several diameters in order to obtain the complete information about the whole surface.

TABLE 1.4. Radius of curvature increment per fringe for several values of the power ratio r/D of the spherical surface being tested with newton fringes.

r/D	$\Delta r/n$ (cm)
1.0	0.00020
2.0	0.00086
3.0	0.00195
4.0	0.00348
5.0	0.00545
6.0	0.00785
7.0	0.01069
8.0	0.01397
9.0	0.01768
10.0	0.02183
20.0	0.08736
30.0	0.19661
40.0	0.34970
50.0	0.54666
60.0	0.78712
70.0	1.07033
80.0	1.39665
90.0	1.77559
100.0	2.18144

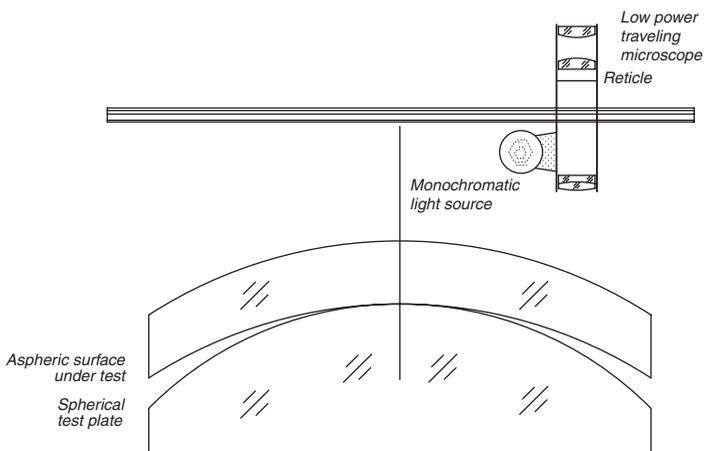


FIGURE 1.14. Schematic arrangement showing the method of measuring aspheric surfaces with a spherical test plate using Newton fringes.

Instead of directly measuring the fringe positions with a microscope, a photograph can be taken, and then the fringe positions can be measured with more conventional procedures.

If the reference surface is spherical and the surface under test is aspherical (hyperboloid or paraboloid), the ideal fringe patterns will be those of a Twyman–Green interferometer for spherical aberration as described in Chapter 2.

The reference surface may also be another aspherical surface that exactly matches the ideal configuration of the surface under test. This procedure is useful when a convex aspheric is to be made, since a concave aspheric can be made and tested more easily than a convex surface. The advantage of this method is that a null test is obtained. It has the disadvantage that the relative centering of the surfaces is very critical because both surfaces have well-defined axes, and these must coincide while testing. This problem is not serious, however, because the centering can be achieved with some experience and with some device that permits careful adjustment.

When mathematically interpreting the interferograms, it should be remembered that the OPD is measured perpendicularly to the surfaces, whereas the surface sagitta z is given along the optical axis. Therefore the OPD is given by $2(z_1 - z_2) \cos \theta$, where $\sin \theta = Sc$.

1.2.7. Measurement of Flatness of Opaque Surfaces

Sometimes we encounter plane surfaces generated on such metal substrates as steel, brass, and copper. An optical flat made of glass should be put on top of such objects for viewing Newton fringes. It is not always the case that the metal object is in the form of a parallel plate. The plane surface may be generated on an otherwise irregular component, and hence some means of holding the component while testing becomes necessary. This can be avoided if we can put the object on top of the optical flat and observe the fringes through the bottom side of the flat. This sort of arrangement is shown in Figure 1.15. Since most metal surfaces have reflectivities that are quite high compared to the value for a glass surface, the contrast of the fringes is not very good. To improve this situation, the optical flat is coated with a thin evaporated film of chromium or inconel having a reflectivity of about 30–40%. This brings about the formation of sharper, more visible fringes.

It is necessary to point out that if the object is very heavy, it will bend the optical flat and the measurement will not be accurate. Therefore, this kind of arrangement is suitable for testing only small, light opaque objects. In dealing with heavy objects, it is preferable to place the optical flat on top of the object.

1.3. FIZEAU INTERFEROMETER

In the Newton interferometer, the air gap between the surfaces is very small, and of the order of a few wavelengths of light. Sometimes it is convenient to obtain fringes similar to the ones obtained in the Newton interferometer, but with a much larger air gap. When the air gap is larger, the surfaces need not be cleaned as thoroughly as they

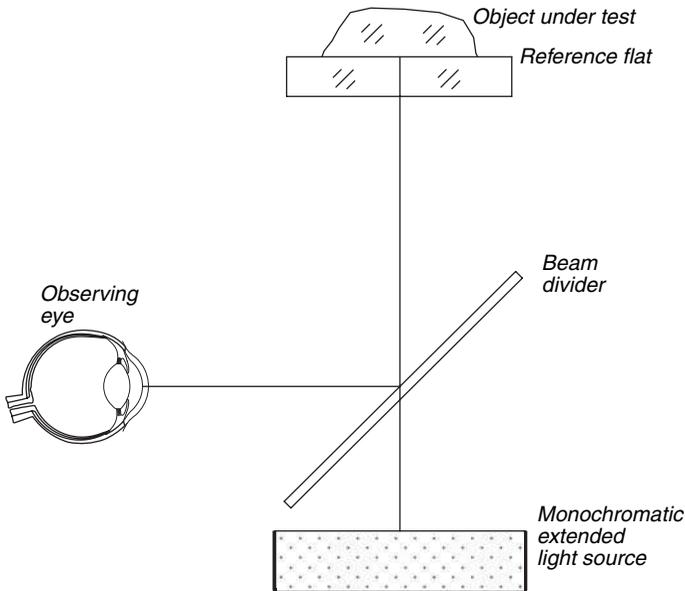


FIGURE 1.15. Schematic arrangement showing the method of testing opaque plane surfaces on irregular objects by placing them on top of the optical flat.

must be before being tested in the Newton interferometer. Also, due to the larger gap, the requirements for the collimation and size of the light source become stronger. This is called a Fizeau interferometer.

The Fizeau interferometer is one of the most popular instruments for testing optical elements. Some of its main applications will be described here, but the basic configurations used for most typical optical elements are identical to the ones used with the Twyman-Green interferometer to be described in Chapter 2. The reader is referred to that chapter for more details.

1.3.1. The Basic Fizeau Interferometer

From the foregoing considerations, it is seen that we should have a collimating system and a smaller light source in a Fizeau interferometer. Figure 1.16 shows the schematic arrangement of a Fizeau interferometer using a lens for collimation. The optical flat that serves as the reference is generally mounted along with the lens and is preadjusted so that the image of the pinhole reflected by the reference surface falls on the pinhole itself. Either the back side of the flat is antireflection coated or (more conveniently) the reference optical flat is made in the form of a wedge (about 10–20 min of arc) so that the reflection from the back surface can be isolated. To view the fringes, a beam divider is located close to the pinhole. The surface under test is kept below the reference flat, and the air gap is adjusted to the smallest value possible; then the air wedge is gradually reduced by manipulating the flat under test. When the air wedge is very large, two distinct images of the pinhole by the two surfaces can be

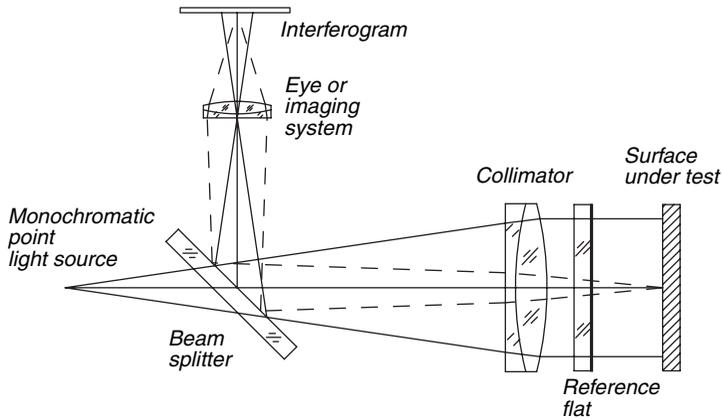


FIGURE 1.16. Schematic arrangement of a Fizeau interferometer using a lens for collimation of light.

seen in the plane P in Figure 1.16. By making use of screws provided to tilt the flat under test, one can observe the movement of the image of the pinhole and can stop when it coincides with that of the reference flat. Then the observer places his eye at the plane P and sees, localized at the air gap, the fringes due to variation in the air gap thickness. Further adjustment, while looking at the fringes, can be made to alter the number and direction of the fringes. The interpretation of these fringes is exactly the same as that for Newton and Twyman-Green fringes.

Figure 1.17 is a schematic of a Fizeau interferometer using a concave mirror as the collimating element. If a long focal length is chosen for the concave mirror, a spherical mirror can be used. For shorter focal lengths, an off-axis paraboloidal

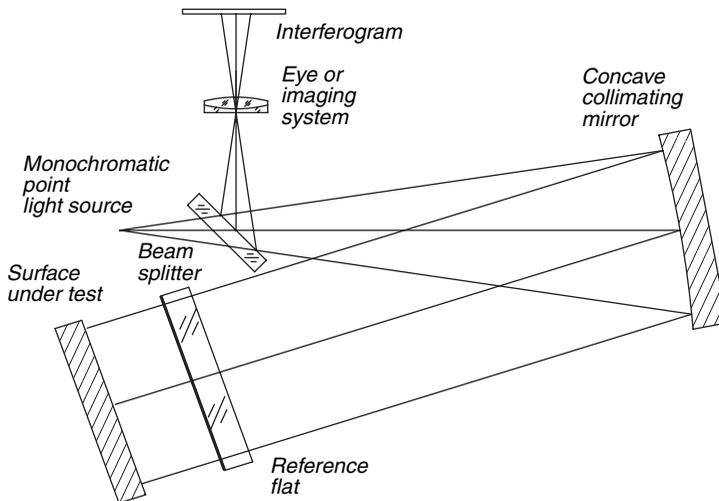


FIGURE 1.17. Schematic arrangement of a Fizeau interferometer using a concave mirror for collimation of the illuminating beam.

mirror may be required. Both the schemes of Figures 1.16 and 1.17 may be arranged in either a vertical (upright and inverted) or a horizontal layout. In the vertical situation the optical flats are horizontal, whereas in the horizontal layout the optical flats stand on their edges.

If the optical system or element under test has a high reflectivity and the reference flat is not coated, then the two interfering beams will have quite different intensities, and thus the fringes will have a poor contrast. On the contrary, if the reference flat is coated with a high reflectivity, but smaller than 100% to allow some light to be transmitted, a confusing system of fringes will appear because of multiple reflections. Commonly, to obtain two-beam interference fringes effectively, the reference surface must be uncoated. Then, to match the intensities, either the reflectivity of the optical element under test also has to be low or the amplitude of the beam under test has to be attenuated. The fact that the two surfaces reflecting the interfering beams have a low reflectivity makes it very important to take all necessary precautions to avoid spurious reflections at some other surfaces, mainly when a laser light source is used.

1.3.2. Coherence Requirements for the Light Source

As in the Newton interferometer, in the Fizeau interferometer the maximum allowed angular size of the light source to be used depends on the length of the air gap. For instance, if the air gap between the flats is 5 mm, and taking $\lambda = 5 \times 10^{-4}$ mm, the permissible value of 2θ given by Eq. (1.12) is 0.01 rad. Such a small angle can be obtained by using a collimator with the entrance pupil of the observer located at the focus, to observe the angle almost perpendicularly to the air gap for all points of the observed flats. Also, either the pupil of the observer or the light source has to be extremely small. Frequently the pupil of the eye has a diameter larger than required, so that it is simpler to have a light source with a pinhole. The larger the air gap is, the smaller the pinhole has to be.

When plane surfaces are tested in the Fizeau interferometer the air gap can be made quite small if desired. The total optical path difference involved does not exceed a few millimeters. Thus, a small low-pressure mercury vapor lamp can be used with a green filter as the source of light. When testing for the wedge of thick plates of glass, the OPD is larger due to the thickness. For gas or metal vapor lamp, this OPD is about the maximum we can use. For plates of greater thickness, the contrast of the interference fringes is greatly reduced because the lamp does not give a very sharp spectral line with a large temporal coherence. Similarly, the same situation of low contrast occurs when thick glass shells are tested or when spherical test plates are tested with one test plate.

This limitation can be eliminated, however, if we can use a source of very high monochromaticity. Fortunately, such a source, the laser, has recently become available. For our application, the low-power (2 mW) helium–neon gas laser operating in a single mode TEM₀₀ and with a wavelength of emission at 632.8 nm is ideal. With this as the source of light, we can tolerate an OPD of at least 2 m and obtain Fizeau fringes of high contrast. Even larger OPDs are possible provided that a properly stabilized laser is chosen and vibration isolation is provided for the instrument.

Most of the coherence requirements for Fizeau interferometers are similar to the requirements for Twyman–Green interferometers as described in section 2.3. There, it is pointed out that a gas laser has perfect spatial coherence, and can have almost perfect temporal coherence and thus we might think that this is the ideal light source for interferometry, but this is not always the case. The reason is that many unwanted reflections from other surfaces in the optical system may produce a lot of spurious fringes that can appear. Also, the laser light produces scattering waves from many small pieces of dust or scratches in the optical elements. To solve this problem, the light source can be extended even when using a laser by introducing a thin rotatory half ground glass close to the point light source. Deck et al. (2000) have proposed an annular shape for the light source by using a diffracting element to produce a small cone of light illuminating the rotating ground glass. The effect of the spurious reflections has been studied by several researchers, for example by Ai and Wyant (1988 and 1993) and by Novak and Wyant (1997).

Another possible effect to be taken into account is that some optical elements or systems to be tested may be retroreflectors, either in one dimension like a porro prism, or in two dimensions like a cube corner prism. The retroreflection has associated an inversion, reversal, or both (which is equivalent to a 180° rotation) of the wavefront. A point of view is that then the interference takes place between two different points on the wavefront, symmetrically placed with respect to the optical axis if the wavefront was rotated, or symmetrically placed with respect to the inversion or reversion axis. The fringes will have a good contrast only if the spatial coherence of the wavefront is high enough. This condition imposes a stronger requirement on the small size of the point light source.

Another equivalent explanation for this retroreflection effect is illustrated in Figure 1.18. Let us consider the pinhole on the light source to have a small finite

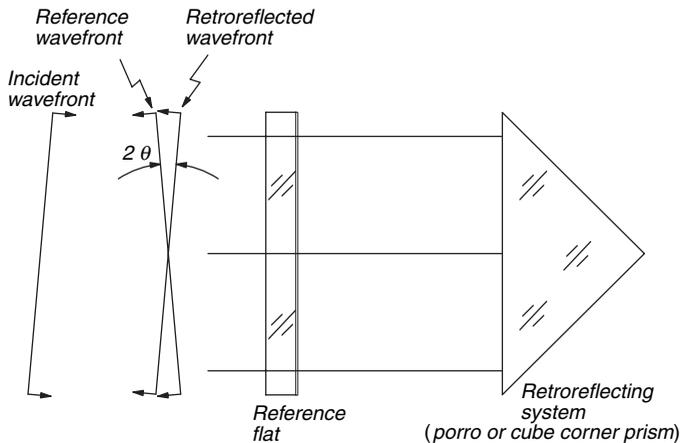


FIGURE 1.18. Interference between the reference wavefront and the wavefront retroreflected by a porro prism under test. Both wavefronts originate at one point on the edge of the small light source. The angle between these two wavefronts reduces the contrast of the fringes.

size and a flat collimated incidence wavefront coming from the edge of that pinhole at a small angle θ . It is easy to see that the two interfering wavefronts will not be parallel to each other, but will make an angle 2θ between them. Of course, there are infinite number of wavefronts coming from different points at the pinhole of the light source, all with different orientations and angles, smaller than θ . This multiplicity of wavefronts with different angles will reduce the contrast of the fringes from a maximum at the center where all the wavefronts intersect, decreasing towards the edge of the pupil. This effect is also present for the same reason in the Twyman–Green interferometer as described in Chapter 2 in more detail.

The strong spatial coherence requirements when a retroreflecting system is tested is difficult to satisfy with gas or metal vapor lamps, but with gas lasers it is always fulfilled.

1.3.3. Quality of Collimation Lens Required

We shall briefly examine the quality of collimating lens required for the Fizeau interferometer. Basically, we are interested in determining the variation in air gap thickness. However, the OPD is a function of not only the air gap thickness but also the angle of illumination, and at a particular point this is $2t \cos \theta$. The air gap t varies because of the surface defects of the flats under test, while the variation of θ is due to the finite size of the source and the aberration of the collimating lens.

For Fizeau interferometers using conventional sources of light, the maximum air gap that is useful is 50 mm. Also, in this case we have to consider the size of the source and the aberration of the lens separately. The effect of the size of the source is mainly on the visibility of the Fizeau fringes. The excess optical path difference $t\theta^2$ should be less than $\lambda/4$ for good contrast of the Fizeau fringes and the pinhole is chosen to satisfy this condition. The effect of the pinhole is uniform over the entire area of the Fizeau fringes. On the contrary, the effect of aberration in the collimating lens is not uniform. Thus, we have to consider the angular aberration of the lens and its effect. If ϕ is the maximum angular aberration of the lens, then $t\phi^2$ should be less than $k\lambda$, where k is a small fraction that depends on the accuracy required in the instrument. Thus, let us set $k = 0.001$, so that the contribution of $t\phi^2$ is 0.001λ . Taking a maximum value of $t = 50$ mm for the ordinary source situation, we have

$$\phi^2 \leq \frac{0.001\lambda}{t} \approx 10^{-8},$$

or

$$\phi = 10^{-4} \text{ rad.} \quad (1.20)$$

This angular aberration is quite large, being of the order of 20 s of arc. Hence, suitable lenses or mirror systems can be designed for the purpose (Taylor, 1957; Yoder, 1957; Murty and Shukla, 1970).

1.3.4. Liquid Reference Flats

It is well known that a liquid surface can be used as a reference flat. Basically the liquid surface has a radius of curvature equal to that of the earth. If the radius of the earth is taken as 6400 km, the sag of the surface is (Grigor'ev et al., 1986; Ketelsen and Anderson 1988)

$$\frac{y^2}{2R} = \frac{y^2}{2 \times 6.4 \times 10^9} \text{ mm} \quad (1.21)$$

where $2y$ is the diameter of the liquid surface considered. If we stipulate that this should not exceed $\lambda/100$ ($\lambda = 5 \times 10^{-4}$), then

$$y^2 \leq 6.4 \times 10^4$$

or

$$2y \leq 512 \text{ mm} \quad (1.22)$$

Thus, a liquid surface of about 0.5 m diameter has a peak error of only $\lambda/100$ as compared to an ideal flat. Therefore, it has been a very attractive proposition to build liquid flats as standard references. In practice, however, there are many problems, mainly in isolating the disturbing influence of vibrations. It is also necessary to exclude the region near the wall of the vessel that holds the liquid and to make sure that no dust particles are settling down on the surface. Possible liquids that can be useful for the purpose are clear and viscous, such as glycerin, certain mineral oils, and bleached castor oil. Water is probably not suitable because of its low viscosity. Mercury may not be suitable because of its high reflectivity; the two interfering beams will have very unequal intensities, resulting in poor contrast of the fringes unless the surface under test is also suitably coated. However, mercury has been used as a true horizontal reference plane reflecting surface in certain surveying and astronomical instruments.

1.3.5. Fizeau Interferometer with Laser Source

We shall now describe a Fizeau interferometer using a source such as the helium-neon gas laser of about 2 mW power lasing at 632.8 nm in the single mode. A schematic diagram is shown in Figure 1.19. A very well corrected objective serves to collimate the light from the pinhole, illuminated by a combination of the laser and a microscope objective. Between the collimating objective and the pinhole (spatial filter), a beam divider is placed so that the fringes can be observed from the side. It is also desirable to provide a screen, upon which the Fizeau fringes are projected, to avoid looking into the instrument as is normally done when conventional light sources are used. The laser has a high radiance compared to other sources, and a direct view may be dangerous to the eye under some circumstances. The reference

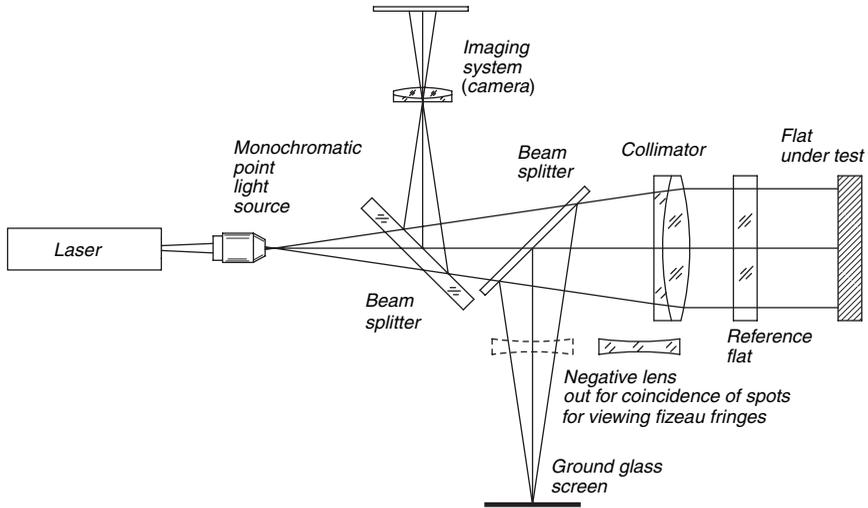


FIGURE 1.19. Schematic arrangement of a Fizeau interferometer using a laser source. The scheme shown here is for plane surfaces. The system is easily aligned with the help of a sliding negative lens.

plane surface is permanently adjusted so that the reflected image of the pinhole is autocollimated. The surface under test is adjusted until the image reflected from it also comes into coincidence with the pinhole. To facilitate preliminary adjustment, the screen is used to project the two pinhole images from the two reflecting plane surfaces. This is accomplished by removing the negative lens between the beam divider and the ground glass screen. The pinhole image from the reference surface is at the center of the screen, whereas the one from the surface under test is somewhere on the screen; by manipulation of this surface, the two spots of light on the screen can be brought into coincidence. Then the negative lens is inserted in the path, and the Fizeau fringes are projected on the screen. These fringes can be further adjusted in direction and number as required. By the use of another beam divider, it is possible to divert part of the beam to a camera for taking a photograph of the fringe pattern. The whole instrument must be mounted on a suitable vibration-isolated platform.

This instrument can be used for various other applications that are normally not possible with conventional sources of light. We describe some such applications in the sections that follow. In addition, many possibilities exist for other applications depending on the particular situations involved.

Several commercial Fizeau interferometers have been available for several years, but probably the two most widely known are the Zygo interferometer (Forman, 1979), shown in Figure 1.20, and the Wyko interferometer, shown in Figure 1.21.

1.3.6. Multiple-Beam Fizeau Setup

If, instead of two-beam fringes, multiple-beam fringes of very good sharpness are required, the reference optical flat and the optical flat under test are coated with a

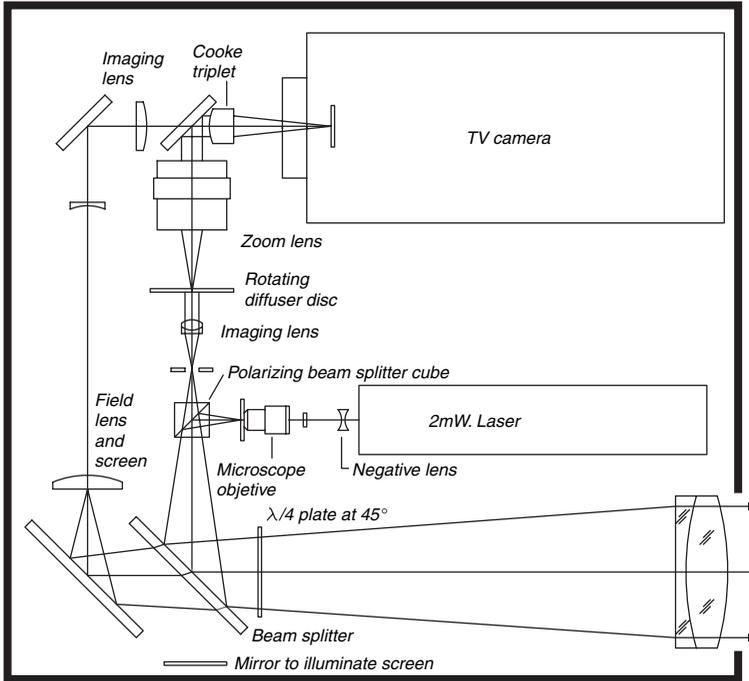


FIGURE 1.20. Fizeau interferometer manufactured by Zygo Corp. (Courtesy of Zygo Corp.).

reflecting material of about 80–90% reflectivity (see Chapter 6) such as aluminum or silver. If higher reflectivities are required, multilayer dielectric coatings can be applied. In fact, the instrument may be provided with several reference flats having coatings of different reflectivities.

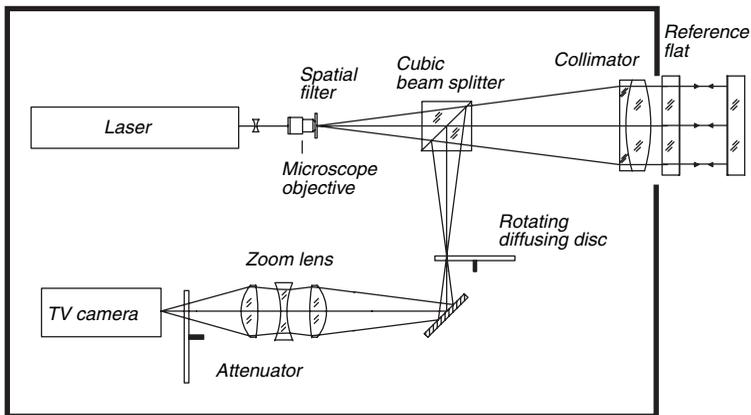


FIGURE 1.21. Fizeau interferometer manufactured by Wyko Corp. (Courtesy of Wyko Corp.).

1.3.7. Testing Nearly Parallel Plates

In many applications, glass plates having surfaces that are both plane and parallel are required. In such cases, the small wedge angle of the plate can be determined by the Fizeau interferometer, and the reference flat of the interferometer need not be used since the fringes are formed between the surfaces of the plate being tested. If α is the angle of the wedge and N is the refractive index of the glass, the angle between the front- and back-reflected wavefronts is given by $2n\alpha$, and hence the fringes can be expressed as

$$2N\alpha = \frac{\lambda}{d}, \quad (1.23)$$

where d is the distance between two consecutive bright or dark fringes. Hence the angle α is given by

$$\alpha = \frac{\lambda}{2nd}. \quad (1.24)$$

To determine the thinner side of the wedge, a simple method is to touch the plate with a hot rod or even with a finger. Because of the slight local expansion, the thickness of the plate increases slightly. Hence a straight fringe passing through the region will form a kink pointing toward the thin side, as shown in Figure 1.22. For instance, if we take $N = 1.5$, $\lambda = 5 \times 10^{-4}$ mm, and $\alpha = 5 \times 10^{-6}$ (1 s of arc), we get for d a value of about 33 mm. Hence a plate of 33 mm diameter, showing one fringe, has a wedge angle of 1 s of arc. If the plate also has some surface errors, we

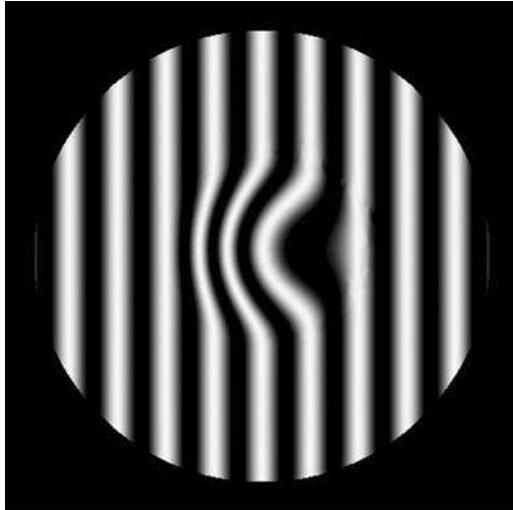


FIGURE 1.22. Kink formation in the straight Fizeau fringes of a slightly wedged plate, obtained by locally heating the plate. The kink is pointing toward the thin side of the wedge.

get curved fringes, indicating both surface and wedge errors. If the surfaces are independently tested and found to be flat, and even in this situation one is getting curved fringes, these should be attributed to variation of the refractive index inside the plate in an irregular manner. In fact, by combining the tests on the Newton interferometer and the Fizeau interferometer for a parallel plate, it is possible to evaluate the refractive index variation (inhomogeneity) (Murty, 1963; Murty, 1964a; Forman, 1964).

1.3.8. Testing the Inhomogeneity of Large Glass or Fused Quartz Samples

The sample is made in the form of a parallel plate. The surfaces should be made as flat as possible with a peak error of not more than λ . Then the plate is sandwiched between two well-made parallel plates of glass with a suitable oil matching the refractive index of the sample. This will make the small surface errors of the sample negligible, and only straight fringe deformation due to the inhomogeneity of the sample will be seen. If the sandwich is kept in the cavity formed by the two coated mirrors, very sharp dark fringes on a bright background are obtained. If, for instance, the maximum fringe deviation from straightness is k and the distance between two fringes is d , the optical path difference is $(k/d)\lambda$. Now the OPD due to the inhomogeneity ΔN and thickness t of the sample is given by $2\Delta N \cdot t$, and hence

$$\Delta N = \left(\frac{k}{d}\right) \left(\frac{\lambda}{2t}\right) \quad (1.25)$$

As an example, if $k/d = 0.25$, $\lambda = 632.8 \text{ nm}$, and $t = 50 \text{ mm}$, we have $\Delta N = 1.6 \times 10^{-6}$. Thus a maximum variation of 1.6×10^{-6} may be expected in the sample for the direction in which it has been tested. Figure 1.23 shows the schematic arrangement of the Fizeau interferometer for the method just described.

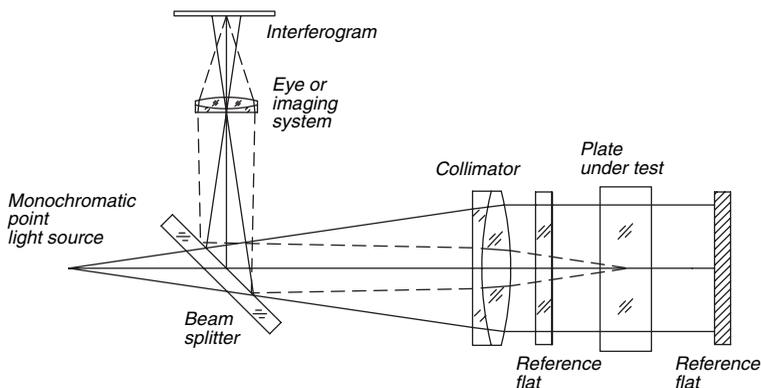


FIGURE 1.23. Schematic arrangement of a Fizeau interferometer for testing the homogeneity of solid samples of glass, fused quartz, and so on.

1.3.9. Testing the Parallelism and Flatness of the Faces of Rods, Bars, and Plates

Frequently, the need for testing the parallelism and the flatness of two opposite faces in a rod, plate, or bar arises. If the plate to be tested is transparent and has a highly homogeneous refractive index, the problem is not so complicated. If the refractive index of the material is inhomogeneous or if it is not transparent, special techniques have to be developed.

Vannoni and Molezini (2004) described a configuration for this purpose, as illustrated in Figure 1.24. The first step is to adjust the interferometer to produce the minimum number of fringes without the plate or rod to be tested. The field of view will show the fringes due to any possible defect in the right angle prism. Then the plate is inserted as shown in the figure.

1.3.10. Testing Cube Corner and Right-Angle Prisms

In their retro-reflective configuration, if the right angles of cube corner and right-angle prisms are exact without any error, they reflect an incident plane wavefront as a single emerging plane wavefront. Otherwise the reflected wavefront consists of several plane wavefronts with different tilts, making possible the measurement of the prism errors. Because of the total internal reflection, the intensity of reflected light from these prisms is very high, nearly 100%. Since the reference flat is not coated, the fringes will have a poor contrast. To optimize the fringe contrast, either the reflectivity of the optical element under test also has to be low or the amplitude of the beam under test has to be attenuated.

To reduce the effective reflectivity of the right angle or cube prism, we can introduce a parallel plate of glass coated with a metallic film having a transmission between 20% and 30%. In this case the intensities of the two beams matched reasonably well, and we get a good contrast of two-beam fringes. The coated plate between the prism and the uncoated reference flat should be tilted

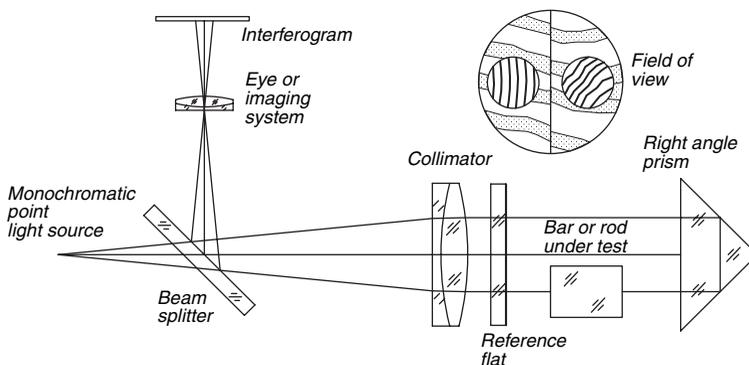


FIGURE 1.24. Schematic arrangement to test an opaque bar or rod for flatness and parallelism of the two opposite faces.

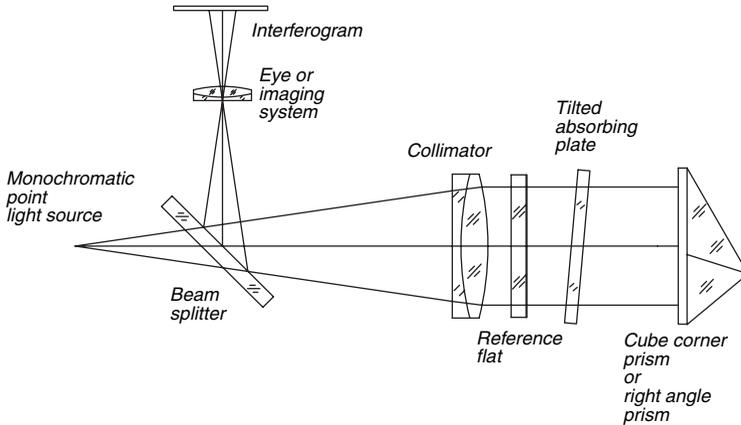


FIGURE 1.25. Schematic arrangement of a Fizeau interferometer for testing cube corner prisms and right-angle prisms. Here an absorbing plate is inserted between the prisms and the reference flat surface to equalize the intensities of the two interfering beams.

sufficiently to avoid the directly reflected beam. This method is shown schematically in Figure 1.25.

Another possible method is to reduce the reflectivity of one of the total reflecting surfaces. This can be done by constructing a special cell in which the prism is mounted, and behind one reflecting surface, a thin layer of water or some other suitable liquid is in contact with the surface. Thus, in effect, the refractive index difference is reduced at one total internal reflecting surface, and hence, the intensity of the wavefront reflected from the prism matches that of an uncoated flat. This method is shown schematically in Figure 1.26.

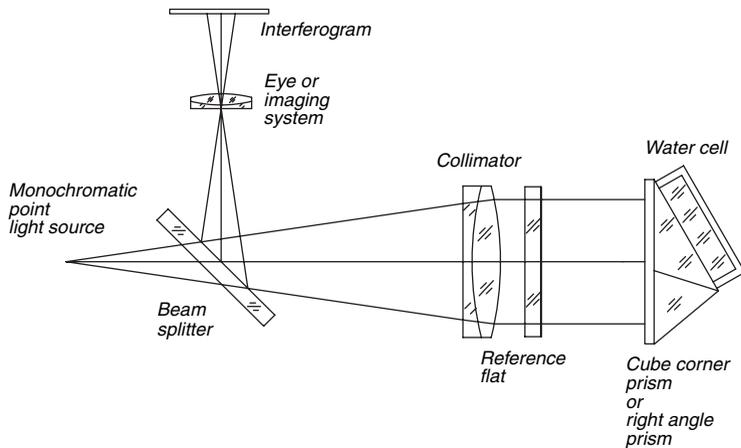


FIGURE 1.26. A scheme for reducing the intensity of reflected light from the corner cube prism and the right-angle prism. One of the total internally reflecting faces is brought into contact with water or some other liquid by the use of a cell behind it.

The interferograms obtained when testing of these prisms are identical to those in the Twyman–Green interferogram. For more details please see Chapter 2.

1.3.11. Fizeau Interferometer for Curved Surfaces

Just as collimated light is employed for testing optical flats on the Fizeau interferometer, it is possible to use either divergent or convergent light for testing curved surfaces. Figure 1.27 shows an arrangement for testing a concave surface against a reference convex surface. The point source of light is located at the center of the curvature of the convex reference surface. The concave surface under test is adjusted until its center of curvature, too, almost coincides with the point source of light. The procedure is exactly the same as before except that to achieve the uniform air gap, we have to provide some translational motion also (Moore and Slaymaker, 1980).

The same setup can be used very easily for checking the uniformity of the thickness (concentricity) of spherical shells. In this case the interfering beams are obtained from the front and back of the two spherical concentric surfaces. Figure 1.28 shows this setup for testing the concentricity of a spherical shell. If the radii of curvature are correct but the shell has a wedge (the centers of curvature are laterally displaced), we get straight fringes characteristic of the wedge. The hot rod or finger touch procedure described in Section 1.2.3 can be adopted to determine which side is thinner. If the two radii are not of proper value ($\sqrt{r_1 - r_2} \neq t$, where r_1 and r_2 are the two radii and t is the center thickness), the value of t is not constant over the entire shell. Hence, we get circular fringes like Newton fringes. If in addition a wedge is present, the center of these circular fringes will be decentered with respect to the center of the shell. In this situation also, we can adopt the hot rod or finger touch procedure to decide whether the shell is thin at the edge or at the center.

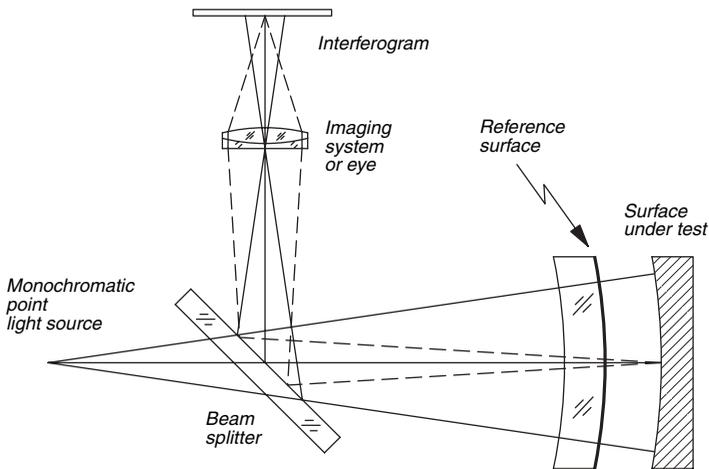


FIGURE 1.27. Fizeau interferometer set up for curved surfaces. Here the convex surface is the reference surface and the concave surface is under test.

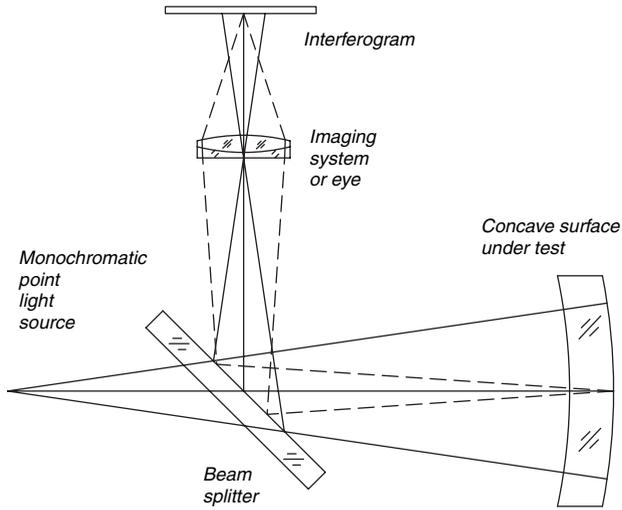


FIGURE 1.28. Fizeau interferometer setup for testing the concentricity of the spherical shell.

We can also have an arrangement for testing convex surfaces against a concave reference surface, as shown schematically in Figure 1.29. Here we use a lens or a group of lenses at finite conjugate distances such that the point source of light is at one conjugate, whereas the common center of curvature of the test surface and the reference surface is at the other conjugate. The concave reference surface is fixed to the instrument, while the convex surface under test is manipulated in the usual manner to obtain a uniform air gap.

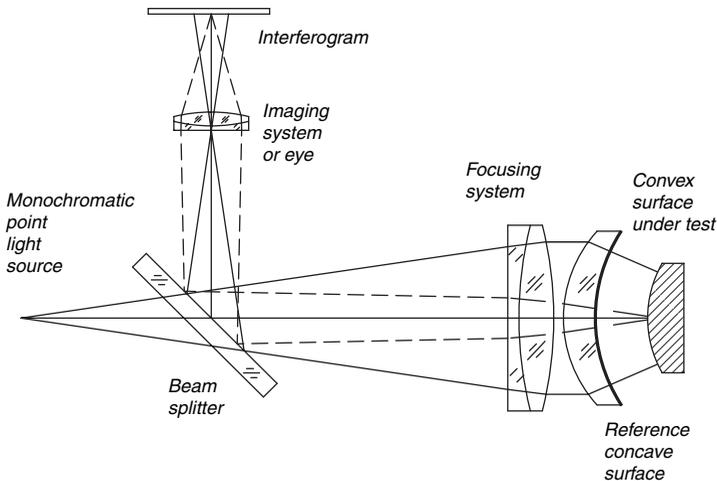


FIGURE 1.29. Fizeau interferometer setup for testing a convex surface against a concave reference surface.

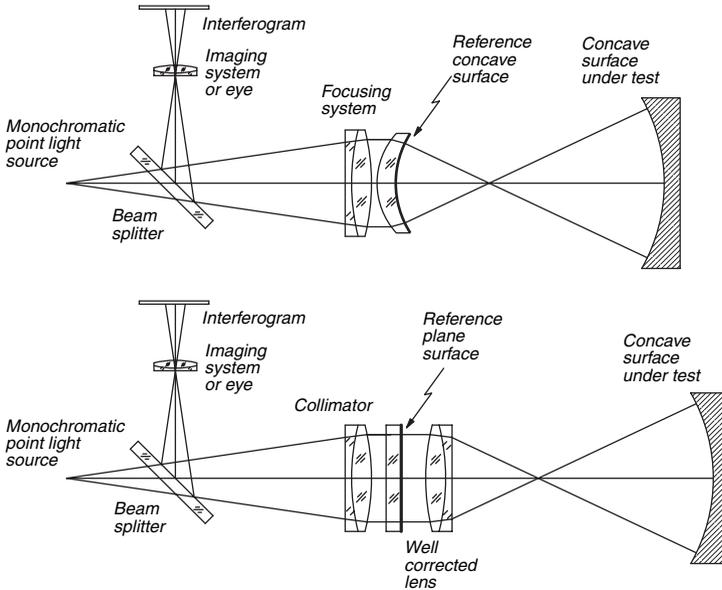


FIGURE 1.30. Schematic diagram of a Fizeau interferometer for testing a concave surface using a concave reference surface or a flat reference surface.

1.3.12. Testing Concave and Convex Surfaces

The reference surface is again the uncoated flat surface that is part of the Fizeau interferometer. The collimated light from the instrument, after passing through the optical flat, is again focused by the use of another highly corrected lens. If the surface is concave, it is set up as shown in Figure 1.30; if convex, as shown in Figure 1.31. When the surface is spherical and the center of curvature coincides with the focus of the lens, a plane wavefront is reflected back. Hence, we should obtain straight fringes due to the interference of the two beams. If the optical reference flat and the spherical surfaces are coated with high reflecting material, we can get very sharp, multiple-beam Fizeau fringes. If the surfaces are not spherical but are aspheric, appropriate null lenses must be used in the interferometer. This setup can also be used to measure the radius of curvature if a length-measuring arrangement is provided.

The testing of convex surfaces with the Fizeau interferometer presents many interesting problems, mainly if the surface is large and/or aspheric, which have been analyzed by several authors, for example by Burge (1995).

Another interferometer, which may be considered as a Fizeau interferometer, was devised by Shack (Shack and Hopkins, 1979; Smith, 1979). The difference is that this scheme uses a He-Ne laser source to give very large coherence length. Hence, the separation between the convex reference surface and the concave surface under test can be very large (typically several meters). Also, the convex reference surface

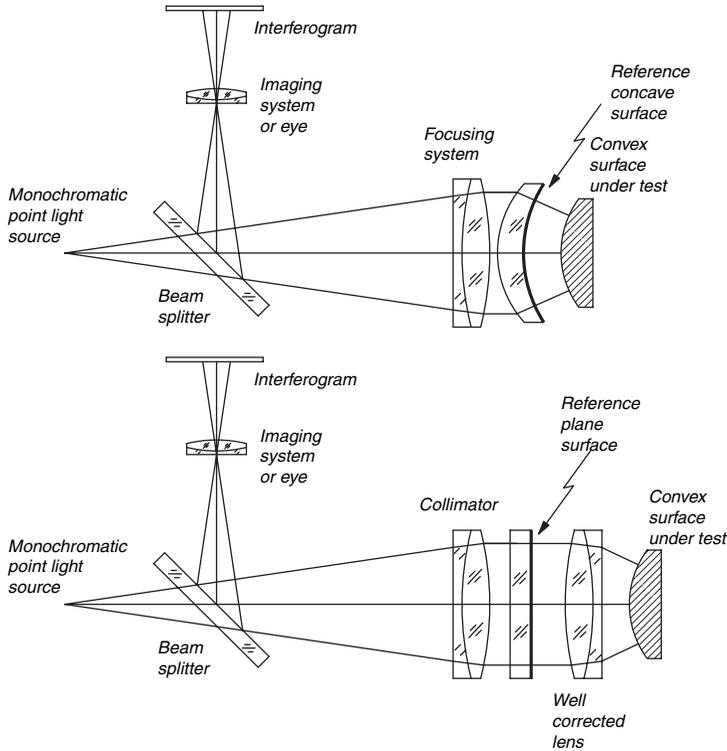


FIGURE 1.31. Schematic diagram of a Fizeau interferometer for testing a convex surface using a flat concave reference surface or a flat reference surface.

becomes a part of the instrument and can be of very short radius of curvature. The scheme, in fact, incorporates the device in the form of a beam-divider cube with one of the faces made into a convex spherical surface. The Shack interferometer is shown schematically in Figure 1.32. It is possible to test a large aspherical surface with this interferometer if a suitable null corrector is inserted between the interferometer and the surface under test.

1.4. HALDINGER INTERFEROMETER

With the Newton and Fizeau interferometers, we are basically interested in finding the variation in the air gap thickness. In these cases, the fringes are referred to as fringes of equal thickness. If, however, the thickness of the air gap is uniform and it is illuminated by a source of large angular size, we get what are called fringes of equal inclination. These fringes are formed at infinity, and a suitable lens can be used to focus them on its focal plane. If the parallel gap is that of air, we have the simple relation $2t \cos \theta = n\lambda$, as given in Eq. (1.9), from which we can easily see that, for a

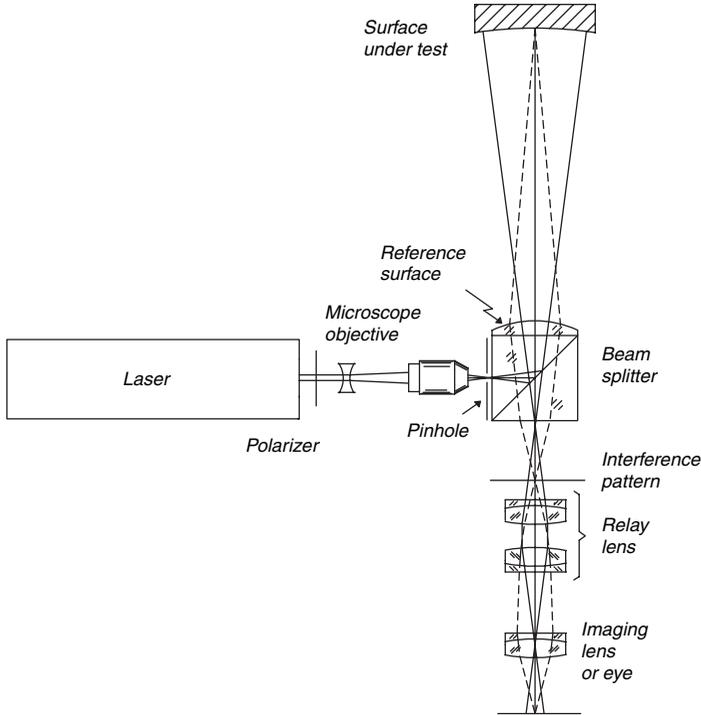


FIGURE 1.32. Schematics of a Shack–Fizeau interferometer.

constant value of t , we obtain fringes of equal inclination that are circles and are formed at infinity.

If the air gap is replaced by a solid plate such as a very good parallel plate of glass, Eq. (1.9) is modified slightly to include the effect of the refractive index N of the plate and becomes

$$2Nt \cos \theta' = n\lambda \quad (1.26)$$

where θ' is the angle of refraction inside the glass plate. For small values of θ' , we may approximate this expression as

$$2Nt = \left(\frac{t}{N}\right)\theta^2 = n\lambda \quad (1.27)$$

To see Haidinger fringes with simple equipment, the following method, illustrated in Figure 1.33 may be adopted. A parallel plate of glass is kept on a black paper and is illuminated by the diffuse light reflected from a white card at 45° . At the center of the white card is a small hole through which we look at the plate. With relaxed accommodation our eyes are essentially focused at infinity, i , and we see a system

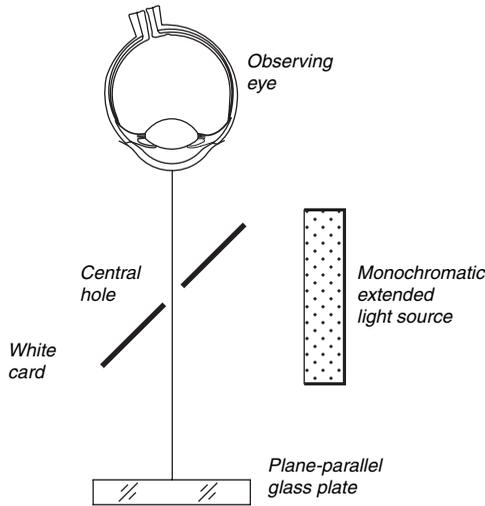


FIGURE 1.33. A simple arrangement to see Haidinger fringes for a nearly parallel glass plate.

of concentric circular fringes. For the light source we can use a sodium or even a fluorescent lamp.

In the situation where the laser is the source of light, there is a much higher limit for the value of t . Even though several meters can be used for t , we shall set $t = 1000$ mm. In this case, using $\lambda = 632.8$ nm, we get for ϕ an upper limit of 5 s of arc. Hence it is not difficult to design a collimating system to satisfy this condition.

Another aspect that is important, especially with large values of t , is the lateral shear one can get in the instrument. To avoid this, the autocollimated pinhole images must coincide with the pinhole itself. Similarly, if the collimating lens is not properly collimated, either a convergent or a divergent beam will emerge. The collimation may be accurately performed by using any of the various devices available, such as the plane parallel plate shearing interferometer (Murty, 1964b).

A somewhat better method is to use a lens for focusing the system of Haidinger fringes on its focal plane. This requires a setup almost identical to that for the Fizeau interferometer. The only difference is that, instead of a pinhole, a wider aperture is used to have a large angular size for the source. The Haidinger fringes are then formed in the focal plane of the lens.

1.4.1. Applications of Haidinger Fringes

The Haidinger fringes may be used as a complementary test to that provided by the Fizeau interferometer. If we are testing a nearly parallel plate, we can find its wedge angle either by the Fizeau or by the Haidinger method. In the Haidinger method we look for the stability of the concentric fringes as we move our line of sight across the plate with a small aperture. If t is slowly varying, the center of the

circular fringe system also appears to change. If t is decreasing, we are moving toward the thinner side of the wedge, and in this case the circular Haidinger fringes appear to expand from the center. On the contrary, the fringes appear to converge to the center if we are moving toward the thick side of the wedge. If we note how many times the center of the fringe system has gone through bright and dark cycles, we can also estimate the wedge angle in the same manner as for the Fizeau situation.

1.4.2. Use of Laser Source for Haidinger Interferometer

A helium–neon laser source of low power is very useful for this interferometer, as it is for the Fizeau instrument. It enables the fringes to be projected on a screen. In this case, the laser can be made to give effectively a point source of light, and consequently, the Haidinger fringes can be considered as the interference from two point sources that are coherent to each other. Hence it is possible to obtain the circular fringes even at a finite distance from the two coherent point sources, and no lens is needed to form the fringes in its focal plane. Figure 1.34 shows the two point images of a point source reflected from a glass plate having a wedge. For the purpose of analysis, it is sufficient for us to consider two point sources of light that are coherent to each other. Then, if we place a screen sufficiently far away and perpendicular to the line joining the two sources, we get a system of concentric circular fringes similar to Newton’s rings and the center of the fringes is collinear with the two point sources. Also, for a glass plate of refractive index N , the distance between the virtual point sources is $2t/N$, where t is the thickness of the plate. Now, if the glass plate has a small wedge, the two virtual sources will also have a slight lateral displacement with respect to each other; this is given by $2N\alpha r$, where α is the wedge angle and r is the distance of the point source from the wedged plate. These various parameters are illustrated in Figure 1.34.

To apply this theory in practice, several methods are available. One method, proposed by Wasilik et al. (1971), is illustrated in Figure 1.35. The laser beam is allowed to pass through a small hole in a white cardboard and is then incident on

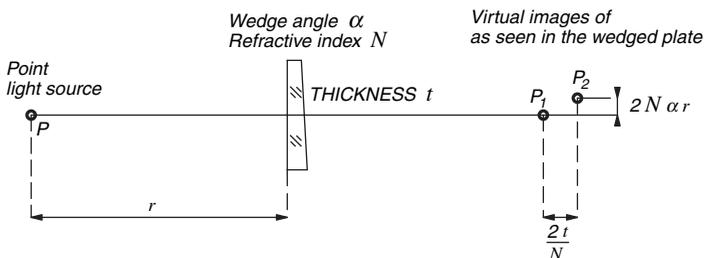


FIGURE 1.34. Various parameters related to the formation of two virtual coherent sources from a single point source by a wedged plate.

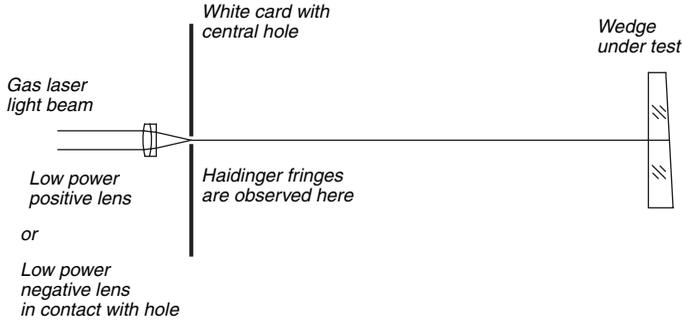


FIGURE 1.35. A schematic arrangement for observing the Haidinger fringes and measuring the displacement of the center. Here a laser beam is passed through a cardboard, and the Haidinger fringes are observed around the hole on the cardboard.

the glass plate under test. To provide some divergence for the laser beam, a negative or positive lens of about 50–100 mm focal length is introduced centrally behind the cardboard. The lens can be fixed in such a manner that it does not deviate the beam but only expands it slightly. This cardboard may be made specially, along with the lens, to fit on the laser. Several concentric circles with known spacing may be drawn on the cardboard for measuring purposes. The plate under test is kept on a platform that can be tilted. The plate is adjusted until the spot of laser light reflected from it goes back through the hole in the cardboard. In this situation concentric circular Haidinger fringes will be seen on the cardboard surrounding the hole. If the plate is free from the wedge, the center of the Haidinger fringe system coincides with the center of the hole. If the plate has a wedge, the center of the Haidinger fringe system is displaced with respect to the center of the hole. An approximate formula relating this displacement to the wedge angle of the wedged glass plate is as follows:

$$d = \frac{2N^2r^2\alpha}{t} \tag{1.28}$$

- where d = displacement of the Haidinger fringe system,
- α = wedge angle of the plate,
- t = thickness of the glass plate,
- N = refractive index of the glass plate,
- r = distance of the point source from the glass plate.

For example, if $\alpha = 1$ s of arc (5×10^{-6} rad), $N = 1.5$, $r = 1000$ mm, and $t = 10$ mm, we have $d = 2.25$ mm, which can be easily detected. Hence this method is quite sensitive and useful.

Another method is illustrated in Figure 1.36. Here the laser beam passes through the wedged-glass plate and falls on a specially prepared ground-glass plate in the center of which a small concave or convex reflector of about

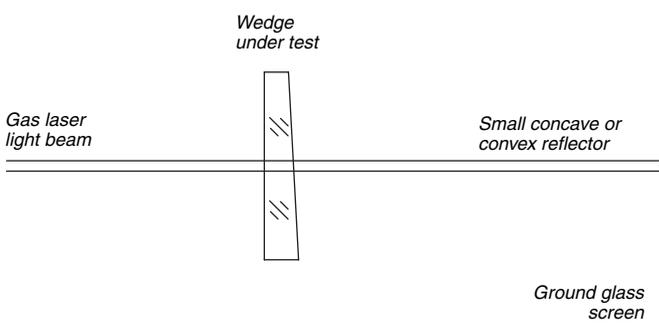


FIGURE 1.36. A schematic arrangement for observing the Haidinger fringes and measuring the displacement of the center. Here the laser beam is directed back into the wedged glass plate by a small concave or convex mirror on the ground glass. The Haidinger fringes are formed on the ground glass.

50–100 mm radius of curvature is cemented. The size of the reflector should be slightly greater than the spot size of the laser. Thus, the laser beam is reflected back onto the glass plate. The wedged plate is adjusted until the reflected spot from it coincides with the small reflecting mirror on the ground glass. Now, Haidinger fringes can be seen on the ground glass, and the center of the fringe system is displaced with respect to the mirror on the ground glass. The same formula, Eq. (1.28), is also valid for this case.

A third method utilizes a beam divider, as shown in Figure 1.37. The laser beam passes through the beam divider, which after reflection from the wedged plate is again reflected at the beam divider and finally falls on a ground-glass screen. The plate under test is adjusted until the laser spot reflected from it goes back on itself. After the position of the spot on the ground glass has been noted, a negative or positive lens is introduced into the laser beam close to the laser side. This widens the beam sufficiently so that circular Haidinger fringes can be seen on

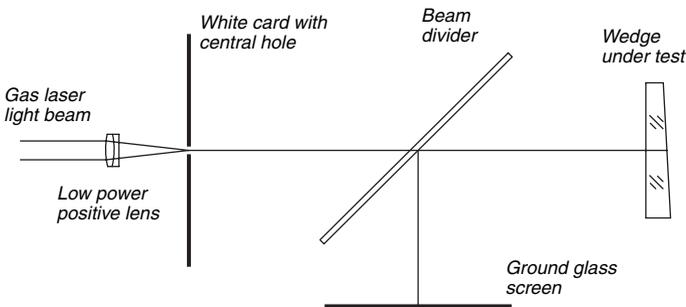


FIGURE 1.37. A schematic arrangement for observing the Haidinger fringes and measuring the displacement of the center. Here the fringes are observed on the ground glass and by means of a beam divider, the central obscuration is avoided.

the ground-glass screen. The displacement of the center of the Haidinger fringe system is measured, and the same formula, Eq. (1.28), can be used for calculating the wedge angle α .

1.4.3. Other Applications of Haidinger Fringes

We have discussed earlier the application of Haidinger fringes for the determination of the very small wedge angle of a nearly parallel plate of glass. There are many types of prisms that can be reduced to equivalent parallel plates and hence can be tested for deviation from their nominal angles. A typical example is a right-angle prism with nominal angles of 90° , 45° , and 45° . In such a prism, it is usually required that the 90° angle be very close to its nominal value and that the two 45° angles be equal to each other. In addition, all the faces of the prism should be perpendicular to a base plane. If not, we say that the prism has pyramidal error that is objectionable in many applications. Figure 1.38 shows how the right-angle prism can be treated as an equivalent parallel plate with a very small wedge angle. If the beam is incident first on the face AC, the beam returning after reflection from the face BC is nearly parallel to the one reflected from the face AC, and hence Haidinger fringes are seen as a result of the interference between these two beams. This arrangement checks the equality of the angles A and B. If there were no pyramidal error and the two angles are equal, the center of the Haidinger fringes will be exactly at the center of the beam spot. If the angles are equal but there is a pyramidal error, the center of the Haidinger fringes will be displaced vertically. If both errors are present, the center will be displaced both vertically and horizontally. The effect of the pyramidal error is to rotate the line of intersection of the two planes of the equivalent wedge so that it is neither vertical nor horizontal. If the beam is incident first on AB, the return beam reflected from the internal face of AB will be nearly parallel to the one reflected from AB externally, and hence we again get Haidinger fringes due to the interference of these two beams. This arrangement checks the exactness of the 90° angle of the angle C. If the center of the Haidinger fringes is not displaced horizontally, the 90° angle is exact; and, if in addition there

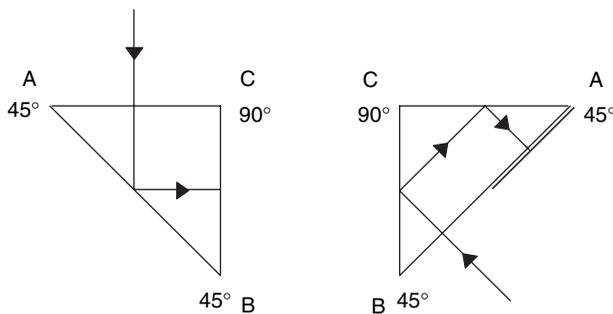


FIGURE 1.38. Schematic of the $45^\circ - 90^\circ - 45^\circ$ prism to be equivalent parallel plates of glass.

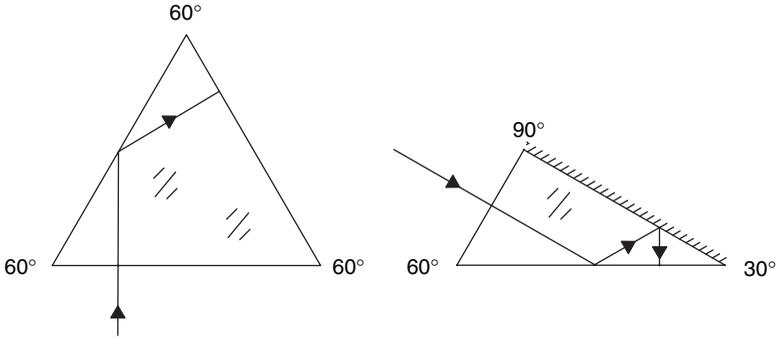


FIGURE 1.39. Schematics of two other prisms to be equivalent parallel plates of glass.

is no vertical displacement, there is no pyramidal error. More details of this method may be found in the paper by Saxena and Yeswanth (1990).

Other examples of prisms that may be treated as equivalent parallel plates are shown in Figure 1.39. Readers may come across other examples depending on particular situations.

The formula given in Eq. (1.27) is applicable in the situations described noting that the displacement has two components, one in the vertical direction and one in the horizontal direction.

1.5. ABSOLUTE TESTING OF FLATS

Until now, we have considered the testing of flats against a “perfect” flat taken as a reference. It is, however, often necessary to make a flat when a good reference flat is not available. In this case, an alternative is to use a liquid flat as mentioned in Section 1.2.2. Another possibility is to make three flats at the same time and test them with several combinations in order to obtain the absolute departure of the three surfaces with respect to an ideal flat.

Let us assume that we have three surfaces that will be tested in many combinations by placing them in pair, one against the other. One of the two glass disks (A) is placed on top of the other by flipping in x by rotation about an axis that is parallel to the y -axis. If the surface deformation is represented by $F_A(x, y)$ as illustrated in Figure 1.40 (a), it is now expressed by

$$[F_A(x, y)]_x = -F_A(-x, y) \quad (1.29)$$

The glass disk at the bottom (B) may be rotated by an angle θ with respect to its original position, as in Figure (1.40)(c). Then, its surface deformation is represented by $[F_B(x, y)]_\theta$ as expressed by

$$[F_B(x, y)]_\theta = F_B(x \cos \theta - y \sin \theta, x \sin \theta + y \cos \theta) \quad (1.30)$$

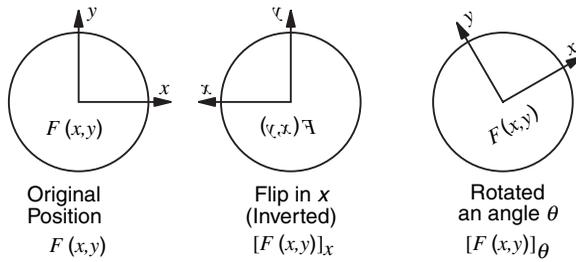


FIGURE 1.40. Three possible orientations for the surfaces whose deviations with respect to an ideal plane are to be determined.

Then, by measuring the fringe pattern we can obtain the value of the difference:

$$G_{BA}(x, y) = [F_B(x, y)]_\theta - [F_A(x, y)]_x \tag{1.31}$$

If, following the treatment by Schulz and Schwider (1976) we take $\theta = 0$ and take the three possible combinations, (see Figure 1.41) we obtain

$$\begin{aligned} G_{BA}(x, y) &= F_B(x, y) + F_A(-x, y) \\ G_{CA}(x, y) &= F_C(x, y) + F_A(-x, y) \\ G_{CB}(x, y) &= F_C(x, y) + F_B(-x, y) \end{aligned} \tag{1.32}$$

This system has more unknowns than equations. Along the y axis, if we make $x = 0$, the system has a simple solution. A solution for all the plane can be obtained only the

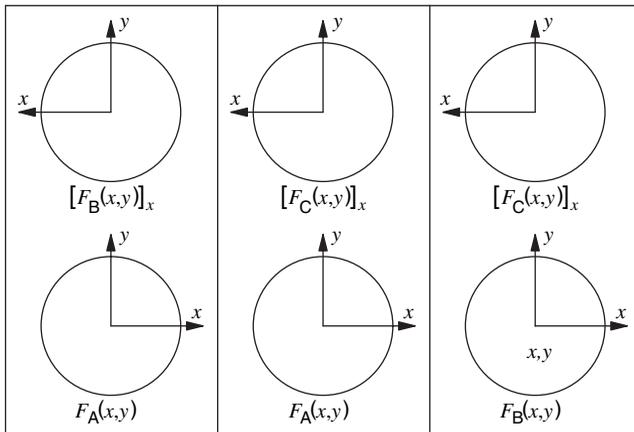


FIGURE 1.41. Three different combinations for the three surfaces to be measured.

if the symmetry about the y axis at least for one surface is assumed, for example for surface B, by taking $F_B(x, y) = F_B(-x, y)$. Then, we may obtain

$$\begin{aligned} F_A(x, y) &= \frac{G_{BA}(-x, y) + G_{CA}(-x, y) - G_{CB}(-x, y)}{2}, \\ F_B(x, y) &= \frac{G_{BA}(x, y) - G_{CA}(x, y) + G_{CB}(x, y)}{2}, \\ F_C(x, y) &= \frac{G_{CA}(x, y) - G_{BA}(x, y) - G_{CB}(x, y)}{2}. \end{aligned} \quad (1.33)$$

Several other methods had been devised, for example by Truax (1988). A specially interesting method is that of Ai and Wyant (1992) as follows next. Let us assume that the shape of one of the surfaces may be represented by the function $F(x, y)$. Any one-dimensional asymmetrical function can be represented by the sum of one even (symmetrical) function and one odd (antisymmetrical) function. On the contrary, for any asymmetrical function $f(x)$ the following properties hold:

$$\begin{aligned} (1) \quad F_e(x) &= f(x) + f(-x) \quad \text{is even,} \\ (2) \quad F_o(x) &= f(x) - f(-x) \quad \text{is odd.} \end{aligned} \quad (1.34)$$

Generalizing this result to two dimensions

$$\begin{aligned} (1) \quad F_{ee}(x, y) &= f(x, y) + f(-x, y) + f(x, -y) + f(-x, -y) \quad \text{is even-even;} \\ (2) \quad F_{eo}(x, y) &= f(x, y) + f(-x, y) - f(x, -y) - f(-x, -y) \quad \text{is even-odd;} \\ (3) \quad F_{oe}(x, y) &= f(x, y) - f(-x, -y) + f(x, y) - f(-x, y) \quad \text{is odd-even;} \\ (4) \quad F_{oo}(x, y) &= f(x, y) - f(-x, y) + f(x, -y) - f(-x, -y) \quad \text{is odd-odd.} \end{aligned} \quad (1.35)$$

The conclusion is that any two-dimensional asymmetric function $F(x, y)$ can always be decomposed into the sum of four functions, even-even, even-odd, odd-even, and odd-odd as follows:

$$F(x, y) = F_{ee} + F_{oo} + F_{oe} + F_{eo}, \quad (1.36)$$

where

$$\begin{aligned} F_{ee}(x, y) &= (F(x, y) + F(-x, y) + F(x, -y) + F(-x, -y))/4, \\ F_{oo}(x, y) &= (F(x, y) - F(-x, y) - F(x, -y) + F(-x, -y))/4, \\ F_{eo}(x, y) &= (F(x, y) + F(-x, y) - F(x, -y) - F(-x, -y))/4, \\ F_{oe}(x, y) &= (F(x, y) - F(-x, y) + F(x, -y) - F(-x, -y))/4. \end{aligned} \quad (1.37)$$

Let us now assume that we test two flats with surface shapes $F_A(x, y)$ and $F_B(x, y)$ by placing one over the other. Following Ai and Wyant, eight combinations are selected,

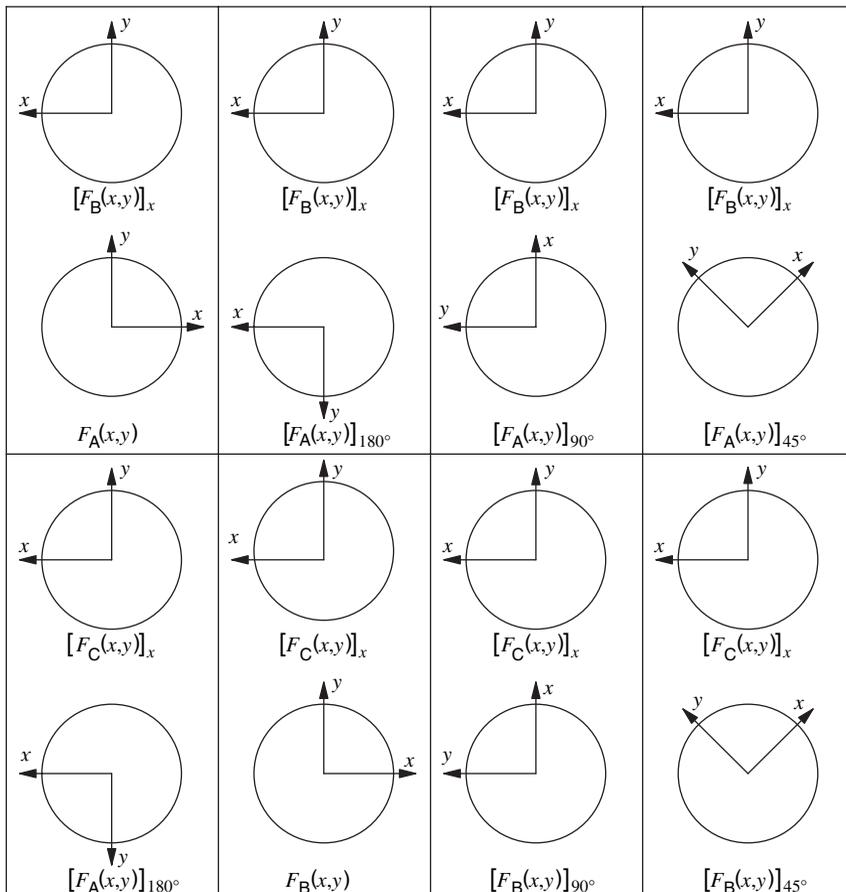


FIGURE 1.42. Eight different combinations for the three surfaces to be measured.

as in Figure 1.42, where the optical path difference producing Newton or Fizeau fringes will be

$$\begin{aligned}
 G_1(x, y) &= [F_B(x, y)]_x - F_A(x, y) \\
 G_2(x, y) &= [F_B(x, y)]_x - [F_A(x, y)]_{180} \\
 G_3(x, y) &= [F_B(x, y)]_x - [F_A(x, y)]_{90} \\
 G_4(x, y) &= [F_B(x, y)]_x - [F_A(x, y)]_{45} \\
 G_5(x, y) &= [F_C(x, y)]_x - [F_A(x, y)]_{180} \\
 G_6(x, y) &= [F_C(x, y)]_x - F_B(x, y) \\
 G_7(x, y) &= [F_C(x, y)]_x - [F_B(x, y)]_{90} \\
 G_8(x, y) &= [F_C(x, y)]_x - [F_B(x, y)]_{45}
 \end{aligned} \tag{1.38}$$

With these expressions, the entire profile of the three planes can be calculated. First, the odd-even, the even-odd, the even-even, and the odd-odd components of the three desired functions $F_A(x, y)$, $F_B(x, y)$, and $F_C(x, y)$ are calculated. The odd-odd component is the most difficult to evaluate, which is obtained using Fourier sine series. Tilt and piston term are not obtained, but this is not a problem since they do not have any practical interest.

Fritz (1983 and 1984) proposed a method using Zernike polynomials to decompose the desired functions into orthogonal functions. Later, Shao et al. (1992) found that by neglecting some high spatial frequencies, the solution can be obtained by using only four combinations.

REFERENCES

- Ai C. and J. Wyant, "Effect of Spurious Reflection on a Phase Shift Interferometry," *Appl. Opt.*, **27**, 3039–3045 (1988).
- Ai C. and J.C. Wyant, "Absolute Testing of Flats Decomposed to Even and Odd Functions," *Proc. SPIE*, **1776**, 73–83 (1992).
- Ai C. and J. Wyant, "Effect of Retroreflection on a Phase-Shifting Interferometer," *Appl. Opt.*, **32**, 3470–3478 (1993).
- Burge J. H., "Fizeau Interferometry for Large Convex Surfaces," *Proc. SPIE*, **2536**, 127–138 (1995).
- Deck L., D. Stephenson, M. Küchel, and C. Evans, "Advances in Laser Fizeau Interferometry for Optical Testing," *TOPS*, **76**, 108–111 (2000).
- Forman P. F., "A Note on Possible Errors Due to Thickness Variations in Testing Nominally Parallel Plates," *Appl. Opt.*, **3**, 646 (1964).
- Forman P. F., "The Zygo Interferometer System," *Proc. SPIE*, **192**, 41 (1979).
- Fritz B. S., "Absolute Calibration of an Optical Flat," *Proc. SPIE*, **433**, 123 (1983).
- Fritz B. S., "Absolute Calibration of an Optical Flat," *Opt. Eng.*, **23**, 379 (1984).
- Grigor'ev V. A., Va. O. Zaborov, and P. P. Ivanov, "Use of a Liquid Mirror for the Calibration of an Interferometer," *Sov. J. Opt. Technol.*, **53**, 613 (1986).
- Hariharan P. and W. H. Steel, "Fringe Localization Depth: A Comment," *Appl. Opt.*, **28**, 29 (1989).
- Karow H. H., "Interferometric Testing in a Precision Optics Shop: A Review of Testplate Testing," *Proc. SPIE*, **192**, 56 (1979).
- Ketelsen D. A. and D. S. Anderson, "Optical Testing With Large Liquid Flats," *Proc. SPIE*, **966**, 365 (1988).
- Malacara D. and A. Cornejo, "Testing of Aspherical Surfaces with Newton Fringes," *Appl. Opt.*, **9**, 837 (1970).
- Moore R. C. and F. H. Slaymaker, "Direct Measurement of Phase in a Spherical Wave Fizeau Interferometer," *Appl. Opt.*, **19**, 2196 (1980).
- Murty M. V. R. K., "A Note on the Testing of Large Aperture Plane Parallel Plates of Glass," *Appl. Opt.*, **2**, 1337 (1963).
- Murty M. V. R. K., "Addendum to: A Note on the Testing of Large Aperture Plane Parallel Plates of Glass," *Appl. Opt.*, **3**, 784 (1964a).

- Murty M. V. R. K., "The Use of a Single Plane Parallel Plate as a Lateral Shearing Interferometer with a Visible Gas Laser Source," *Appl. Opt.*, **3**, 531 (1964b).
- Murty M. V. R. K. and R. P. Shukla, "Some Considerations of the Fizeau Interferometer," *Bull. Opt. Soc. India*, **4**, 13 (1970).
- Novak E., C. Ai and C. Wyant, "Errors Caused by Nearly Parallel Optical Elements in a Laser Fizeau Interferometer Utilizing Strictly Coherent Imaging," *Proc. SPIE*, **3134**, 456–460 (1997).
- Primak W., "Optical Flatness Standard," *Opt. Eng.*, **23**, 806 (1984).
- Primak W., "Optical Flatness Standard II: Reduction of Interferograms," *Proc. SPIE*, **954**, 375 (1989a).
- Primak W., "Optical Flatness Standard: Comment," *Opt. Eng.*, **28**, 934 (1989b).
- Saxena A. K. and L. Yeswanth, "Low Cost Method for Subarcsecond Testing of a Right Angle Prism," *Opt. Eng.*, **29**, 1516–1520 (1990).
- Schulz G. and J. Schwider, in *Progress in Optics*, vol. **13**, E. Wolf, Ed., North Holland, Amsterdam, 1976, Chap. IV.
- Schulz G. and J. Schwider, "Comments on the Paper "Optical Flatness Standard," *Opt. Eng.*, **26**, 559 (1987).
- Shack R. V. and G. W. Hopkins, "The Shack Interferometer," *Opt. Eng.*, **18**, 226 (1979).
- Simon J. M. and S. A. Comatri, "Fringe Localization Depth," *Appl. Opt.*, **26**, 5125 (1987).
- Smith W. S., "Versatile Interferometer for Shop Use," *Proc. SPIE*, **192**, 13 (1979).
- Taylor W. G. A., "Spherical Aberration in the Fizeau Interferometer," *J. Sci. Instrum.*, **34**, 399 (1957).
- Truax B. E., "Absolute Interferometric Testing of Spherical Surfaces," *Proc. SPIE*, **966**, 130 (1988).
- Vannoni M. and G. Molezini, "Joint Interferometric Measurement of Planarity and Parallelism," *Opt. Eng.*, **43**, 1215–1220 (2004).
- Wasilik H., T. V. Biomquist, and C. S. Willet, "Measurement of Parallelism of the Surfaces of a Transparent Sample Using Two Beam Nonlocalized Fringes Produced by Laser," *Appl. Opt.*, **10**, 2107 (1971).
- Wyant J. C., "Fringe Localization," *Appl. Opt.*, **17**, 1853 (1978).
- Yoder P. R., Jr. and W. W. Hollis, "Design of a Compact Wide Aperture Fizeau Interferometer," *J. Opt. Soc. Am.*, **47**, 858 (1957).

2

Twyman–Green Interferometer

D. Malacara

2.1. INTRODUCTION

The Twyman–Green interferometer is a modification of the Michelson interferometer used to test optical components. It was invented and patented by Twyman and Green (1916) for the testing of prisms and microscope objectives and was later adapted and applied to the testing of camera lenses (Twyman, 1919). The first publications on this instrument were those of Twyman (1918a, 1918b, 1920, 1920–1921, 1923). The instrument has been very useful and so popular that many review papers (Briers, 1972) and books (Candler, 1951; Twyman, 1957; U.S. Department of Defense, 1963; Cook, 1971; Horne, 1972) describe it in detail. One of the basic Twyman–Green configurations is illustrated in Figure 2.1. After the system has been illuminated with a quasi-monochromatic point light source, the light is collimated by means of lens L_1 in order to form a flat wavefront. The wavefront is divided in amplitude by means of a beam-splitter plate. After reflection, light from both mirrors M_1 and M_2 impinges again on the beam splitter. Two interference patterns are then formed, one going to lens L_2 and the other going back to the light source. Lens L_2 permits all of the light from the aperture to enter the eye so that the entire field can be seen. The observed fringes are of equal thickness type.

It is easy to see that if the beam splitter is all dielectric, the main interference pattern is complementary to the one returning to the source; in other words, a bright fringe in one pattern corresponds to a dark fringe in the other. This has to be so because of the conservation of energy principle, even though the optical path difference is the same for both patterns. Phase shifts upon reflection account for this complementarity. The case of a absorbing beam splitter has been treated by Parmigiani (1981).

It is interesting that Michelson (1918) did not consider the instrument applicable to the testing of large optics, pointing out at the same time that the arrangement we now know as an unequal-path interferometer was impractical because of the lack of sufficiently coherent light sources. To answer Michelson's comments, Twyman

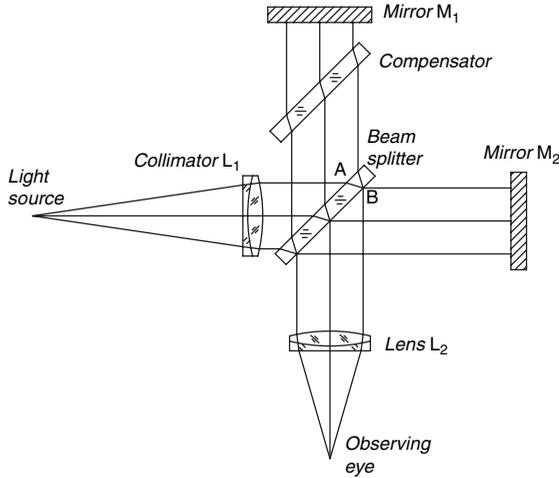


FIGURE 2.1. Basic Twyman–Green interferometer configuration.

(1918) pointed out that the arrangement shown in Figure 2.2 had been suggested in his patent (Twyman and Green, 1916) for the testing of large mirrors or lenses. This procedure eliminates the need for a large collimator and beam splitter but unfortunately requires (for sources of limited coherence) a concave spherical mirror as large as the optical element under test. This kind of arrangement is often referred to as a Williams interferometer (De Vany, 1965; Grigull and Rottenkolber, 1967) because Burch (1940) attributed it to Williams. A Twyman–Green interferometer for general laboratory usage is shown in Figure 2.3.

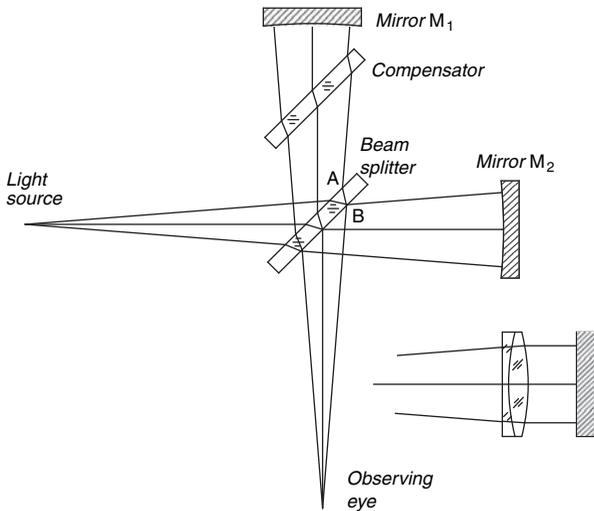


FIGURE 2.2. Twyman–Green interferometer (Williams type).

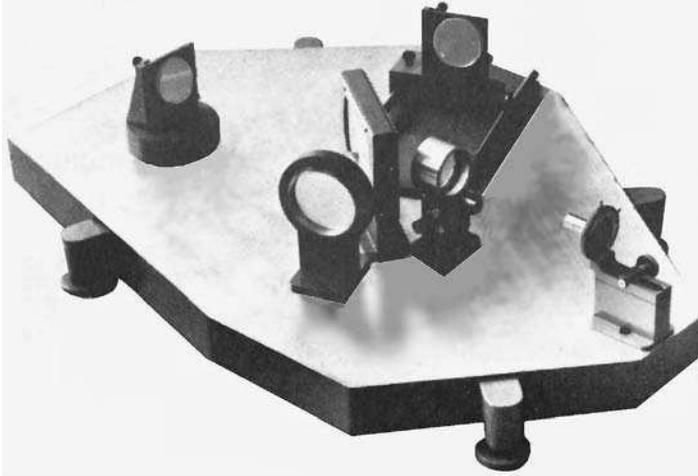


FIGURE 2.3. A general purpose Twyman-Green interferometer.

2.2. BEAM-SPLITTER

The beam-splitter can take the form of glass plate as shown in Figure 2.4 or a cube. It is made in such a way that the splitting face reflects the light in appropriate amount by means of a partially reflective coating. If it is a plate, the other face should not reflect any light. To avoid reflections on the second face, a multilayer antireflection coating

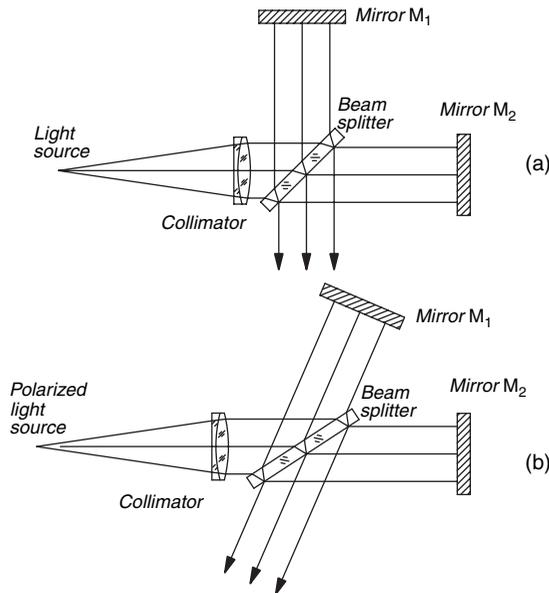


FIGURE 2.4. Twyman Green interferometer with beam splitter (a) at 45° and (b) at the Brewster's angle.

can be used. However, an easier solution is to place the plate at Brewster's angle as shown in Figure 2.4(b) and to use a source of light with p polarization, which goes through the interface without any reflection. Still another solution is to introduce a wedge angle on the plate so that the unwanted reflected light escapes from the system. The contrast of the fringes does not depend on the reflectance R of the beam splitter; only the irradiance I on the fringe maxima is affected, since

$$I = 4I_0RT, \quad (2.1)$$

where I_0 is the irradiance of the incident wavefront and T is the transmittance. If there is no absorption (nonmetallic coating), $T \approx 1 - R$ and there is a maximum value of I for $R = T = 1/2$.

If one of the two mirrors has a higher reflectance than the other, for example if one of them is a highly reflective mirror and the other is still uncoated, the intensity of one of the two interfering beams can be about 25 times higher than the other. Under these conditions, the fringe contrast is greatly reduced. A solution is to use a beam splitter with a different reflectivity. Another possibility is to use both mirror without coating.

2.2.1. Optical Path Difference Introduced by a Beam Splitter Plate

We can show with a few algebraic steps and the law of refraction that a beam splitter or compensating plate shifts the optical axis laterally and parallel to itself by the following amount:

$$d = t \sin \theta \left[1 - \frac{\cos \theta}{(n^2 - \sin^2 \theta)^{1/2}} \right], \quad (2.2)$$

where θ is the incidence angle, t is the plate thickness, and n is the refractive index.

The light going to the observer from mirror M_1 has traversed the beam splitter only once, whereas the light from mirror M_2 has gone through it three times. An interferometer that has more glass in one arm than in the other, as in this case, is said to be uncompensated. The interferometer can be compensated by inserting another piece of glass in front of mirror M_1 as shown in Figures 2.1 and 2.2.

The importance of compensating an interferometer is clearly seen in the following section. Adjustable compensators for Williams configurations (Steel, 1963) and Twyman–Green configurations (Connes, 1956; Mertz, 1959; Steel, 1962) have been described in the literature.

As pointed out before, an interferometer is said to be uncompensated when it has more glass in one of its arms than in the other, because (a) an optical component (lens or prism) is present in one arm in order to test it or (b) the light travels once through the beam splitter in one path and three times in the other path, and the compensating plate is absent. Both of these situations can be included in a general case in which an inclined plane glass plate is placed in one of the arms. The unfolded optical paths for both arms of the interferometer

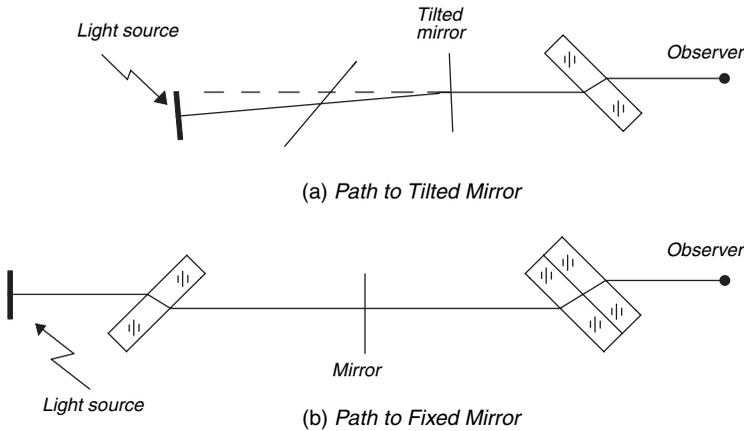


FIGURE 2.5. Light paths for both interfering beams in an uncompensated interferometer.

are shown in Figure 2.5. Here we may see that the complete effect is equivalent to going through a system of two plates. The optical path difference (OPD) introduced by one passage through a glass plate is a function of the angle of incidence of the light, as shown in Figure 2.6, yielding

$$OPD = n(AB) + (BC) - (AD) \tag{2.3}$$

and then

$$OPD(\varphi) = t_N(n \cos \varphi - \cos \varphi) \tag{2.4}$$

If the plates are inclined at an angle φ_0 with respect to the optical axis and the ray direction is defined by the angles θ and ψ as shown in Figure 2.7, the OPD introduced by both passages may be computed by

$$OPD(\varphi) = t_N(n \cos \varphi'_1 + n \cos \varphi'_2 - \cos \varphi'_1 - \cos \varphi'_2) + 2t_0 \cos \theta \tag{2.5}$$

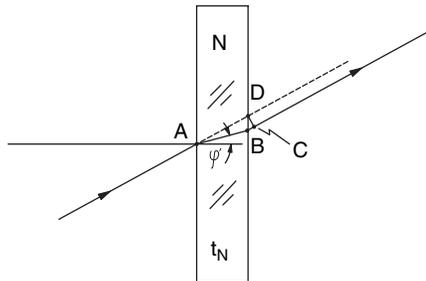


FIGURE 2.6. Optical path difference introduced by a plane parallel plate.

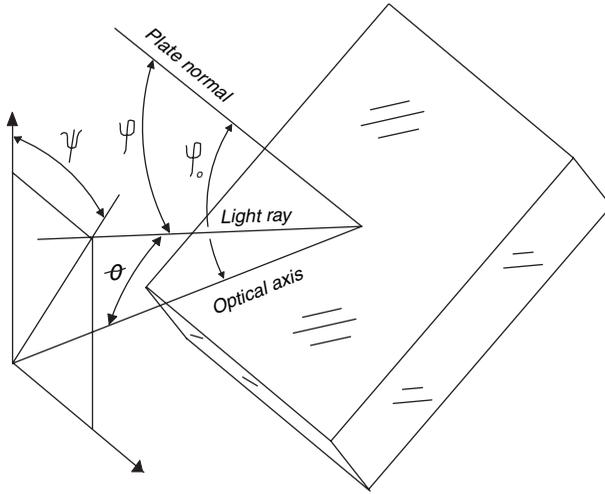


FIGURE 2.7. Light passing through an inclined plane parallel plate.

where subscripts 1 and 2 designate the first and second passages, respectively, through the plate. The last term corresponds to an additional OPD introduced by a shift t_0 of one of the mirrors along the optical axis. Angles φ_1 and φ_2 are obtained from angles φ_1 , ψ , and φ_0 by means of the relations

$$\begin{aligned} \cos \varphi_1 &= \cos \varphi_0 \cos \theta + \sin \varphi_0 \sin \theta \cos \psi \\ \cos \varphi_{21} &= \cos \varphi_0 \cos \theta - \sin \varphi_0 \sin \theta \cos \psi \end{aligned} \tag{2.6}$$

If the glass plate is normal to the optical axis, $\varphi_0 = 0$ and $\varphi_1 = \varphi_2 = \theta$ thus $\varphi'_1 = \varphi'_2 = \theta'$. In this case Eq. (2.5) reduces to

$$\text{OPD}(\theta) = 2t_N(n \cos \theta' - \cos \theta) + 2t_0 \cos \theta \tag{2.7}$$

2.2.2. Required Accuracy in a Beam Splitter Plate

When constructing a Twyman–Green interferometer with a beam splitter plate, it is very important to determine the required surface quality in each one of the two faces of the beam splitter, for a desired interferometer accuracy. These may be found considering the following model of a beam splitter with three small localized defects with thicknesses δt_1 , δt_2 , and δt_3 as shown in Figure 2.8.

In this model we use the fact that in the external reflection in defect 1, if the reflection is displaced a small distance t from the plate, then the optical path is reduced by an amount

$$\text{OPD} = 2t \cos \theta. \tag{2.8}$$

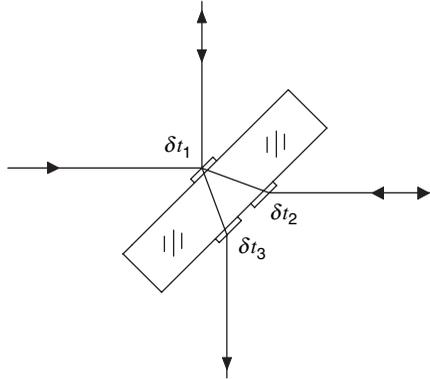


FIGURE 2.8. Effect of errors in the faces of the beam splitter of a Twyman-Green interferometer.

For beams reflected in the two faces of glass plate with plane and parallel faces, there is an optical path difference given by

$$\text{OPD} = 2nt \cos \theta'. \quad (2.9)$$

Thus, in the internal reflection at the localized defect 1, the optical path is increased. Upon transmission through a glass plate with thickness t , the additional optical path introduced by the plate is

$$\text{OPD} = t_N(n \cos \theta' - \cos \theta). \quad (2.10)$$

Thus, the optical path introduced by the plate defects for the beam reflected from mirror A is

$$\delta \text{OP}_A = -2\delta t_1 \cos \theta + (\delta t_1 + \delta t_3)[n \cos \theta' - \cos \theta], \quad (2.11)$$

and the optical path introduced by the plate defects for the beam reflected from mirror B is

$$\delta \text{OP}_B = -2n\delta t_1 \cos \theta' + (\delta t_1 + 2\delta t_2 + \delta t_3)[n \cos \theta' - \cos \theta]. \quad (2.12)$$

Thus, the difference between these two quantities is the optical path difference OPD introduced by the plate errors:

$$\delta \text{OPD} = -2\delta t_1(\cos \theta + n \cos \theta') + 2\delta t_2(n \cos \theta' - \cos \theta) \quad (2.13)$$

As expected, this optical path difference does not depend on the error δ_3 , since it is common for both beams. Of course, any error δ_2 it is also an error δ_3 for some other ray.

If OPD is the desired interferometer accuracy, the required accuracy δ_1 on the plates face 1 is given by

$$\delta t_1 = \frac{\delta \text{OPD}}{2(\cos \theta + n \cos \theta')}, \tag{2.14}$$

and the required accuracy δt_2 on the plate face 2 is

$$\delta t_2 = \frac{\delta \text{OPD}}{2(\cos \theta' + n \cos \theta)}. \tag{2.15}$$

These results mean that, roughly, the reflecting face must be polished with about twice the interferometer required accuracy, while the other face must have a quality of only about half the interferometer accuracy.

2.2.3. Polarizing Cube Beam Splitter

The beam splitter can also take the form of a glass cube. It can be a non-polarizing beam splitter or a polarizing beam splitter. The second option has many advantages that we will describe in Section 2.8.1, as pointed out by Bruning and Herriott (1970) and Bruning et al. (1974).

Two important properties of the cube beam splitter are that the interferometer is automatically compensated and that most beam splitter faces are all dielectric with no absorption. If polarized light is used, some other important characteristics are present. Figure 2.9 shows an interferometer using a polarizing cube beam splitter, where all the P polarized component of the light beam is transmitted while the S polarized component of the light beam is reflected. If the

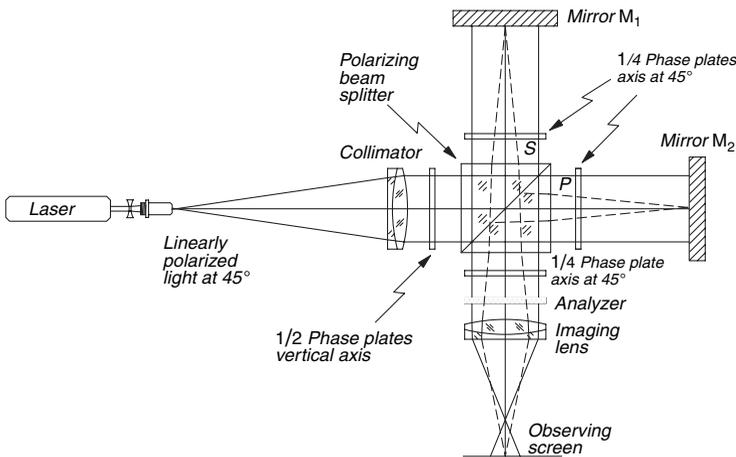


FIGURE 2.9. Twyman Green interferometer with a polarizing cube beam splitter.

incident light beam is linearly polarized in a plane at 45° with respect to the square cube edges the reflected intensity is equal to the transmitted intensity. However, as mentioned before, if the mirrors M_1 and M_2 have different reflectivities, a half wave phase plate can be inserted before the cube beam splitter to maximize the fringe contrast. If the angle between the slow or fast axis of the phase plate forms an angle $\theta/2$ with respect to the plane of polarization of the incident beam, this plane of polarization will rotate an angle θ . If the fast and slow axes of the phase plate are interchanged by rotating the phase plate 90° , the phase or the output beam whose plane of polarization is rotated changes 180° . The S and the P components will have different intensity as desired depending on this angle.

When the transmitted and the reflected light go to mirrors M_1 and M_2 , both beams pass twice through quarter wave phase plates with their axes at 45° before returning to the beam splitter. Thus, both planes of polarization will rotate by 90° . This allows the returning beams to go to the observing screen instead of returning to the light beam. So, there is no returning complementary interference pattern when using a nonpolarizing beam splitter. Both interference patterns go to the observing screen, but we can separate them later, as we will next describe. Figure 2.10 illustrates the polarization states at different points along the light trajectories. We can see that after recombining and exiting the beam splitter, the two beams are orthogonally polarized. So, interference cannot take place. Let us assume that these two beams with S and P polarizations have the amplitudes

$$E_x = Ae^{ikz} \quad \text{and} \quad E_y = Ae^{ik(z+OPD)}, \quad (2.16)$$

where OPD is the optical path difference between the two beams. A $\lambda/4$ phase plate with its axis at 45° is placed after the beam splitter to transform these two beams into

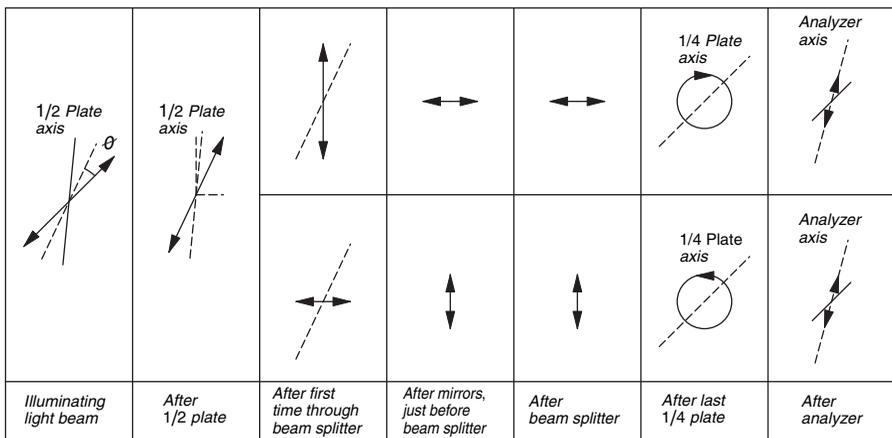


FIGURE 2.10. Polarization states at different locations in the interferometer in Fig. 2.9.

two circularly polarized beams with opposite sense. The total electric field due to the superposition of the two fields along the slow axis can be shown to be

$$E_{Slow} = \frac{1}{\sqrt{2}} A(1 + e^{-ik \text{OPD}})e^{ikz}, \quad (2.17)$$

and along the fast axis

$$E_{Fast} = \frac{1}{\sqrt{2}} A(1 - e^{-ik \text{OPD}})e^{i(kz+\pi/2)}. \quad (2.18)$$

If we place a linear polarizer (analyzer) in front of these circularly polarized beams, only the components along the axis of the polarizer forming an angle α with the slow axis of the phase plate will pass through. Thus, the amplitude along this axis of the polarizer is

$$E_{\alpha} = E_{Slow} \cos \alpha + E_{Fast} \sin \alpha, \quad (2.19)$$

which can be transformed into

$$E_{\alpha} = \frac{1}{\sqrt{2}} A(e^{i\alpha} + e^{-i\alpha} e^{-ik \text{OPD}})e^{ikx}, \quad (2.20)$$

or equivalently into

$$E_{\alpha} = \sqrt{2} A \cos\left(\alpha + \frac{k \text{OPD}}{2}\right) e^{ik(x-(\text{OPD}/2))}. \quad (2.21)$$

Hence, the interferogram irradiance can be shown to be

$$I_{\alpha} = E_{\alpha} E_{\alpha}^* = A^2 [1 + \cos(k \text{OPD} + 2\alpha)]. \quad (2.22)$$

The conclusion is that the orientation α of the axis of the analyzer will change the phase difference between the two interfering beams without modifying the contrast. The phase difference between the two interfering beams changes linearly with the angle. By turning the analyzer by an angle α , the phase difference will change by 2α . This effect is frequently used in phase shifting interferometers.

2.2.4. Nonpolarizing Cube Beam Splitter

A simplified version of the previously described interferometer configuration where the main beam splitter is nonpolarizing and only one phase plate is used is illustrated in Figure 2.11. The light source is linearly polarized at 45° , which means that the x and y components of the electric vector are in phase. The $\lambda/8$ phase plate is traversed twice, sending back to the beam splitter circularly polarized light, which is equivalent to saying that the x and y components have a phase difference equal to 90° .

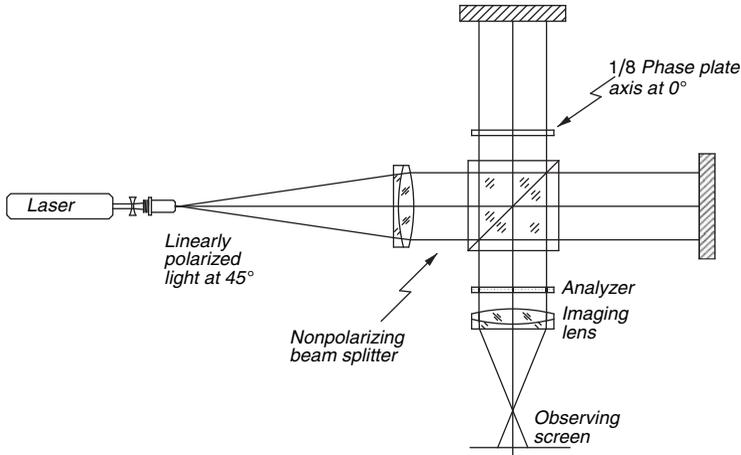


FIGURE 2.11. Twyman Green interferometer with a nonpolarizing cube beam splitter.

As in the previous interferometer, the exiting light beams cannot interfere, since one of them is linearly polarized at 45° and the other is circularly polarized. If an analyzer is rotated in front of the interfering light beams coming out of the beam splitter, the interference pattern appears, since only the components along the axis of the polarizer, forming an angle θ with the x axis will pass through it. It can be proved that the interference patterns transmitted with the analyzer at an angle θ and at an angle $\theta + 90^\circ$ are complementary. The phase depends on the angle of the analyzer, but for this arrangement with a nonpolarizing beam splitter, the fringe contrast is maximum only at the analyzer angles $\theta = \pm 45^\circ$.

2.3. COHERENCE REQUIREMENTS

The size (spatial coherence) and monochromaticity (temporal coherence) of the light source must satisfy certain minimum requirements that depend on the geometry of the system, as described by Hansen (1955, 1984) and by Birch (1979). It is interesting to know that if the optical element under test has very steep reflections, the state of polarization of the light may change in the reflection, introducing changes in the contrast (Ferguson, 1982). However, in most of the cases, the important factor in the contrast is the coherence of the light source.

2.3.1. Spatial Coherence

The light source for interferometry must satisfy some minimum requirements of spatial as well as temporal coherence, depending on the interferometer configuration and the specific application and needs. As described in Chapter 1, Section 1.1.2., some gas or vapor lamps can be used in conjunction with a small pinhole to illuminate an interferometer. These lamps with the pinhole do not have perfect

spatial and temporal coherence. A gas laser, however, has perfect spatial coherence and can have almost perfect temporal coherence. We might think at first that this is the ideal light source for interferometry, but this is not always the case. The coherence length is, in general, so large that many unwanted reflections from other surfaces in the optical system may produce a lot of spurious fringes in addition to the speckle noise that make it difficult to analyze the interferogram. On the contrary, the perfect spatial coherence produce scattering spherical waves from many unavoidable small pieces of dust or scratches on the optical elements, which in turn produce many spurious rings of fringes. This problem has been studied in detail by Schwider (1999).

The conclusion is that, quite frequently, it is a better option to use a gas or vapor light source instead of a laser. However, if the optical path difference is large, it is unavoidable to use a gas laser. In this section we will study the coherence requirements for the light source.

There are two cases for which the collimated wavefront has ray lights spread over a solid angle with diameter 2θ , and hence the final accuracy of the interferometry or the contrast will be reduced:

- (1) The collimator has spherical aberration, in which

$$\theta = \frac{TA}{f}, \quad (2.23)$$

where TA is the maximum value of the transverse spherical aberration of the collimator at its best focus position. This aberration might limit the accuracy of the interferometer unless the OPD remains constant with changes in the angle θ . Otherwise, given the maximum value of θ , the maximum change in the OPD should be smaller than the desired accuracy.

- (2) The light source is not a mathematical point but has a small diameter $2a$; then

$$\theta = \frac{a}{f}, \quad (2.24)$$

where f is the focal length of the collimator.

Fringes with high contrast are obtained, using an extended thermal source, only if the OPDs for the two paths from any point of the source with different value of θ differ by an amount smaller than $\lambda/4$ according to the Rayleigh criterion. On the contrary, radiometric considerations usually require as large a source as possible that will not degrade the contrast of the fringes.

When the beam splitter is a glass plate and is not compensated by another identical glass plate, we may show that the maximum light source size has an elliptical shape. This is the reason why the fringes are elliptical in an uncompensated Michelson interferometer. The shape and the size of the ellipse not only are functions of φ_0 , θ , and ψ but also depend very critically on t_0 .

The simpler case of a glass plate with its normal along the optical axis can be analyzed with more detail as will be shown. The OPD is given by Eq. (2.7). As shown

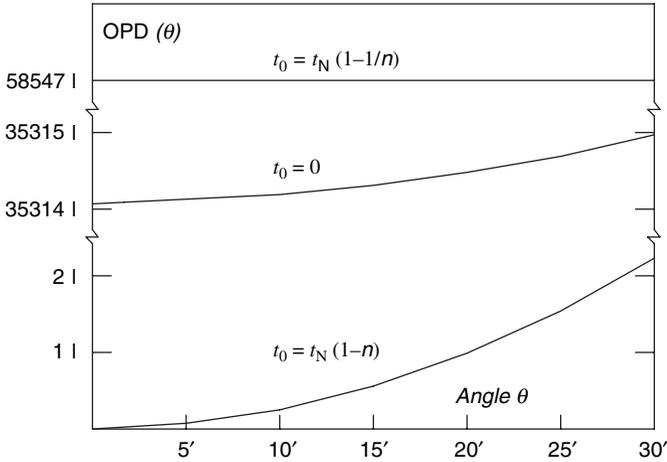


FIGURE 2.12. Optical path difference introduced by a plane parallel plate normal to the optical axis ($t = 2$ cm, $n = 1.52$, $\lambda/4 = 589$ nm.).

in Figure 2.12, the value of the OPD changes with the value of θ depending on the value of t_0 . The maximum allowed value of the angular semidiameter θ of the light source as seen from the collimator is that which gives a variation of the OPD equal to $\lambda/4$. On the contrary, the maximum allowed value of the angle θ due to spherical aberration of the collimator is that which gives a variation of the OPD equal to the accuracy desired from the interferometer.

When testing small optics using a nonmonochromatic light source, the optical path difference can be adjusted to be zero. Then, it is convenient to choose

$$t_0 = t_N(1 - n), \quad (2.25)$$

so that $\text{OPD}(0^\circ) = 0$, but this situation will require an even smaller light source. It should be pointed out that when testing large optics, the value of t_0 cannot be changed at will, because in general it will be very large.

If an extended quasi-monochromatic light is used, a good condition in order to make the optical path difference insensitive to the angle θ is

$$\frac{d \text{OPD}(\theta)}{d\theta} = 0, \quad (2.26)$$

yielding

$$t_0 = t_N \left(1 - \frac{1}{n} \right). \quad (2.27)$$

It is interesting to see that this equation is equivalent to the condition that the apparent distance of the image of the collimator (or the light source in a Michelson

interferometer) to the observer is the same for both arms of the interferometer. This condition seems reasonable if we consider that then the angular size of the two images of the light source is the same as pointed out by Steel (1962) and Slevogt (1954).

When the light source is extended and the interferometer is compensated in this manner, the fringes are localized at a certain plane in space. To find this plane, the system may be unfolded as studied by Hansen (1942, 1955). For an interferometer with plane mirrors, this location for the fringes is near the plane mirrors because of the way the image of the light source moves when one of the mirrors is tilted in order to obtain the fringes as shown in Figure 2.5. Thus, the viewing system must be focused near the mirrors to see the fringes.

As it will be described later in this chapter, to test an optical system, one of the plane mirrors is replaced by the system to be tested, plus some auxiliary optics to send back a collimated beam to the interferometer, just like one single mirror would do. The returning collimated beam has to have the same diameter. Thus, it is easy to conclude that this whole system, including the element under test has the following general characteristics:

- (1) It is afocal.
- (2) Its magnification is either one or minus one. If it is minus one, the returning wavefront will be rotated with respect to the incident wavefront.
- (3) The system is symmetric and hence it is always free of coma.
- (4) Entrance and the exit pupil are symmetrically placed with respect to the system and have the same diameter.

With these properties we see that since the system is reflective (it is retroreflector only if the magnification is minus one), the entrance and exit pupil are at the same plane. An important conclusion is that the fringes should be observed at this entrance and exit pupil plane. This problem has been studied with detail by Schwider and Falkenstörfer (1995).

It should be noticed that the entrance pupil of the whole system is not necessarily the same as the pupil of the lens under test. However, when testing a lens, the fringes are to be observed at the pupil, which ideally should be the same. This does not happen with a single mirror; therefore, the mirror should be as close as possible to the lens. This the reason why a convex mirror with the longest possible radius of curvature is desirable (Steel, 1966) when testing telescope objectives. On the contrary, the entrance pupil of a microscope objective is at infinity; hence, the exit pupil is at the back focus. Dyson (1959) described an optical system such as the one to be described in Chapter 12, which images the mirror surface on the back focus of the microscope objective, where the fringes are desired.

The limitation on the size of a pinhole source was examined in a slightly different manner by Guild (1920–1921) as explained below. Imagine that the small source is greatly enlarged to form an extended source. Then fit an eyepiece in front of lens L_2 (see Fig. 2.1) to form a telescope. Under these conditions, equal inclination fringes in

the form of concentric rings (like the ones normally observed in the Michelson interferometer) are observed. If the mirrors are exactly perpendicular to their optical axes, the rings will be exactly centered. The ideal size of the source is that which allows only the central spot on the fringe system to be observed. The size of the central spot increases when the OPD (θ) reduces its dependence on θ by one of the adjustments described above, making possible the use of a larger source, although the effective size of the spot is then limited by the pupil of the observing eye or the camera.

In all the foregoing considerations, the two interfering wavefronts are assumed to have the same orientation, that is, without any rotations or reversals with respect to each other. In other words, if one of the beams is rotated or reversed, the other should also be rotated or reversed. A wavefront can be rotated 180° by means of a cube corner prism or a cat's-eye retroreflector formed by a convergent lens and a flat mirror at its focus. The wavefront can be reversed upon reflection on a system of two mutually perpendicular flat surfaces, e.g., in a Porro prism. Murty (1964) showed that if one of the wavefronts is rotated or reversed with respect to the other, then, to have fringes with good contrast and without phase shifts, the pinhole diameter 2α should satisfy the condition

$$2\alpha \leq \frac{1.2\lambda f}{D}, \quad (2.28)$$

so that diametrically opposite points over the wavefront are coherent to each other. Here, f and D are the collimator's focal length and diameter, respectively. Then 2α is extremely small and therefore an impractical size for some sources. However, there is no problem if a gas laser is used, because its radiance and spatial coherence are extremely high. This subject will be examined with more detail in Chapter 5.

When testing an optical element as it will be described in the following section, the wavefront is sometimes inverted (up-down) or reversed (left-right) or rotated (both), which is equivalent to a rotation of one of the wavefronts by 180° . Then, the spatial coherence requirements increase. If a laser is used, no problem arises. If a gas or vapor source is used, the reference wavefront has also to be inversed, reversed, or rotated, like the wavefront under test.

When there is no alternative but to use a gas laser source, due to a large optical path difference, speckle noise and spurious fringes may be reduced by artificially reducing the spatial coherence of the light a little. This is possible by placing a small rotating ground glass disc on the plane of the pinhole as described by Murty and Malacara (1965), Schwider and Falkenstörfer (1995), and Schwider (1999).

2.3.2. Temporal Coherence

The OPD (θ) given by Eq. (2.5) also imposes some minimum requirements on the monochromaticity of the light source. Considering first the case of an interferometer that is uncompensated because of the lack of a compensating plate or the presence of

an optical system with lenses or prisms in one of the arms, from Eq. (2.5) we can write

$$\Delta \text{OPD}(0^\circ) = 2t_N \left(\frac{dn}{d\lambda} \right) \Delta\lambda \quad (2.29)$$

and, using again the Rayleigh criterion ($\Delta \text{OPD} \leq \lambda/4$),

$$\Delta\lambda \leq \frac{\lambda}{8t_N \left(\frac{dn}{d\lambda} \right)} \quad (2.30)$$

Therefore, since the glass dispersion ($dn/d\lambda$) is never zero, and t_n is also nonzero, the bandwidth $\Delta\lambda$ must not have a very large value if the interferometer is not compensated. If the interferometer is exactly compensated, white light fringes can be observed when OPD is nearly zero; otherwise, a highly monochromatic light source such as a low vapor pressure lamp or (even better) a gas laser must be used.

If many different kinds of glasses are present in both arms of the interferometer, we may take a more general approach by considering that the interferometer is compensated for the bandwidth $\Delta\lambda$ of the light, if the phase difference for the light following the two paths in the interferometer is independent of the wavelength. According to Steel (1962), we can say that, if each arm of the interferometer contains a series of optical components of thickness t and refractive index n , the phase difference for the two arms is

$$\phi = \frac{2\pi}{\lambda} \left(\sum_1 nt - \sum_2 nt \right) \quad (2.31)$$

This relative phase is independent of the wavelength when $d\phi/d\lambda = 0$, thus giving

$$\sum_1 \tilde{n}t = \sum_2 \tilde{n}t \quad (2.32)$$

where \tilde{n} is the “group refractive index,” defined by

$$\tilde{n} = n - \lambda \frac{dn}{d\lambda} \quad (2.33)$$

Thus, the interferometer is compensated for the bandwidth $\Delta\lambda$ when the “group optical path” for both arms is the same. Steel (1962) pointed out that the compensation for the bandwidth of the light source can be examined by looking at the fringes formed by a white light source through a spectroscope with its slit perpendicular to the fringes. The spectrum is crossed by the fringes and their inclination shows the change of fringe position with wavelength. The fringes will be straight along the direction of dispersion if the bandwidth compensation is perfect. Otherwise, the glass optical paths can be adjusted until the fringes show a maximum (zero slope) at the

wavelength to be used. If the bandwidth of the light source is very large, a detailed balancing of the glass paths has to be made by using the same types and thicknesses of the glass on both arms.

If the $OPD(0)$ is very large, as in the unequal-path interferometer (described in Section 2.5), the last term in Eq. 2.5 dominates, and we can write

$$OPD(0^\circ) = 2t_0 = m\lambda \quad (2.34)$$

but from the Rayleigh criterion, the order number m should not change from one end of the wavelength bandwidth to that of the other by more than $1/4$; thus

$$m\lambda = \left(m + \frac{1}{4}\right)(\lambda - \Delta\lambda) \quad (2.35)$$

where $\Delta\lambda$ is the maximum allowed bandwidth. Thus we can write the approximation

$$\Delta\lambda \leq \frac{\lambda}{8t_0} \quad (2.36)$$

Since the length of a train of waves with bandwidth $\Delta\lambda$ is equal to $\lambda^2/\Delta\lambda$, this condition is equivalent to saying that the $OPD(0^\circ)$ should be smaller than one fourth of the length of the wavetrain (or wavelength of the modulation). In uncompensated interferometers, this condition is incompatible with the condition for an extended light source.

A very interesting and practical case occurs when the light source is a gas laser, but this discussion is left to Section 2.5 on the unequal-path interferometer.

2.4. USES OF A TWYMAN-GREEN INTERFEROMETER

Many different kinds of optical components can be tested with this instrument. The simplest one to test is a plane parallel plate of glass, as shown in Figure 2.13. The OPD introduced by the presence of the glass plate is given by

$$OPD = 2(n - 1)t \quad (2.37)$$

where t is the plate thickness and n is the refractive index. If the interferometer is adjusted so that no fringes are observed before introducing the plate into the light beam, all the fringes that appear are due to the plate. If the field remains free of fringes, we can say that the quantity $(n - 1)t$ is constant over all the plate. If straight fringes are observed, we can assume that the glass is perfectly homogeneous (nconstant) and that the fringes are due to an angle ϵ between the two flat faces, given by

$$\epsilon = \frac{\alpha}{2(n - 1)} \quad (2.38)$$

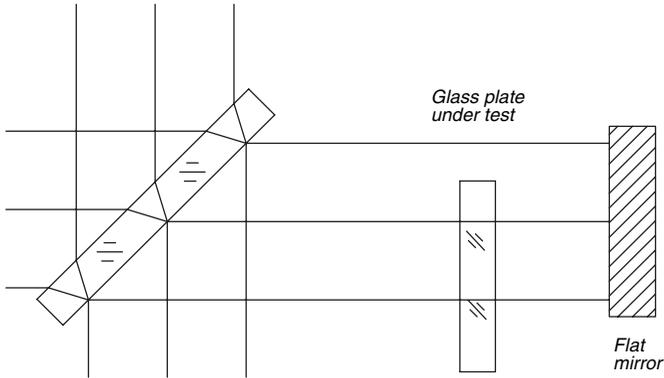


FIGURE 2.13. Testing a glass plate.

where α is a small angle between the two interfering wavefronts, which can be determined from

$$\alpha = n\lambda \tag{2.39}$$

Here m is the number of interference fringes per unit length being observed.

The fringes, however, may not be straight but quite distorted, as shown in Figure 2.14, because of bad surfaces or inhomogeneities in the index, since the only quantity we can determine is $(N - 1)t$. To measure the independent variations of n and t , we must complement this test with another made in a Fizeau interferometer, which measures the values of nt (Kowalik 1978). Many different kinds of

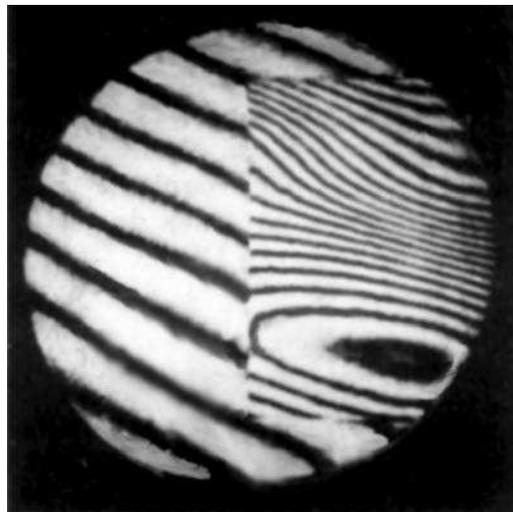


FIGURE 2.14. Interferogram of a glass plate.

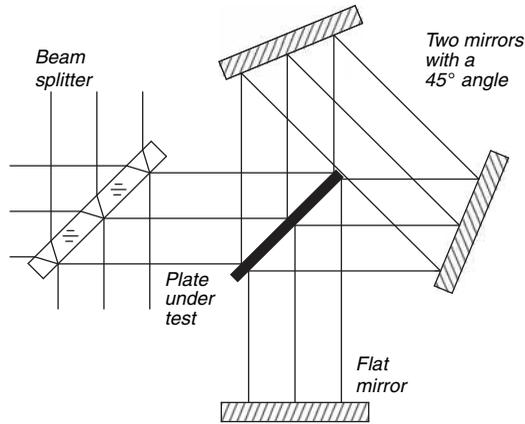


FIGURE 2.15. Testing a thin and flexible opaque glass plate.

material can be tested with this basic arrangement (Adachi et al., 1961, 1962; Masuda et al., 1962; Twyman and Dalladay, 1921–1922).

In many instruments, when using a glass window, the important requirement is that the optical path difference introduced by its presence is a constant for the whole aperture as in Eq. (2.30). However, sometimes the plate may not be transparent in the visible, only in the infrared, where it is used, and an infrared interferometer is not available. If a constant index of refraction is assumed, the important parameter is a constant thickness. But an independent measurement of the flatness of the faces does not permit this evaluation because the plate is frequently so thin that it may bend, which, on the contrary, is not important in its operation. For these cases, Williamson (2004) has described a configuration as shown in Figure 2.15. The interesting characteristic of this configuration is that if the plate bends or curves in any way, the change in one of its two faces is canceled out by the corresponding change in the other.

2.4.1. Testing of Prisms and Diffraction Rulings

The Twyman–Green interferometer is a very useful instrument for testing prisms. Its application for testing the accuracy of the 90° angle between two of the faces of a right angle (Porro) prism, a roof (Amici) prism, or a cube corner prism is especially interesting. As explained before, the relative rotation or reversal of the wavefronts should be corrected, as shown in Figure 2.16, if a gas laser is not used. The arrangements in Figure 2.17 can be used when a gas laser source is employed.

A very good cube corner prism will give rise to an interferogram like that shown in Figure 2.18. The fringes are straight throughout the aperture. A cube corner prism with angular errors produces an interferogram such as that shown in Figure 2.19, in which the straight fringes abruptly change their direction. Thomas and Wyant (1977) made a complete study of the testing of cube corner prisms.

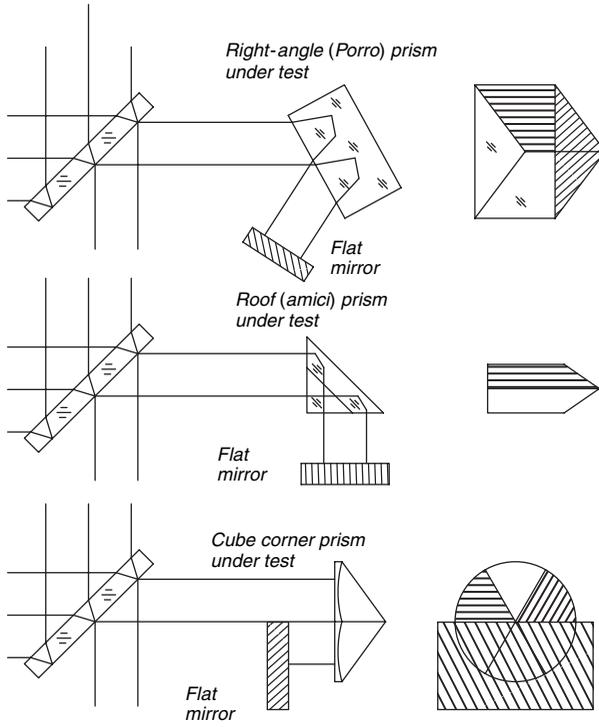


FIGURE 2.16. Testing some prisms.

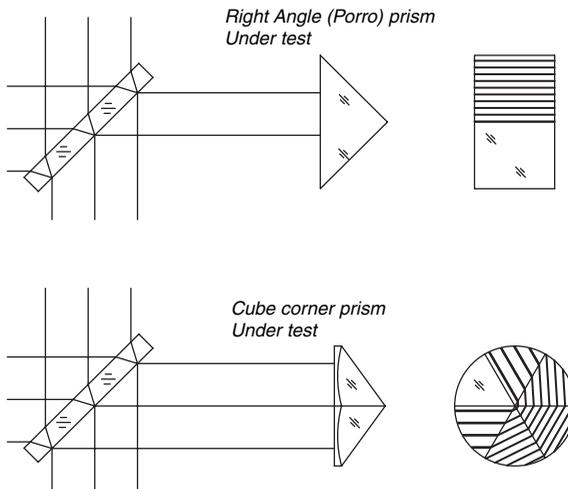


FIGURE 2.17. Testing some prisms by retroreflection with laser illumination.

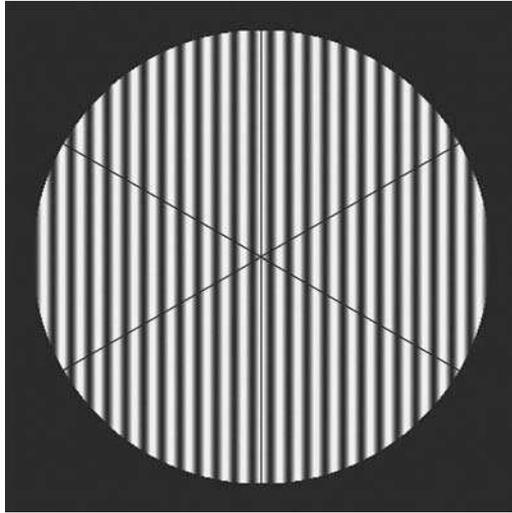


FIGURE 2.18. Interferogram of a good corner cube prism tested in the retroreflecting configuration and with some tilt.

Figures 2.20 and 2.21 show similar situations for a right-angle prism of no error and of some angular error, respectively. If, in addition to angle errors, the surfaces are not flat or the glass is not homogeneous, an interferogram with curved fringes is obtained. When a right angle or porro prism is tested in the retroreflective configuration and the surface flatness as well as the 90° angle is correct, the fringes look

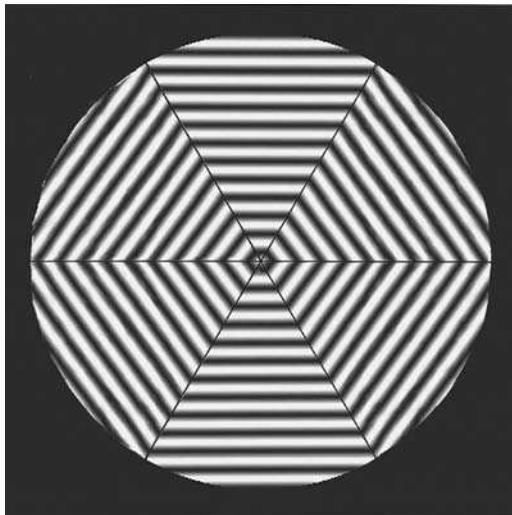


FIGURE 2.19. Interferogram of a corner cube prism with some errors in the angles of the faces tested in the retroreflecting configuration and with an average tilt equal to zero.

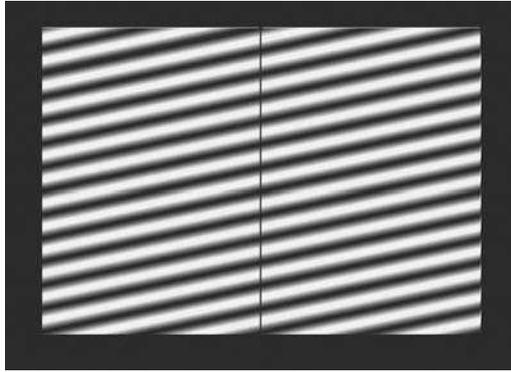


FIGURE 2.20. Interferogram of a good porro prism tested in the retroreflecting configuration and with some tilt.

straight and parallel as in Figure 2.20. If the right angle has an error, the fringes look like those shown in Figure 2.21 and can be manipulated to look like those in Figure 2.22. We describe here a brief method for obtaining the angular error in a right angle prism. If $2L$ is the width of the face of the prism, $\pi/2 \pm \epsilon$ is the angle of the prism, d is the distance between two successive fringes, k is the deviation of the fringe from the straight fringe after bending, n is the refractive index of the prism, and λ is the wavelength used. As shown in Figure 2.23, the error is given by

$$\epsilon = \frac{\alpha}{2N} = \left(\frac{k}{d}\right) \left(\frac{\lambda}{4nL}\right)$$

where α is the angle between the two exiting wavefronts. For example, for a prism of 100 mm face width and $k/d = 0.25$, the error ϵ of the 90° angle is about 1 s of arc. In regard to the sign of the error, the hot rod or finger procedure described before can be used.

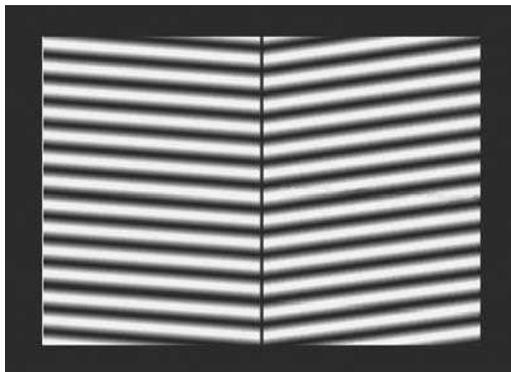


FIGURE 2.21. Interferogram of a porro prism with a small error in the angle tested in the retroreflecting configuration. It has some tilt about the x axis and an average tilt equal to zero about the y axis.

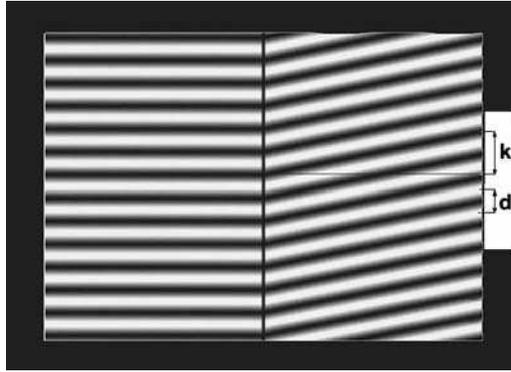


FIGURE 2.22. Interferogram of a porro prism with a small error in the angle tested in the retroreflecting configuration. It has some tilt about the x axis and a zero tilt about the y axis on the left side.

Luneburg (1964) showed that the angular error ϵ in a roof face of a prism is

$$\epsilon = \frac{\alpha}{4mn \sin \theta} \quad (2.40)$$

where n is the refractive index of the material, α is the angle between the two exiting wavefronts in a single pass through the prism, θ is the angle between the roof edge and the incident beam, and m is the number of times the light is reflected on the roof face. For the arrangements shown in Figs. 2.15 and 2.16 we have the values in Table 2.1. The angle α is determined from Eq. (2.32), but with the interferometer adjusted in such a way that all the fringes in one of the faces are eliminated.

A dispersive prism can also be tested as shown in Figure 2.24(a). This arrangement of smoothly changing inhomogeneities in the glass may be compensated for by appropriately figuring the faces. An axicon may be tested in a Twyman-Green interferometer using the method described by Fantone (1981) as well as reflexicons (Hayes et al., 1981).

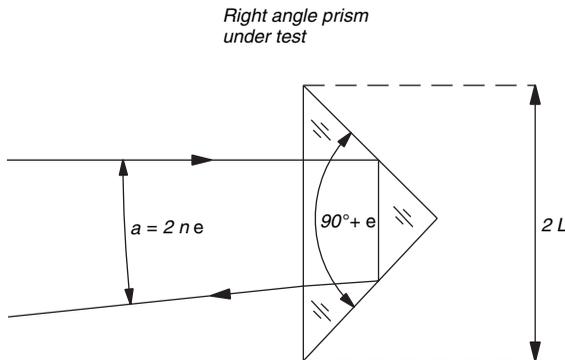


FIGURE 2.23. Testing a porro prism with a small error in the angle in a retroreflective configuration.

TABLE 2.1 Values of angle θ and of K for prisms in Figures 2.16 and 2.17.

Prism	Figure	θ	$\sin \theta$	K
Porro	2.10	60°	0.8662	2
Amici	2.10	45°	$1/\sqrt{2}$	2
Cube Corner	2.10	54.7°	$\sqrt{2/3}$	2
Porro	2.11	90°	1	1
Cube Corner	2.11	54.7°	$\sqrt{2/3}$	1

In 1935, Bisacre and Simeon suggested a method whereby a diffraction grating could be tested by means of a Twyman–Green interferometer. Unfortunately, they never published their work (Candler, 1951). They used the arrangement shown in Figure 2.24(b). The interferometer is initially adjusted to obtain horizontal fringes in the first order. Then the grating is rotated to pass to the third order, in which the ghosts, if any, are stronger. If there are ghosts and a tilt about an axis along the grating chromatic dispersion is introduced, the fringes have a sawtooth appearance. When the spacing between the horizontal fringes is increased by removing the tilt, the teeth become larger and larger until they form a system of vertical fringes due to the interference between the zero order and the ghost wavefront. Using this interferometer, Jaroszewics (1986) has also tested the spacing error of a plane diffraction grating.

2.4.2. Testing of Lenses

One of the early applications of the Twyman–Green interferometer was the testing of lenses and camera objectives (Twyman, 1920), including the measurement of the chromatic aberration (Martin and Kingslake, 1923–1924). Any of the arrangements in Figure 2.25 can be used to test a convergent lens. A convex spherical mirror with

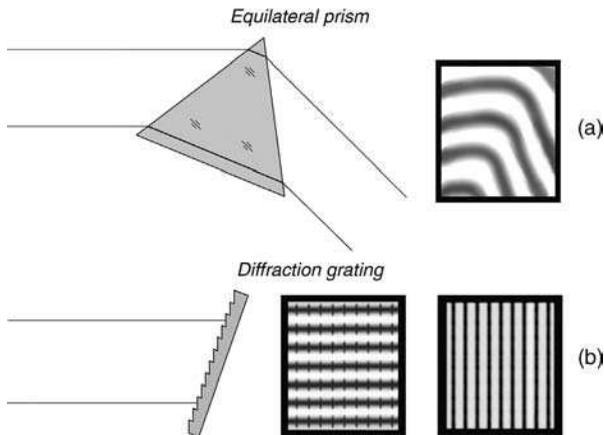


FIGURE 2.24. Testing a dispersive prism and a diffraction grating.

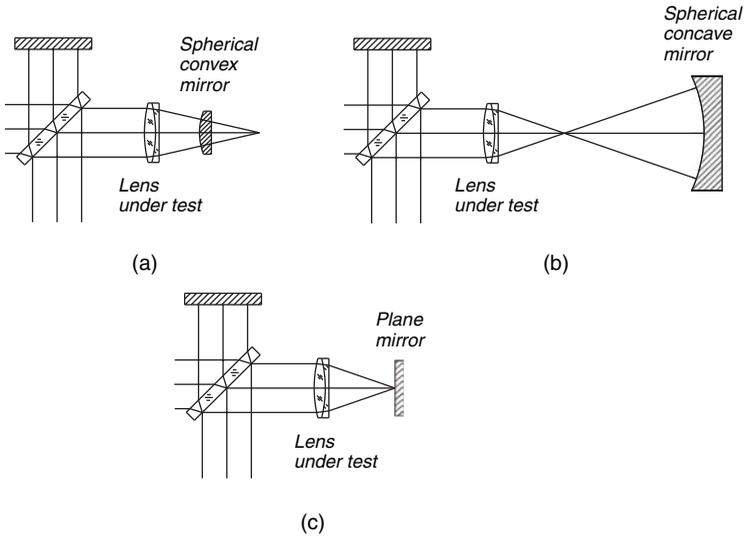


FIGURE 2.25. Three possible arrangements to test a lens.

its center of curvature at the focus of the lens is used for lenses with long focal lengths and a concave spherical mirror for lenses with short focal lengths. A small flat mirror at the focus of the lens can also be employed to great advantage, since the portion of the flat mirror being used is so small that its surface does not need to be very accurate. However, because of the spatial coherence requirements described in Section 2.3.1, the same arrangement or a cube corner prism must be employed on the other interferometer arm, if a laser is not used. Another characteristic of this method is that asymmetric aberrations like coma are canceled out, leaving only symmetric aberrations like spherical aberration and astigmatism.

When a lens is to be tested off axis, it is convenient to mount it in a nodal lens bench as shown schematically in Figure 2.26. The lens L under test is mounted in a rotating mount so that the lens can be rotated about the nodal point N . Since the focal surface is usually designed to be a plane and not a sphere, mirror M is moved

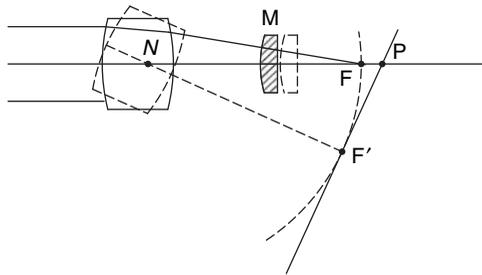


FIGURE 2.26. Testing a lens with a nodal bench.

backward a small distance $\overline{F\bar{P}}$ by pushing the mirror support against a metallic bar $\overline{F\bar{P}}$, fixed with respect to lens L. Interferograms obtained with lenses having third order aberrations will be shown in Section 2.7.

Testing a large lens on the Twyman–Green interferometer requires the use of a beam-splitter plate even larger than the lens. To avoid this difficulty, according to Burch (1940), Williams suggested later by Hopkins (1962) for use with a gas laser in an unequal path configuration.

Complete small telescopes can also be tested with good results as shown by Ostrovskaya and Filimonova (1969).

2.4.3. Testing of Microscope Objectives

Twyman (1920, 1923) also used his instrument for the testing of microscope objectives with good success. The arrangement is essentially the one used for a lens, but a convex mirror, in general, cannot be employed because of the short focus of the objective. Since the microscope sometimes works at a finite tube length, a negative lens is added to change the collimated light and simulate a light source 16 cm away. This lens must be corrected for spherical aberration, but it is not necessary to correct it quasi-monochromatic. As shown in Figure 2.27, several arrangements can be used to test an objective, the most common being a spherical concave mirror with its center of curvature at the focus of the objective. A solid spherical reflector slightly thicker than a hemisphere can serve to simulate the presence of a cover glass.

A plane mirror at the focus of the objective can also be used, but in this case the wavefront is rotated 180°. Therefore, we should either use a laser light source or rotate the wavefront on the other arm. This can be done by means of a cube corner prism or with another microscope objective with the same flat mirror arrangement. It should be pointed out that the interferogram in this case represents the difference between the aberrations of the two objectives.

A fourth arrangement is formed by two oppositely placed microscope objectives. In this case, the interferogram represents the sum of the aberrations of the two objectives. However, when the aberrations to be measured are not small or the pinhole is not small, the best arrangement is the one with the Dyson’s system described in Chapter 12.

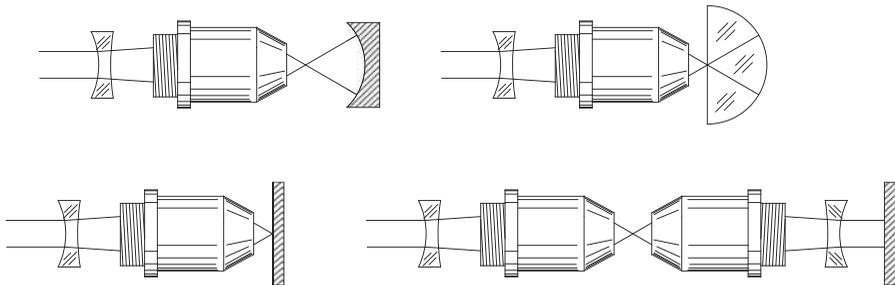


FIGURE 2.27. Testing microscope objectives.

2.5. COMPENSATION OF INTRINSIC ABERRATIONS IN THE INTERFEROMETER

A Twyman-Green interferometer may easily have aberrations of its own due to defective optical components, so that the interferometric pattern is the result of the addition of the aberrations of the optical system under test and the intrinsic interferometer aberrations. Basically, the final aberration may be the superposition of three sources, the reference path W_{ref} , the testing path W_{test} , and the element or optical surface under test W_{surf} . A method to isolate the optical surface aberration from the instrument aberration has been proposed by Jensen (1973). His procedure takes three different measurements with different positions and orientations of the surface under test. To describe it let us assume that these three measurements are as follows.

- (a) At the normal testing position, as in Figure 2.28(a). Then, the interferogram aberration can be written as

$$W_{0^\circ} = W_{surf} + W_{ref} + W_{test} \quad (2.41)$$

- (b) At the normal testing position, but rotating the surface under test 180° , as shown in Figure 2.28(b), the interferogram aberration now is

$$W_{180^\circ} = \bar{W}_{surf} + W_{ref} + W_{test} \quad (2.42)$$

where the bar on top of W_{surf} means that this wavefront aberration has been rotated by 180° .

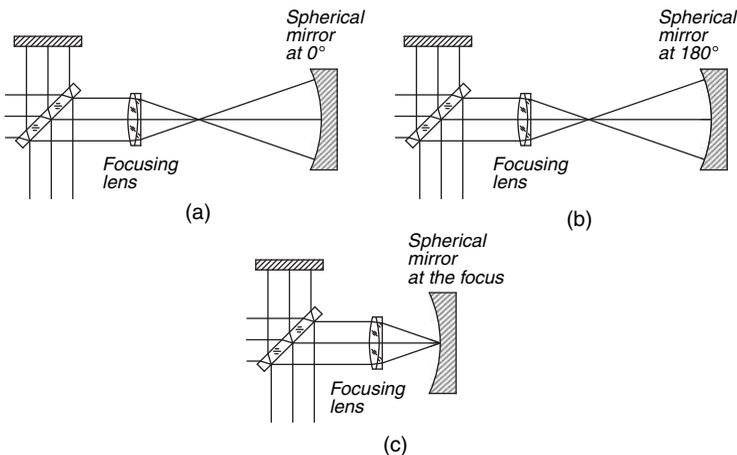


FIGURE 2.28. Calibration of a Twyman-Green Interferometer by absolute testing a concave sphere.

- (c) The vertex of the optical surface under test is placed at the focus of the focusing lens, as illustrated in Figure 2.28(c). Then, the surface aberrations do not appear on the interferogram. However, The reflected wavefront is rotated. Now we have

$$W_{focus} = W_{ref} + \frac{1}{2}(W_{test} + \overline{W}_{test}) \quad (2.43)$$

In any of these three equations, we can rotate all wavefronts in the same expression and it will remain valid. This is done either by placing the bar on top of the W that does not have it or by removing it if it is already there. By using this property, it is possible to obtain

$$W_{surf} = \frac{1}{2}(W_{0^\circ} + \overline{W}_{180^\circ} - W_{focus} - \overline{W}_{focus}) \quad (2.44)$$

With this expression the intrinsic interferometer aberrations are subtracted, making the instrument as if it did not have any aberration of its own. If a large number of similar spherical surfaces are to be tested, the intrinsic instrumental aberration can be expressed as

$$W_{ref} + W_{test} = \frac{1}{2}(W_{0^\circ} - \overline{W}_{180^\circ} + W_{focus} + \overline{W}_{focus}) \quad (2.45)$$

Once the interferometer is calibrated, this intrinsic aberration can be subtracted if the surface under test has a radius of curvature close to the one of the mirror used to make the calibration.

Unfortunately, as shown by Creath and Wyant (1992), this method is quite sensitive to experimental errors due to misalignments, such as decentrations and tilts in the rotation and shifting of the surface under test. In view of this, they proposed a simpler method where both the intrinsic interferometer aberration and the aberration of the spherical surface are almost rotationally symmetric. Then, Eq. (2.37) reduces to

$$W_{surf} = W_{0^\circ} - W_{focus} \quad (2.46)$$

and the intrinsic aberration is just W_{focus} . It is important to point out that this method works for Twyman–Green well as for Fizeau interferometers.

2.6. UNEQUAL-PATH INTERFEROMETER

In Section 2.3, we discussed the coherence requirements of a Twyman–Green interferometer and pointed out that, when a laser light source is used, extremely large OPDs can be introduced (Morokuma et al., 1963). In explaining this, let us first consider the spectrum of the light emitted by a laser. As shown in Figure 2.29(a), the

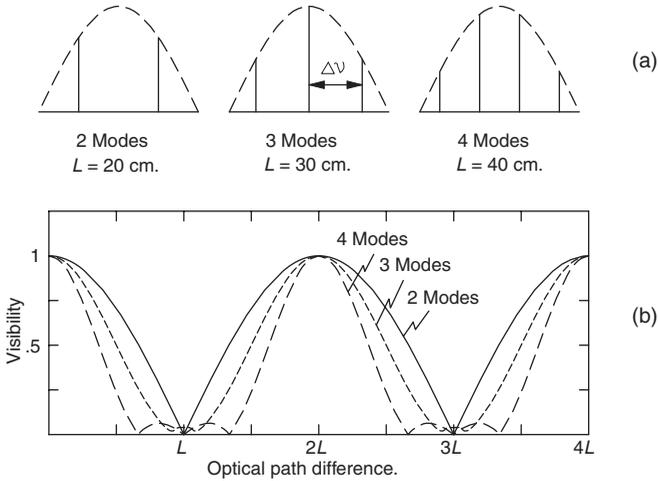


FIGURE 2.29. Visibility in the interferometer using a multimode gas laser.

light emitted by a gas laser usually consists of several spectral lines (longitudinal modes) (Sinclair and Bell, 1969) spaced equally at a frequency interval $\Delta\nu$ given by

$$\Delta\nu = \frac{c}{2L} \quad (2.47)$$

If the cavity length L of a laser changes for some reason (thermal expansion or contractions, mechanical vibrations, etc.), the lines move in concert along the frequency scale, preserving their relative distances $\Delta\nu$, but always with intensities inside the dotted envelope (power gain curve) as shown in Figure 2.29(a).

Lasers that have only one spectral line are called single-mode or single-frequency lasers. They produce a perfect unmodulated wavetrain, but because of instabilities in a cavity length L the frequency is also unstable. By the use of servomechanisms, however, single-frequency lasers with extremely stable frequencies are commercially produced. They are the ideal source for interferometry since an OPD as long as desired can be introduced without any loss in contrast. Unfortunately, these lasers are very expensive and have very low power outputs (less than 1 mW). Even so, a 1-mW laser has a higher radiance than any other type of interferometric source.

It can be shown (Collier et al., 1971) that the theoretical visibility in an interferometer, when a laser source with several longitudinal modes is used, is as illustrated in Figure 2.29(b). Therefore, to have good contrast, the OPD (0) has to be near an integral multiple of $2L$; thus

$$\text{OPD}(0_0) = 2t_0 = 2ML \quad (2.48)$$

Hence lasers are very convenient for Twyman-Green interferometry provided that the mirrors in the interferometer can be adjusted to satisfy this condition. Because of mechanical instability, the laser cavity normally vibrates, producing a continuous

instability in the frequency of the lines. This does not represented any serious problem for small OPDs of the order of 1 m. If, however, the OPD is very large, of the order of 10 or more meters, an almost periodic variation of the contrast is introduced as shown by Batishko and Shannon (1972). To overcome the inconvenience of this effect, they recommend taking photographs with exposures of the order of 1/250 ms. This exposure is fast enough to stop the vibration of the fringes but slow enough so that the contrast variation is integrated out.

A laser with two longitudinal modes can be stabilized to avoid contrast changes by a method recommended by Balhorn et al. (1972), Bennett et al. (1973), and Gordon and Jacobs (1974).

Some suggestions for aligning and adjusting the unequal path interferometer had been given by Zielinski (1978,1979).

2.6.1. Some Special Designs

With the advent of the laser, it became practical to use Twyman–Green interferometers with large optical path differences. Probably the first one to suggest this was Hopkins (1962). An instrument of this type following a Williams arrangement was made by Grigull and Rottenkolder (1967) for wind-tunnel observations and the testing of spherical mirrors.

A very versatile unequal-path interferometer for optical shop testing was designed by Houston et al. (1967). A schematic diagram of this interferometer is shown in Figure 2.30. The beam-splitter plate, which is at the Brewster angle, has a wedge angle of 2–3 min of arc between the surfaces. The reflecting surface of this plate is located to receive the rays returning from the test specimen in order to preclude astigmatism and other undesirable effects. A two-lens beam diverger can be placed in one arm of the interferometer. It is made of high index glass, all the surfaces being spherical, and has the capability for testing a surface as fast as $f/1.7$. A null lens can be used to test an aspheric element, with the combination beam diverger and null lens spaced and aligned as depicted in Figure 2.31 (see Chapter 12)

Another unequal-path interferometer was designed by Kocher (1972). This instrument, shown in Figure 2.32, is quite similar to the Twyman–Green interferometer in

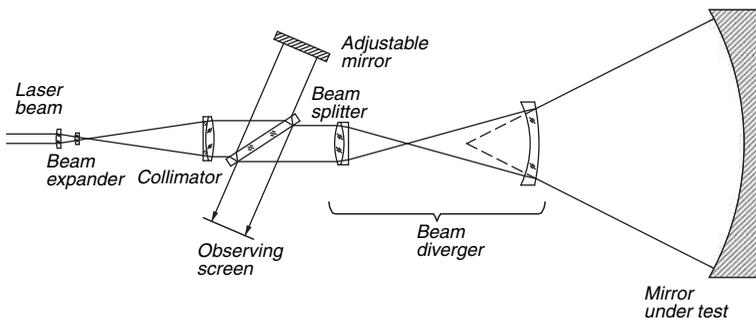


FIGURE 2.30. Houston's unequal path interferometer.

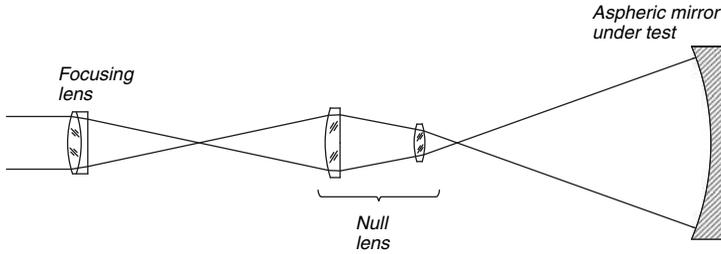


FIGURE 2.31. Null lens and lens diverger for unequal path interferometer.

Figure 2.2. A significant feature is the use of an optically thick beam-splitter substrate in the diverging beam. Such a plate introduces aberrations, but they are intentionally made equal on both arms. To a first approximation there is no effect on the fringe pattern, mainly if the total thickness is kept small. Buin et al. (1969) reported a successful industrial use of unequal-path interferometers.

2.6.2. Improving the Fringe Stability

The unequal-path interferometer frequently has the problem that the fringes are very unstable due to vibrations of the surface under test. Most mirror vibrations have mainly tip and tilt components more than a piston component. Bending is not usually important mainly if the optical elements are thick enough. Thus, one obvious approach is to configure the optical system so that the antisymmetric components of the wavefront from the vibrating system under test are canceled out. The disadvantage of this is that the antisymmetric aberrations, such as coma, are not detected.

One possible implementation of this concept is the arrangement in Figure 2.33. The diverging beam of light going out from the interferometer illuminates the concave surface under test and then the reflected convergent beam gets reflected in a small flat mirror near the center of curvature. Then, the light returns to the mirror, but the returning wavefront has been rotated 180° with respect to the incident

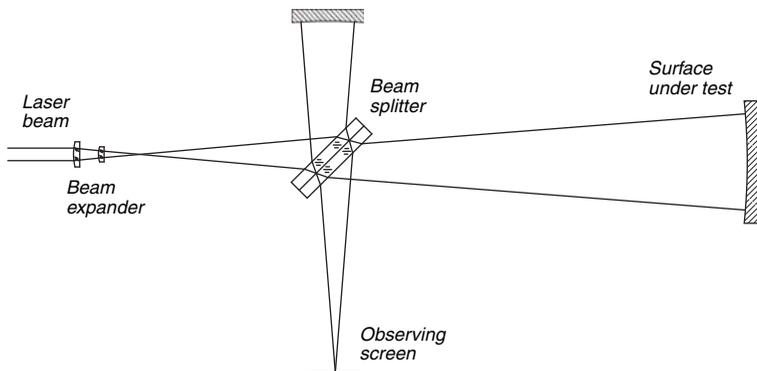


FIGURE 2.32. Kocher's unequal path interferometer.

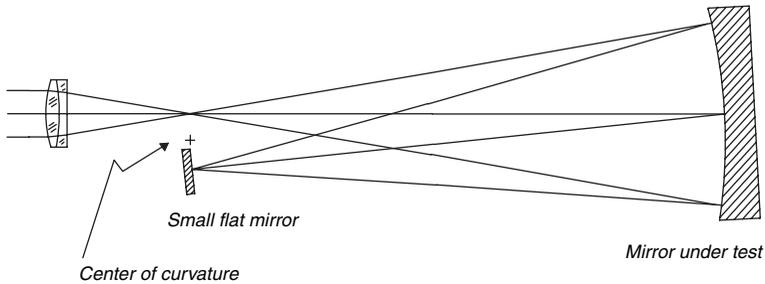


FIGURE 2.33. Stabilization of interference fringes by eliminating antisymmetric errors in the wavefront, by means of a reflection in a small flat mirror.

wavefront. This rotation of the wavefront eliminates all antisymmetric components of the wavefront error, remaining only the symmetric components.

The mirror vibrations may also be eliminated from the interferogram by introducing exactly the same vibrations in the reference wavefront. This is done with an arrangement described by Hardesty (1979), where the reference arm of the interferometer is made as long as the test arm and the reference mirror is placed as close as possible to the tested element, so that they vibrate together.

A third method to dampen the vibrations is by sensing this vibrations with a light detector and with this signal move the reference mirror in the opposite sense (Cole et al., 1997). This is a closed loop servo system that uses a phase sample frequency of several kilo Hertz.

A fourth method is to capture the interferogram image as fast as possible, so that vibrations do not affect the image. This can be done, but the limitation might come if several images with different phase are required, as in phase shifting interferometry. Some interferometers had been designed that allow the simultaneous measurement of four interferograms with different phase.

2.7. OPEN PATH INTERFEROMETERS

An optical element under test in a Twyman–Green interferometer is traversed twice. However, sometimes it is necessary that the sample is traversed by the light beam only once. This can be done if the interferometer configuration is modified by unfolding the light paths. Several possible configurations including the well-known Mach-Zehnder will be described.

2.7.1. Mach-Zehnder Interferometers

As pointed out before, the Mach-Zehnder configuration shown in Figure 2.34 has some advantages with respect to the Twyman–Green configuration. For example, if the sample under test has a large aberration, it is better to pass the beam of light only once through it. Another advantage some times is that this interferometer is automatically compensated because it has two beam splitters.

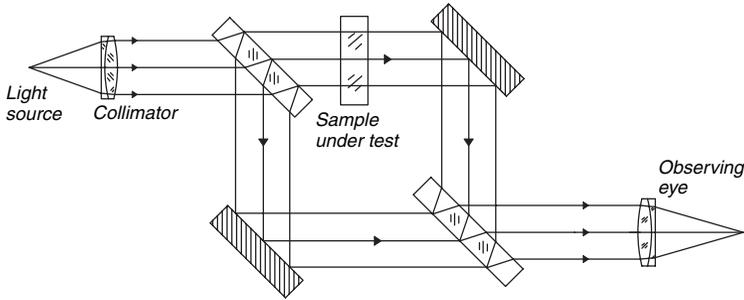


FIGURE 2.34. Basic Mach-Zehnder interferometer configuration.

A commercial version of this interferometer configuration, manufactured by Wyko Corp., uses a pinhole in one of the branches to generate a perfect wavefront, in order to test the wavefront quality of the light source, as shown in Figure 2.35. (Leung and Lange, 1983; Creath, 1987). A Mach-Zehnder has also been used to test off-axis paraboloids (Gerth et al., 1978). Cuadrado et al. (1987) have described a method to align a Mach-Zehnder interferometer using equilateral hyperbolic zone plates and Flack (1978) has analyzed the errors that result from a test section misalignment.

Sometimes, to save one mirror, a triangular configuration as illustrated in Figure 2.36 is used.

2.7.2. Oblique Incidence Interferometers

Another kind of two beam interferometers have triangular paths, so that one of the beams is obliquely reflected on the flat surface under test. It may easily be proved that under those conditions a small error with height h on the surface under test introduces an error equal to $2h \cos \theta$, where θ is the incidence angle. Thus, the interferometer is desensitized by a factor $\cos \theta$. Another consequence of the oblique incidence is that the reflectivity of the surface under test is greatly increased. Thus, an interferometer with oblique incidence is ideally suited for testing ground or mate flat surfaces,

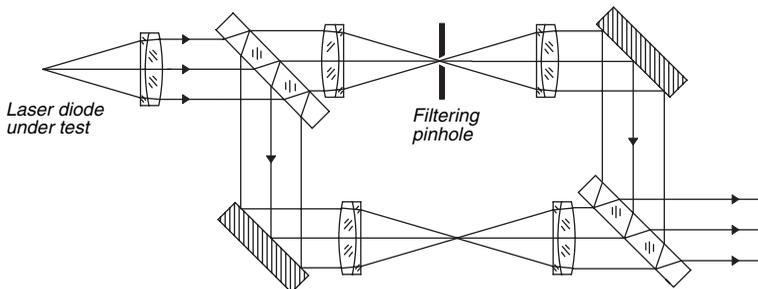


FIGURE 2.35. Mach-Zehnder interferometer used to test the wavefront quality of a laser diode.

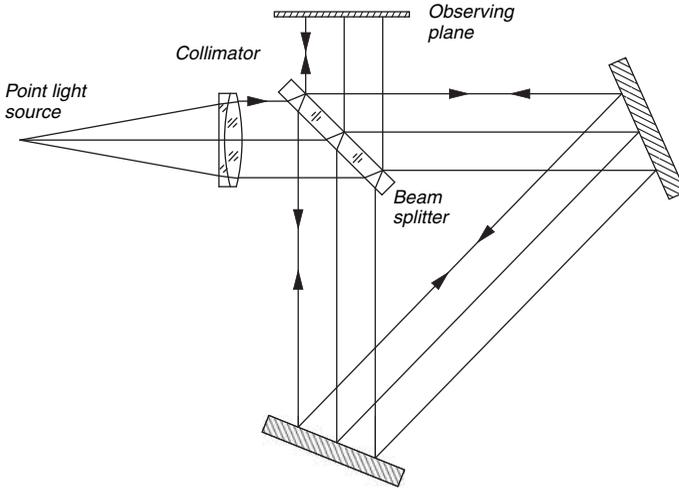


FIGURE 2.36. Triangular path interferometer.

whose flatness and reflectivity are not good enough to be tested by conventional interferometry.

Oblique incidence interferometers have been described by Linnik (1942), Saunders and Gross (1959), by Birch (1973,1979), by Hariharan (1975), and by MacBean (1984). Some of these interferometers use diffraction gratings as beam splitters, as the one designed by Hariharan, as shown in Figure 2.37. Small ground and almost flat aspherical surfaces may be tested with oblique incidence interferometers as shown by Jones (1979).

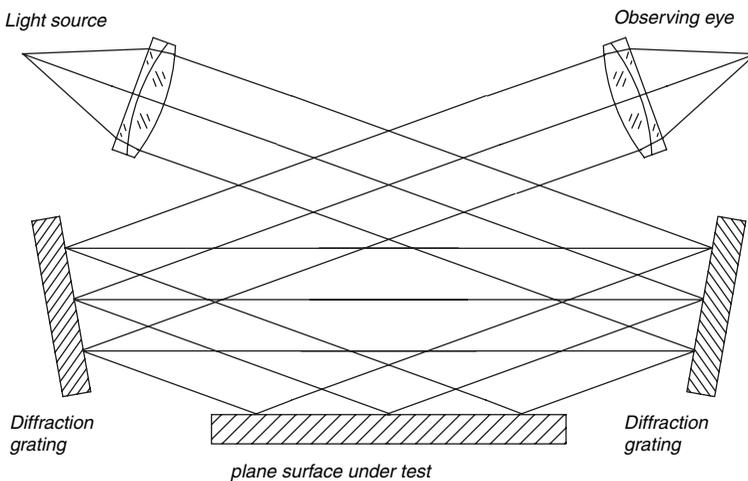


FIGURE 2.37. Grazing incidence Interferometer using a diffraction grating beam splitter.

2.8. VARIATIONS FROM THE TWYMAN–GREEN CONFIGURATION

Many variations of the Twyman–Green interferometer have been invented; interesting among them are some small, compact interferometers designed by Van Heel and Simons (1967) and by Basile (1979).

A carbon dioxide laser has been used as a light source of long wavelengths (10.67 μm) in a Twyman–Green interferometer (Munnerlyn et al., 1969, Kwon et al., 1979, and Lewandowski et al., 1986) in order to measure unpolished or rough surfaces.

Another interesting development is the invention of a cheap interferometer whose defects are corrected by means of a hologram (Rogers, 1970). First, a photograph is taken in a very imperfect and inexpensive interferometer, introducing a large tilt. This gives rise to an exceedingly fine set of fringes invisible to the naked eye but capable of being photographed. Then a second exposure is taken on the same photographic plate, after introducing into the interferometer the plate to be tested. A moiré pattern appears on the developed plate, giving the contours of the surface quality of the plate under test.

Another interesting holographic Twyman–Green interferometer has been described by Chen and Breckinridge (1982). In this design, a single holographic optical element combines the functions of a beam splitter, beam diverger, and null compensating lens. McDonell and DeYoung (1979) designed a large aperture interferometer using a holographic compensator.

2.8.1. Multiple Image Interferometers

An application of the systems using polarized light in the interferometer is to obtain Twyman–Green multiple simultaneous fringe patterns with different values of the constant phase difference by selecting the phase with the orientation of the analyzer. The interferometers described in Section 2.2.4 can produce two complementary interferograms as illustrated in Figure 2.38. If the two light beams returning to the light source are observed with another polarizing beam splitter, we will have four different fringe patterns with four different phases. Interferometers like this had been made with four independent cameras, but they are difficult to align. Another system that does not have this problem, described by Millerd, Brock, Hayes, Kimbrough, Novak, and North-Morris (2005), uses a CCD detector with a specially designed screen in front of it as illustrated in Figure 2.39, where the pixels in the CCD camera have analyzers oriented in four different directions to capture the four phase shifted interferogram simultaneously, with a holographic element that projects the same image region to a set of every four pixels.

2.8.2. Interferometers with Diffractive Beam Splitters

Some two beam interferometers may be thought as modifications from the basic Twyman–Green interferometer. Some of them use diffraction gratings as beam splitters, as the one shown in Figure 2.40, described by Molesini et al. (1984). The

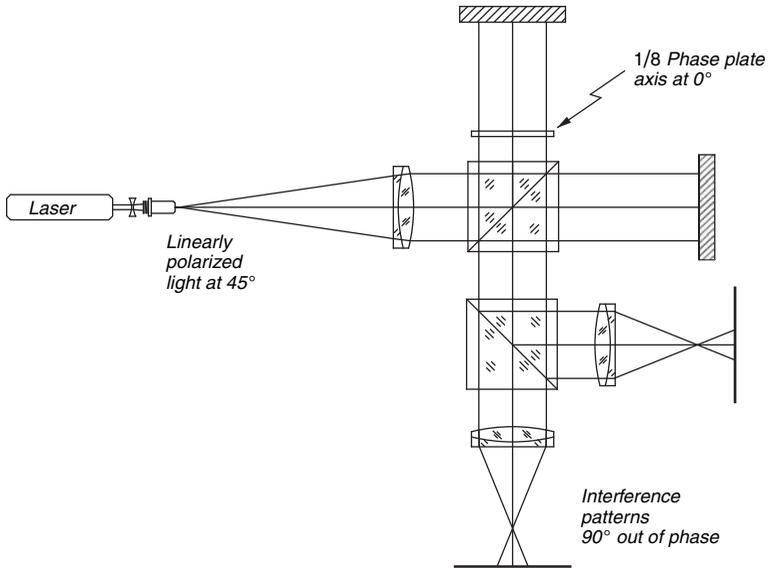


FIGURE 2.38. Twyman-Green interferometer with a cube nonpolarizing beam splitter.

advantage of this particular configuration is that relatively large errors in the grating flatness may be tolerated. In the following section we will see some other interferometers using diffraction gratings as beam splitters.

2.8.3. Phase Conjugating Interferometer

Phase conjugating mirrors are very useful tools in interferometry. They eliminate the need for a perfect reference wavefront. A Twyman-Green interferometer as shown in

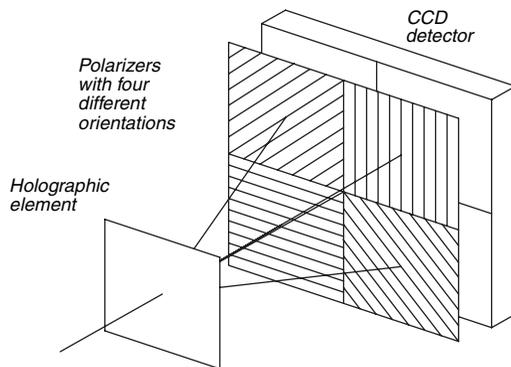


FIGURE 2.39. Arrangement to simultaneously produce for interferograms with different phase differences using a pixelated CCD detector with polarizing element in front of it.

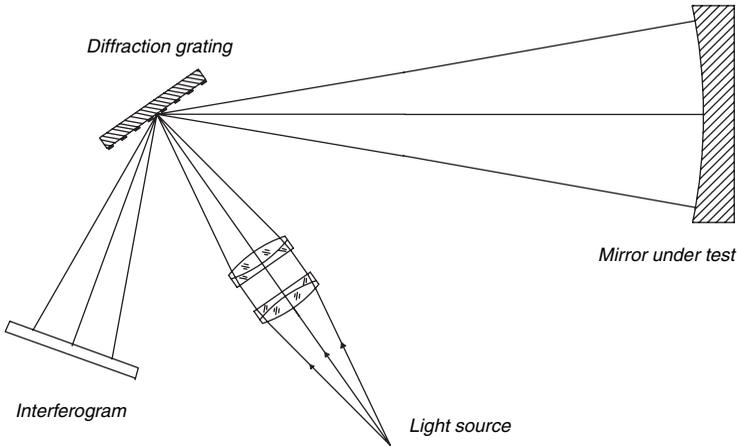


FIGURE 2.40. Oblique incidence interferometer using reflecting diffraction gratings as beam splitters.

Figure 2.41, using a phase conjugating mirror has been described by Feinberg (1983) and Howes (1986a, 1986b). The phase conjugating mirror is formed by a BaTiO_3 crystal, with the C axis parallel to one of its edges and inclined 20° with respect to a plane perpendicular to the optical axis. The phase conjugation is obtained by four wave mixing. These pumping beams are automatically self-generated from a 30-mW argon laser ($\lambda = 514.5 \text{ nm}$) incident beam by internal reflection at the crystal faces. Thus, it is a self-pumped phase conjugating mirror.

The property of this self-conjugating mirror is that the wavefront incident to the mirror is reflected back along the same ray directions that the incident wavefront has. Thus, the wavefront deformations change sign. Since the returning rays have the same directions as the incident rays, the quality of the focusing lens is not important.

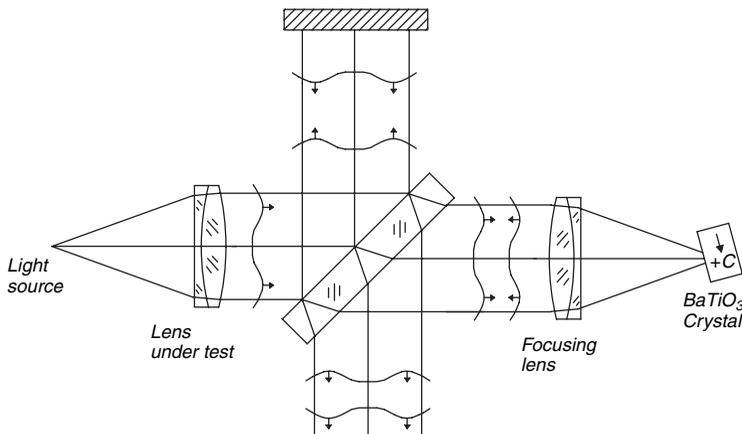


FIGURE 2.41. Twyman-Green Interferometer with a phase conjugating mirror.

However, the quality of the light source collimator is important. Any wavefront distortions produced by this collimator will appear in the final interferogram, but duplicated. In other words, the wavefront is not tested against a flat reference but against another wavefront with deformations opposite in sign.

Then the lens under test is the collimator and the sensitivity is the same as that in the common Twyman–Green interferometer, but with only a single pass through the lens. The advantage is that no perfect lenses are necessary. The disadvantage is that an argon laser is required.

2.9. TWYMAN–GREEN INTERFEROGRAMS AND THEIR ANALYSIS

The interferograms due to the primary aberrations can be described by using the wavefront function by Kingslake (1925–1926), which is given by

$$\text{OPD} = A + Bx + Cy + D(x^2 + y^2) + E(x^2 + 3y^2) + Fy(x^2 + y^2) + G(x^2 + y^2)^2, \quad (2.49)$$

where these coefficients represent:

<i>A</i>	Constant (piston) term
<i>B</i>	Tilt about the <i>y</i> axis
<i>C</i>	Tilt about the <i>x</i> axis
<i>D</i>	Reference sphere change, also called defocus
<i>E</i>	Sagittal astigmatism along the <i>y</i> axis
<i>F</i>	Sagittal coma along the <i>y</i> axis
<i>G</i>	Primary spherical aberration

In polar coordinates (ρ, θ) , Eq. 2.36 can also be written $(x = \rho \cos \theta; y = \rho \sin \theta)$ as

$$\text{OPD} = A + B\rho \cos \theta + C\rho \sin \theta + D\rho^2 + E\rho^2(1 + 2 \sin^2 \theta) + F\rho^3 \sin \theta + G\rho^4, \quad (2.56)$$

This expression is designed to represent the wavefronts produced in the presence of the primary aberrations of a centered lens whose point source and image are displaced in the *y* direction. Thus, the wavefront is always symmetric about the *y* axis. Also, the coma and astigmatism terms are referred to the Petzval surface, which is not of a great relevance in most interferograms. When testing an optical surface or a decentered system no symmetry can, in general, be assumed and a more general wavefront representation has to be considered.

Additionally, it is convenient for the mathematical analysis that the average tilt of all aberrations is zero with the exception of the two tilts. This is equivalent to selecting the optimum tilts of the reference wavefront for each aberration. Also, the average curvature of all aberrations must be zero for all aberrations, with the

exception of the spherical curvature, also called defocus. This is equivalent to selecting the optimum value of the focus setting for each aberration. These aberrations are the Zernike polynomials to be described with detail in Chapter 13. In terms of these aberrations, the wavefront shape up to the fourth order terms can be written as

$$\begin{aligned}
 \text{OPD} = & A_1 + A_2x + A_3y + A_4(x^2 + y^2 - 0.5) + A_5xy + A_6(x^2 - y^2) \\
 & + A_7y(3x^2 + 3y^2 - 2) + A_8x(3x^2 + 3y^2 - 2) + A_9y(3x^2 - y^2) \\
 & + A_{10}x(x^2 - 3y^2) + A_{11}[6(x^2 + y^2)^2 + 6(x^2 + y^2) + 1] \\
 & + A_{12}(x^2 - y^2)(4x^2 + 4y^2 - 3) + A_{13}xy(4x^2 + 4y^2 - 3) \\
 & + A_{14}[(x^2 - y^2)^2 - 8x^2y^2] + A_{15}xy(x^2 - y^2),
 \end{aligned} \tag{2.49}$$

or in polar coordinates

$$\begin{aligned}
 \text{OPD} = & A_1 + A_2\rho \cos \theta + A_3\rho \sin \theta + A_4(\rho^2 - 1) + A_5\rho^2 \sin 2\theta \\
 & + A_6\rho^2 \cos 2\theta + A_7\rho(3\rho^2 - 2) \sin \theta + A_8\rho(3\rho^2 - 2) \cos \theta \\
 & + A_9\rho^3 \sin 3\theta + A_{10}\rho^3 \cos 3\theta + A_{11}(6\rho^4 - 6\rho^2 + 1) \\
 & + A_{12}\rho^2(4\rho^2 - 3) \cos 2\theta + A_{13}\rho^2(4\rho^2 - 3) \sin 2\theta \\
 & + A_{14}\rho^4 \cos 4\theta + A_{15}\rho^4 \sin 4\theta,
 \end{aligned} \tag{2.50}$$

where

A_1	Constant (Piston) term
A_2	Tilt about the y axis
A_3	Tilt about the x axis
A_4	Spherical term, also called defocus
A_5	Astigmatism with axis at $\pm 45^\circ$
A_6	Astigmatism with axis at 0° or 90°
A_7	Third order coma along y axis
A_8	Third order coma along x axis
A_9	Triangular astigmatism with base parallel to x axis
A_{10}	Triangular astigmatism with base parallel to y axis
A_{11}	Primary spherical aberration
A_{12}	High order astigmatism at 0° or 90°
A_{13}	High order astigmatism at $\pm 45^\circ$
A_{14}	Quadrangular (ashtray) astigmatism 0° or 90°
A_{15}	Quadrangular (ashtray) astigmatism at $\pm 45^\circ$.

In computing interferograms, a normalized entrance pupil with unit semidiameter ρ can be assumed. The great advantage of this normalization is that a value of all the aberration coefficients will represent the same maximum wavefront deformation at the edge of the pupil.

The relative simplicity of the Kingslake expression allows us an easy and intuitive analysis of the interferograms, as we will see with some examples. The interferograms for some aberrations were simulated by calculating the irradiance at a two-dimensional array of points. A wavelength equal to 632.8 nm was used in these interferograms, the pupil diameter is 20.0 mm but the values of the coefficients are defined for a normalized pupil ($\rho = 1$).

1. *Perfect lens.* The patterns for a perfect lens without tilts ($B = C = 0$) and with tilt ($B = 5.0 \times 10^{-3}$) are shown in Figures 2.42(a,b). A perfect lens with defocusing ($D = 3.0 \times 10^{-3}$) and with defocusing and tilt ($D = 3.0 \times 10^{-3}$, $B = 5.0 \times 10^{-3}$) is illustrated in Figures 2.42(c,d).

2. *Spherical aberration.* The patterns for pure spherical aberration were computed assuming that $G = 5.0 \times 10^{-3}$. They are shown at the paraxial focus ($D = 0$), without tilts ($B = C = 0$) and with tilt ($B = 5.0 \times 10^{-3}$) in Figures 2.43(a,d). The patterns at the marginal focus are obtained by setting in Eq. (2.43), only A and D different from zero,

$$\frac{d\text{OPD}}{d\rho} = 4G\rho^3 + 2B\rho = 0. \quad (2.57)$$

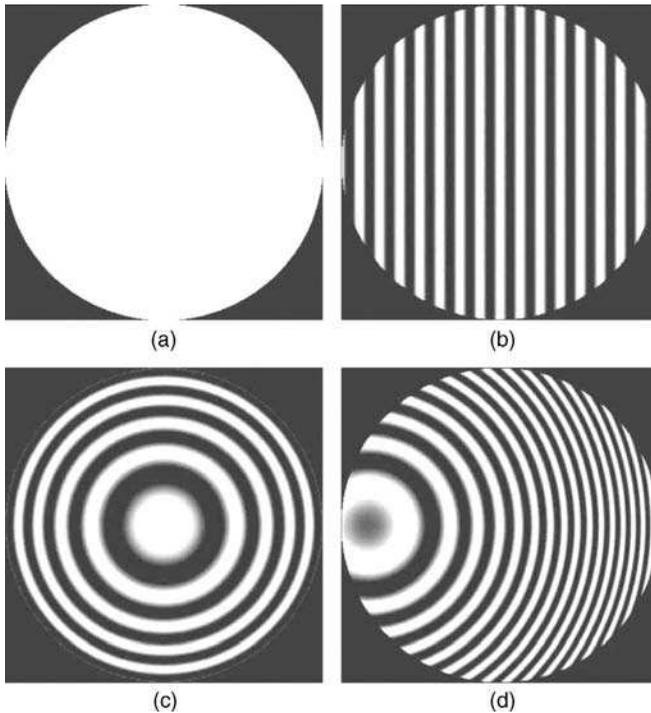


FIGURE 2.42. Interferograms for a perfect lens. (a) With no tilt or defocusing. (b) With tilt. (c) With defocusing. (d) With tilt and defocusing.

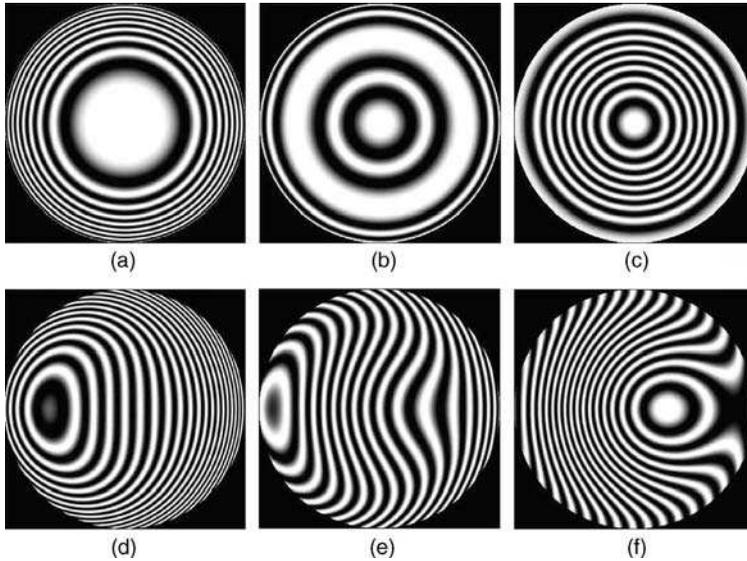


FIGURE 2.43. Interferograms for a lens with spherical aberration at the paraxial, medium, and marginal foci.

Therefore, we set the defocusing coefficient $B = -5.0 \times 10^{-3}$ and the spherical aberration coefficient $G = 5.0 \times 10^{-3}$. These interferograms without ($B = C = 0$) and with ($B = 5.0 \times 10^{-3}$) tilt are shown in Figures 2.43(c,f). The fringe patterns at the medium focus with $B = -10.0 \times 10^{-3}$ are in Figures 2.43(b,e).

3. *Coma*. All the patterns for coma were obtained using $F = 5.0 \times 10^{-3}$. Figure 2.44 shows them for the paraxial focus ($D = 0$) and Figure 2.45 with a small defocusing ($D = 5.0 \times 10^{-3}$). In both figures the central pattern has no tilt ($E = F = 0$) and the surrounding pictures are for different tilt combinations ($B = \pm 5.0 \times 10^{-3}, C = \pm 5.0 \times 10^{-3}$).

4. *Astigmatism*. All the patterns for astigmatism were computed for $C = 3.0 \times 10^{-3}$. If $D = 0$, we obtain the Petzval focus. The OPD for astigmatism can be written from Eq. (2.36) as

$$\text{OPD} = (D + E)x^2 + (D + 3E)y^2. \quad (2.58)$$

Therefore, the sagittal focus is obtained for $D + E = 0$ and the tangential focus for $D + 3E = 0$. The medium focus is obtained for $D + E = -(D + 3E)$; hence $D = -2E$.

Figure 2.46 shows the patterns at the Petzval focus with tilts in all directions ($B = \pm 5.0 \times 10^{-3}, C = \pm 5.0 \times 10^{-3}$). Figures 2.47–2.49 show the patterns at the sagittal, medium, and tangential foci, respectively, also with tilts in all directions.

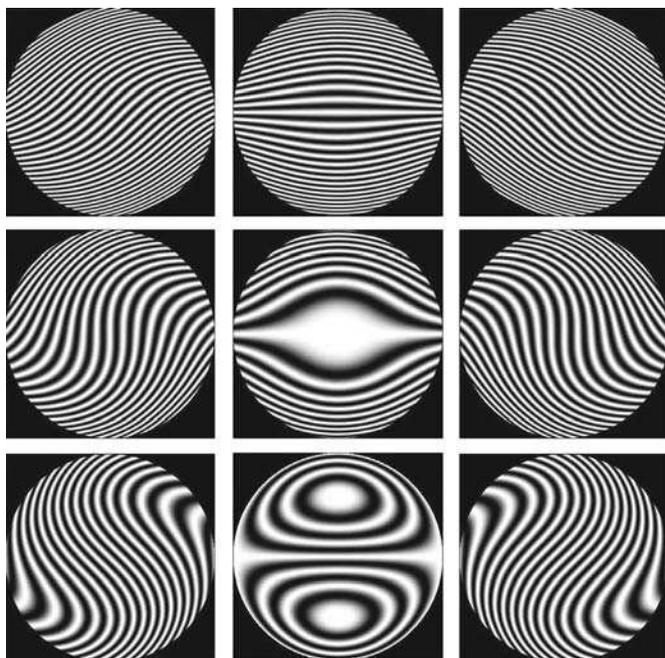


FIGURE 2.44. Interferograms for a lens with coma at the paraxial focus.

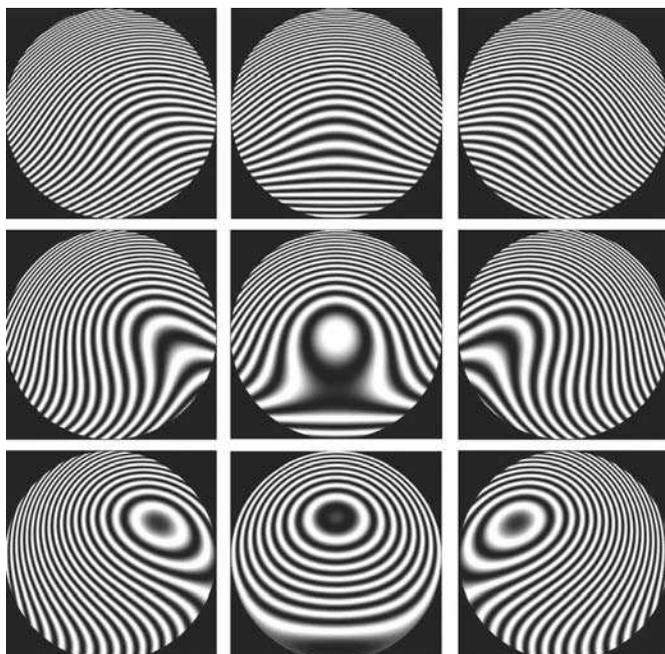


FIGURE 2.45. Interferograms for a lens with coma and a small defocusing.

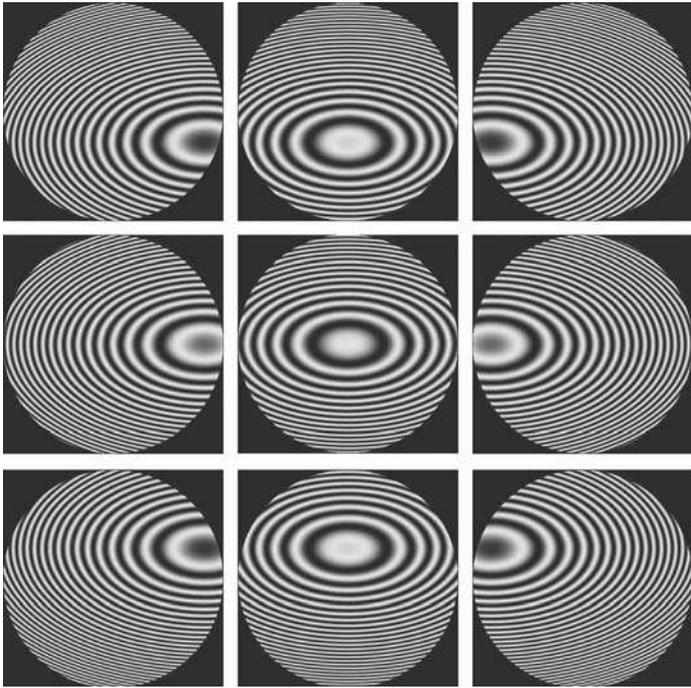


FIGURE 2.46. Interferograms for a lens with astigmatism at the Petzval focus.

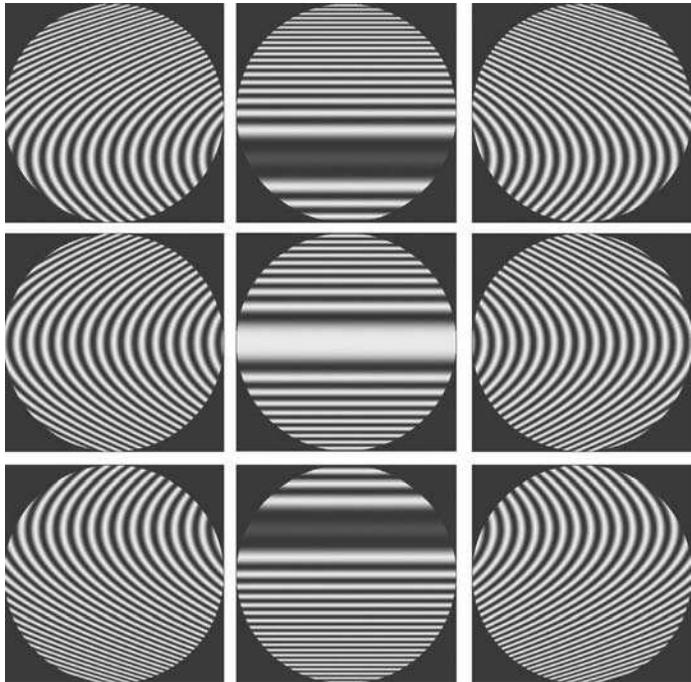


FIGURE 2.47. Interferograms for a lens with astigmatism at the sagittal focus.

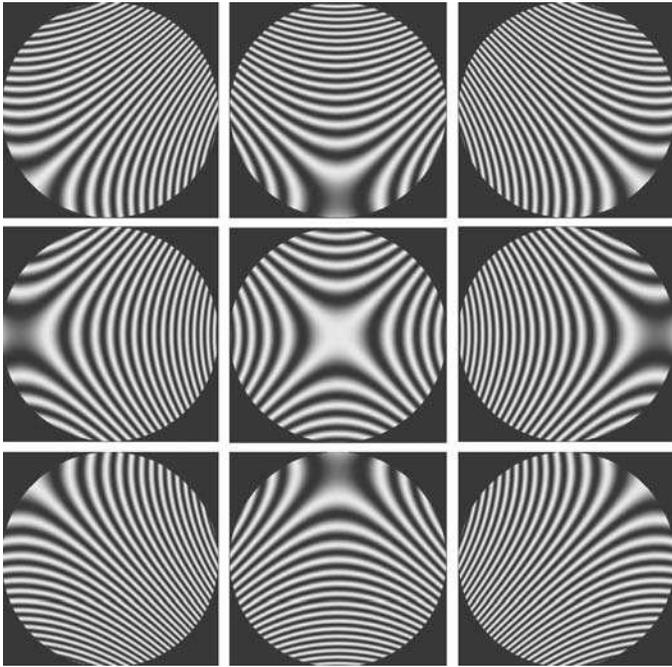


FIGURE 2.48. Interferograms for a lens with astigmatism at the best focus.

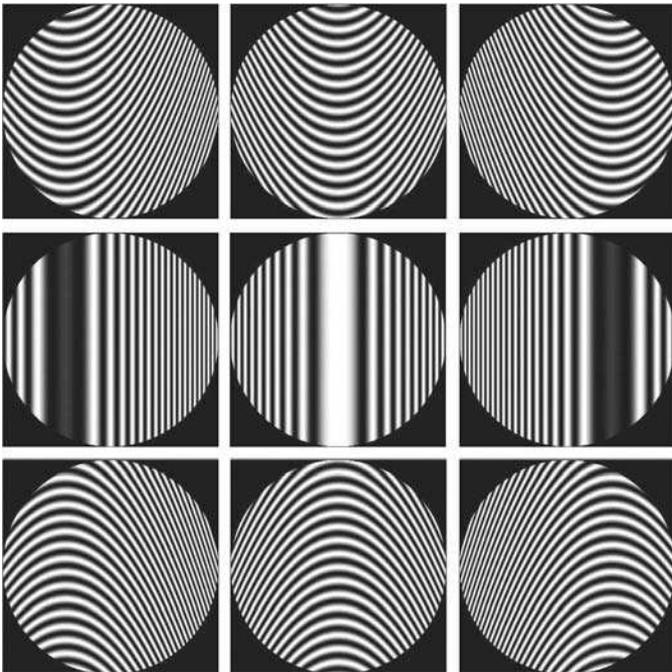


FIGURE 2.49. Interferograms for a lens with astigmatism at the tangential focus.

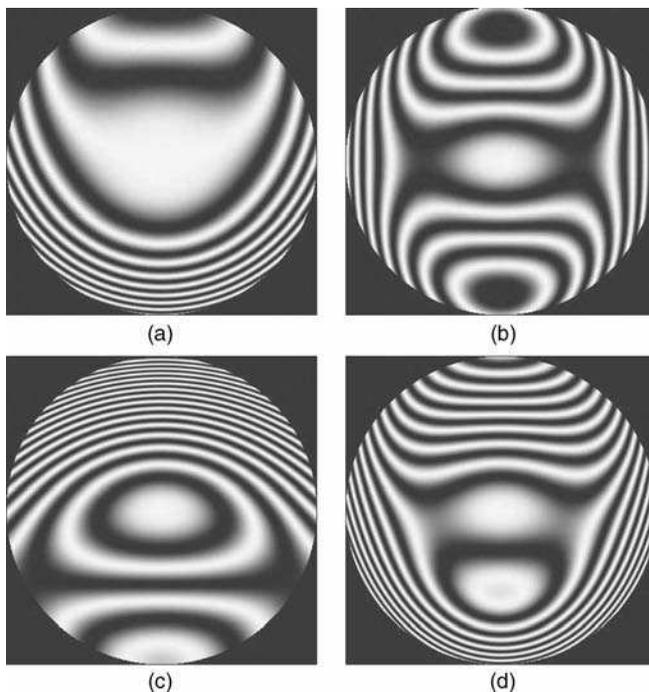


FIGURE 2.50. Interferograms for a lens with combined aberrations. (a) Spherical aberration with coma. (b) Spherical aberration with astigmatism. (c) Coma with astigmatism. (d) Spherical aberration with coma and astigmatism.

5. *Combined Aberrations.* Figure 2.50 shows the patterns for combined aberrations: spherical aberration plus coma ($G = 2.0 \times 10^{-3}$ and $F = 3.0 \times 10^{-3}$) in Figure 2.50(a), spherical aberration plus astigmatism ($G = 4.0 \times 10^{-3}$ and $E = -2.0 \times 10^{-3}$) in Figure 2.50(b), coma plus astigmatism ($F = -2.0 \times 10^{-3}$, $E = 4.0 \times 10^{-3}$) in Figure 2.50(c), and, finally, spherical aberration plus coma plus astigmatism ($G = 5.0 \times 10^{-3}$, $F = -2.0 \times 10^{-3}$, $E = 4.0 \times 10^{-3}$) in Figure 2.50(d).

Pictures of typical interferograms are shown in a paper by Marechal and Dejonc (1950). These interferograms can be simulated by beams of fringes of equal inclination on a Michelson interferometer (Murty, 1960) using the OPDs introduced by a plane parallel plate and cube corner prisms instead of mirrors or by electronic circuits on a CRT (Geary et al., 1978 and Geary, 1979).

This type of interferogram was first analyzed by Kingslake (1926–1927). He measured the OPD at several points on the x and y axes just by fringe counting. Then, solving a system of linear equations, he computed the OPD coefficients B , C , D , E , F , G . Another method for analyzing a Twyman–Green, interferogram was proposed by Saunders (1965). He found that the measurement of four appropriately chosen points is sufficient to determine any of the three primary aberrations. The

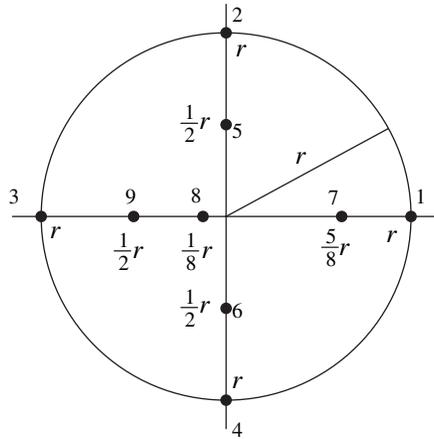


FIGURE 2.51. Distribution of reference points for evaluation of primary aberrations.

points were selected as in Figure 2.51 and then the aberration coefficients were computed as

$$A = \frac{128}{81r^3} [P_1 - P_9 + 2(P_8 - P_7)], \tag{2.59}$$

$$B = \frac{2}{3r^3} [P_2 - P_4 + 2P_6 - 2P_5], \tag{2.60}$$

$$C = \frac{1}{4r^2} [P_2 + P_4 - P_1 - P_3], \tag{2.61}$$

where P_i is the interference order at point i .

If a picture of the interferogram is not taken, the aberration coefficients can be determined by direct reading on the interferogram setting, looking for interference patterns with different foci and tilts (Perry, 1923-1924). To make these readings easier, some optical arrangements may be used to separate symmetrical and asymmetrical wavefront aberrations as shown by Hariharan and Sen (1961).

2.9.1. Analysis of Interferograms of Arbitrary Wavefronts

The problem of determining the shape of a wavefront with arbitrary shape from a single Twyman–Green interferogram has been considered very generally and briefly by many authors, for example, by Berggren (1970) and more completely by Rimmer et al. (1972). The procedure consists in measuring the positions of the fringes at many points over the interferograms and taking readings of the position (x, y) and the order of interference m . Since the measurements are taken at a limited number of points, an interpolation procedure must be adopted. The interpolation may be performed by the use of splines or a least squares procedure with polynomial fitting as described in detail in Chapter 16.

If the wavefront is smooth enough, it is very convenient to express the final wavefront $W(x, y)$ in terms of a linear combination of Zernike polynomials. Then the process of removing or adding defocussing or tilts becomes much simpler. With the final results, it is an extremely simple matter to plot level maps of the wavefront. Further details can be found in the book by Malacara, Servín and Malacara (2005).

REFERENCES

- Adachi I., T. Masuda, and S. Nushiya, "A Testing of Optical Materials by the Twyman Type Interferometer," *Atti. Fond. Giorgio Ronchi Contrib. Ist. Naz. Ottica*, **16**, 666 (1961).
- Adachi I., T. Masuda, T. Nakata, and S. Nushiya, "The Testing of Optical Materials by the Twyman Type Interferometer III," *Atti. Fond. Giorgio Ronchi Contrib. Ist. Naz. Ottica*, **17**, 319 (1962).
- Balhorn R., H. Kunzmann, and F. Lewosky, "Frequency Stabilization of Internal Helium-Neon Lasers," *Appl. Opt.*, **11**, 742 (1972).
- Basile G., "Two-Wave Interferometers for Use with a Monochromatic Source," *Appl. Opt.*, **18**, 422, (1979).
- Batishko C. R. and R. R. Shannon, "Problem in Large-Path Difference Laser Interferometry;" *Appl. Opt.*, **11**, 195 (1972)
- Bennett S. J., R. E. Ward, and D. C. Wilson, "Comments on Frequency Stabilization of Internal Mirror He-Ne Lasers," *Appl. Opt.*, **12**, 1406 (1973).
- Berggren R., "Analysis of Interferograms," *Opt. Spectra.*, **4**, 22 (1970).
- Birch K. G., "Oblique Incidence Interferometry Applied to Non-Optical Surfaces," *J. Phys. E.*, **6**, 1045 (1973).
- Birch K. G., "Interferometric Examination of Lenses and Other Components," *Proc. SPIE*, **163**, 112, (1979).
- Briers J. D., "Interferometric Testing of Optical Systems and Components: A Review," *Opt. Laser Technol.*, **4**, 28 (1972).
- Bruning J. and D. R. Herriott, "A Versatile Laser Interferometer," *Appl. Opt.*, **9**, 2180 (1970).
- Bruning J. H., D. R. Herriott, J. E. Gallagher, D. P. Rosenfeld, A. D. White and D. J. Brangaceio, "Digital Wavefront Measuring Interferometer for Testing Surfaces and Lenses," *Appl. Opt.* **13**, 2693–2703 (1974).
- Buin A. P., M. P. Semenova, and Kriuyukhina, "Inspection of the Surface Quality of Large Scale Optical Components of an Unequal Arm Interferometer," *Sov. J. Opt. Technol.*, **36**, 720 (1969).
- Burch C. R., "The Williams Interferometer," *Monthly Not. R. Soc.*, **100**, 488 (1940).
- Candler C., *Modern Interferometers*, Hilger and Watts, London, 1951, Chaps. 6 and 7.
- Chen C. W. and J. B. Breckinridge, "Holographic Twyman–Green Interferometer," *Appl. Opt.*, **21**, 25-63 (1982).
- Cole G. C., J. H. Burge, and L. R. Dettmann, "Vibration Stabilization of a Phase Shifting Interferometer for Large Optics," *Proc. SPIE*, **3134**, 438–446 (1997).
- Collier R. T., C. B. Brukhardt, and L. H. Lin, *Optical Holography*, Academic Press, New York, 1971, p. 146.

- Connes P., "Aumentation du Produit Luminosite X resolution des Interferometers par l' Emploi d' une Difference de Marche Independente de l' Incidence," *Rev. Opt.*, **35**, 37 (1956).
- Creath K., "Wyko Systems for Optical Metrology," *Proc. SPIE*, **816**, 111 (1987).
- Creath K., "Testing Spherical Surfaces: A Fast, Quasi-Absolute Testing," *Appl. Opt.*, **31**, 4350-4354 (1992).
- Creath K. and J. C. Wyant, "Testing Spherical Surfaces: A Fast, Quasi-Absolute Technique," *Appl. Opt.*, **31**, 4350-4354 (1992).
- Cuadrado Javier M., M. Va. Pérez, C. Gómez-Reino, "Equilateral Hiperbolic Zone Plates: Their use in the Alignment of a Mach-Zehnder Interferometer," *Appl. Opt.*, **26**, 1527 (1987).
- De Vany, A. S., "On Using a Williams Interferometer for Making a Divider Plate," *Appl. Opt.*, **4**, 365 (1965).
- Dorband B. and H. J. Tiziani, "Auslegung von Kompensationsystemen zur Interferometrischen Prüfung Asphärischer Flächen," *Optik*, **467**, 1, (1) (1984).
- Dyson J., "Unit Magnification Optical System without Seidel Aberrations," *J. Opt. Soc. Am.*, **49**, 713 (1959).
- Fantone S. D., "Simple Method for Testing an Axicon," *Appl. Opt.*, **20**, 3685 (1981).
- Feinberg J., "Interferometer with a Self-Pumped Phase-Conjugating Mirror," *Opt. Lett.*, **8**, 569 (1983).
- Ferguson T. R., "Polarization Effects in Interferograms of Conical Optical Elements," *Appl. Opt.*, **21**, 514 (1982).
- Flack Jr., Ronald D., "Mach-Zehnder Interferometer Errors Resulting from Test Section Misalignment," *Appl. Opt.*, **17**, 985 (1978).
- Geary J. M., "Real Time Interferogram Simulation," *Opt. Eng.*, **18**, 39 (1979).
- Gerth H. L., R. E. Sladky, M. J. Besik, and C. A. Washington, "Fabrication of Off Axis Parabolic Mirrors," *Opt. Eng.*, **17**, 588 (1978).
- Gordon S. K. and S. F. Jacobs, "Modification of Inexpensive Multimode Lasers to Produce a Stabilized Single Frequency Beam," *Appl. Opt.*, **13**, 231 (1974).
- Grigull V. and H. Rottenkolber, "Two Beam Interferometer Using a Laser," *J. Opt. Soc. Am.*, **57**, 149 (1967).
- Guild J., "Fringe Systems in Uncompensated Interferometers," *Proc. Phys. Soc.*, **33**, 40 (1920-1921).
- Hansen G., "Die Sichtbarkeit der Interferenzen beim Twyman-Interferometer," *Zeiss-Nachrichten*, **4**, 109-121 (1942).
- Hansen G., "Die Sichtbarkeit der Interferenzen beim Twyman-Interferometer," *Optik*, **12**, 5-16 (1955).
- Hansen G. and W. Kinder, "Abhängigkeit des Kontrastes der Fizea-Streifen im Michelson-Interferometer vom Durchmesser der Aperturblende," *Optik*, **15**, 560-564 (1958).
- Hansen G., "On Twyman-Interferometers," *Optik*, **67**, 79 (1984).
- Hardesty C., "Vibration Insensitive Laser Unequal Path Interferometer (LUPI) Test," *Proc. SPIE*, **192**, 93 (1979).
- Hariharan P., "Improved Oblique Incidence Interferometer," *Opt. Eng.*, **14**, 257 (1975).

- Hariharan P. and D. Sen, "The Separation of Symmetrical and Asymmetrical Wavefront Aberrations in the Twyman Interferometer," *Proc. Phys. Soc.*, **77**, 328 (1961).
- Hopkins R. E., "Re-Evaluation of the Problem of Optical Design," *J. Opt. Soc. Am.*, **52**, 1218 (1962).
- Horne D. F., *Optical Production Technology*, Adam Hilger, London, 1972, and Crane Russak, New York, 1972, Chap. 11.
- Houston J. B., Jr., C. J. Buccini, and P. K. Neill, "A Laser Unequal Path Interferometer for the Optical Shop," *Appl. Opt.*, **6**, 1237 (1967).
- Howes W. L., "Lens Collimation and Testing Using a Twyman-Green Interferometer with a Self-Pumped Phase-Conjugation Mirror," *Appl. Opt.*, **25**, 473 (1986a).
- Howes W. L., "Large Aperture Interferometer with Phase-Conjugate Self-Reference Beam," *Appl. Opt.*, **25**, 3167 (1986b).
- Jarosewics Z., "Interferometric Testing of the Spacing Error of a Plane Diffraction Grating," *Opt. Commun.*, **60**, 345 (1986).
- Jones R. A., "Fabrication of Small Nonsymmetrical Aspheric Surfaces," *Appl. Opt.*, **18**, 1244 (1979).
- Jensen A. E., "Absolute calibration Method for Twyman-Green Wavefront Testing Interferometers," *J. Opt. Soc. Am.*, **63**, 1313A (1973).
- Kingslake R., "The Interferometer Patterns Due to the Primary Aberrations," *Trans. Opt. Soc.*, **27**, 94 (1925-1926).
- Kingslake R., "The Analysis of an Interferogram," *Trans. Opt. Soc.*, **28**, 1 (1926-1927).
- Kocher D. G., "Twyman-Green Interferometer to Test Large Aperture Optical Systems," *Appl. Opt.*, **11**, 1872-1873 (1972).
- Kowalik W., "Interference Measurement of Continuous Heterogeneities in Optical Materials," *Appl. Opt.*, **17**, 2956 (1978).
- Kwon O. Y., J. C. Wyant, and C. R. Hayslett, "Long-Wavelength Interferometer in the Optical Shop," *Proc. SPIE*, **192**, 88 (1979).
- Linnik V. P., "An Interferometer for Controlling Large Mechanical Details," *Comptes Rendus (Doklady) de l'Academie des Sciences de l'URSS*, **35**, 16 (1942).
- Leung K. M. and S. Lange, "Wavefront Evaluation on Laser Diodes Using a Phase Measurement Interferometer," *Proc. SPIE*, **429**, 27 (1983).
- Lewandowski J., B. Mongeau, M. Cormier, and J. La-Pierre, "Infrared Interferometers at 10.7 μ m," *J. Appl. Phys.*, **66**, 132 (1986).
- Luneburg, R. K., *Mathematical Theory of Optics*, University of California, Press, Berkeley, 1964, Appendix 2, P. 372.
- MacBean M. D. A., "Oblique Incidence Interferometry of Rough Surfaces Using a Novel Dove-Prism Spectrometer," *Appl. Opt.*, **23**, 4024 (1984).
- Malacara D. and C. Menchaca, "Imaging of the Wavefront Under Test in Interferometry," *Proc. SPIE*, **540**, 34 (1985).
- Malacara D., M. Servin and Z. Malacara, *Interferogram Analysis for Optical Testing*, 2nd Ed., Taylor and Francis, Boca Raton FL, 2005.
- Marechal A. and P. Dejonc, "Quelques Aspects de Franges de Twyman" (Some Aspects of Twyman Fringes), *Rev. Opt. Theor. Instrum.*, **29**, 430 (1950).
- Martin L. C. and R. Kingslake, "The Measurement of Chromatic Aberration on the Hilger Lens Testing Interferometer," *Trans. Opt. Soc.*, **25**, 213 (1923-1924).

- Masuda T., S. Nishiyama, T. Nakata, and I. Adachi, "The Testing of Optical Materials by the Twyman Type Interferometer. II," *Atti Fond. Giorgio Ronchi Contrib. Ist. Naz. Ottica*, **17**, 197 (1962).
- McDonell M. M. and T. F. DeYoung, "Inexpensive Large Aperture Interferometer," *Proc. SPIE*, **192**, 145, (1979).
- Mertz L., International Commission for Optics Conference, Stockholm, 1959.
- Michelson A. A., "On the Correction of Optical Surfaces," *Astrophys J.*, **47**, 283 (1918).
- Millerd J., N. Brock, J. Hayes, B. Kimbrough, M. Novak, M. North-Morris, "Modern Approaches in Phase Measuring Metrology," *Proc. SPIE*, **5856**, 14-22 (2005).
- Molesini G., G. Pedrini, and F. Quercioli, "Laser Unequal Path Interferometer Configurations by Grating Splitting at the Fourier Plane," *Opt. Eng.*, **23**, 646 (1984).
- Morokuma T., K. F. Neftin, T. R. Lawrence, and T. M. Klucher, "Interference Fringes with Long Path Difference Using He-Ne Laser," *J. Opt. Soc. Am.*, **53**, 394 (1963).
- Munnerlyn C. K., M. P. Givens, and R. E. Hopkins, "Interferometric Measurement of Optically Rough Surfaces," *IEEE J. Quantum Electron.*, **QE-5**, 359 (1969).
- Murty M. V. R. K. and D. Malacara-Hernández "Some Applications of the Gas Laser as a Source of Light for the Testing of Optical Systems," *Jap. J. Appl. Phys.*, **4**, Supplement 1, 106-112 (1965).
- Ostrovskaya M. A. and F. M. Filimonova, "Use of the Gas Laser for Interferometric Quality Control in Telescope Manufacture," *Sov. J. Opt. Technol.*, **36**, 563-565 (1969).
- Parmigiani F., "Phase Dependence of Michelson Interferometer Outputs on The Absorbing Beam Splitter Thickness," *Opt. Commun.*, **38**, 319 (1981).
- Perry J. W., "The Determination of Aberrations, as Expressed in Geometrical Optics, from the Indications of the Hilger Interferometer," *Trans. Opt. Soc.*, **25**, 97 (1923-1924).
- Rimmer M. P., D. M. King, and D. G. Fox, "Computer Program for the Analysis of Interferometric Test Data," *Appl. Opt.*, **11**, 2790 (1972).
- Rogers G. L., "The Equivalent Interferometer in Holography," *Opt. Acta*, **17**, 257 (1970).
- Saunders J. B., "Precision Method for Evaluating Primary Aberrations of Lenses with a Twyman Interferometer," *J. Res. Nat. Bur. Stand.*, **69C**, 251 (1965).
- Saunders J. B. and F. L. Gross, "Interferometer for Large Surfaces," *J. Res. Nat. Bur. Stand.*, **62**, 137 (1959).
- Schwider J. and O. Falkenstörfer, "Twyman-Green Interferometer for Testing Microspheres," *Opt. Eng.*, **34**, 2972-2975 (1995).
- Schwider J., "Partially Coherent Illumination in Interferometry for Optical Testing" *Proc. SPIE*, **3745**, 2-13 (1999).
- Sinclair D. C. and W. E. Bell, *Gas Laser Technology*, Holt, Rinehart and Winston, New York, 1969, Chap. 5.
- Slevogt H., "Zur geometrischen Optik der Zweistrahl-Interferometer," (About the Geometrical Optics of Two-Beam Interferometers), *Optik*, **11**, 366 (1954).
- Steel W. H., "Adjustable Compensators for Two-Beam Interferometers," *Opt. Acta*, **9**, 111-119 (1962).
- Steel W. H., "The Compensation of a Williams Interferometer," *Opt. Acta*, **10**, 205-208 (1963).
- Steel W. H., "Two-Beam Interferometry," in *Progress in Optics*, Vol. 5, E. Wolf, Ed., North-Holland, Amsterdam, 1966, Chap. 3.

- Thomas D. A. and J. C. Wyant, “Determination of the Dihedral Angle Errors of a Corner Cube from its Twyman–Green Interferogram,” *J. Opt. Soc. Am.*, **67**, 467 (1977).
- Twyman F. and A. Green, *British Patent* (prisms and microscopes) 103832 (1916).
- Twyman F., “Correction of Optical Surfaces,” *Astrophys. J.*, **48**, 256 (1918a).
- Twyman F., *British Patent* (camera lens) 130224 (1919).
- Twyman F., “Interferometers for the Experimental Study of Optical Systems from the Point of View of the Wave Theory,” *Philosophical Magazine*, **35**, 49–58 (1918b).
- Twyman F., “The Testing of Microscope Objectives and Microscopes by Interferometry,” *Trans. Faraday Soc.*, **16**, 208 (1920).
- Twyman F., “An Interferometer for Testing Camera Lenses,” *Trans. Opt. Soc.*, **22**, 174 (1920-1921); also appeared in *Philos. Mag.*, **42**, 777 (1921).
- Twyman F., and A. J. Dalladay, “Variation in Refractive Index Near the Surfaces of Glass Melts,” *Trans. Opt. Soc.*, **23**, 131 (1921-1922).
- Twyman F., “The Hilger Microscope Interferometer,” *J. Opt. Soc. Am.*, **7**, 635–656 (1923).
- Williamson R., “Novel Interferometer,” *Proc. SPIE*, **5178**, 124–130 (2004).
- Zielinski R. J., “Unequal Path Interferometer Alignment and Use,” *Opt. Eng.*, **18**, 479 (1979).
- Zielinski R. J., “Unequal Path Interferometer Alignment and Use,” *Proc. SPIE*, **153**, 51 (1978).

3

Common-Path Interferometers

S. Mallick and D. Malacara

3.1. INTRODUCTION

In the general type of interferometer, such as the Twyman–Green or Mach–Zehnder, the reference and test beams follow widely separated paths and are, therefore, differently affected by mechanical shocks and temperature fluctuations. Thus, if suitable precautions are not taken, the fringe pattern in the observation plane is unstable and measurements are not possible. The problems are particularly acute when optical systems of large aperture are being tested. Most of the difficulty can be avoided by using the so-called common-path interferometers, in which the reference and test beams traverse the same general path. These interferometers have the additional advantage that they do not require perfect optical components (the master) of dimensions equal to those of the system under test for producing the reference beam. Furthermore, the path difference between the two beams in the center of the field of view is, in general, zero, making the use of white light possible.

In certain common-path interferometers, the reference beam is made to traverse a small area of the optical system under test and is, therefore, unaffected by system aberrations. When this beam interferes with the test beam, which has traversed the full aperture of the optical system, explicit information about the system defects is obtained. However, in most common-path interferometers both the reference and test beams are affected by the aberrations, and interference is produced by shearing one beam with respect to the other. The information obtained in this case is implicit and some computations are needed to determine the shape of the aberrated wavefront.

The beam splitting is brought about by amplitude division with the help of a partially scattering surface, a doubly refracting crystal, or a semireflecting surface. We consider a few examples of these instruments in this chapter.

*This chapter has been updated and a few sections have been included by the second author.

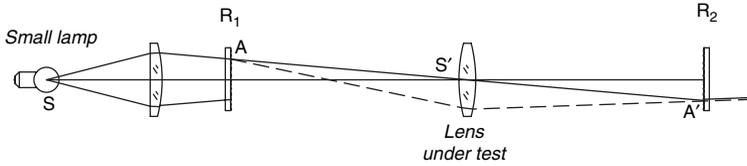


FIGURE 3.1. Burch interferometer testing a lens with magnification one, using two identical scatter plates.

3.2. BURCH’S INTERFEROMETER EMPLOYING TWO MATCHED SCATTER PLATES

This interferometer was first described by Burch (Burch 1953, 1962, 1969) and later reviewed by several authors, for example, by Rubin (1980). An arrangement to test a convergent lens working with unit magnification is illustrated in Figure 3.1. A lens forms an image of a small source S at S' on the lens. The magnification of the lens under test has to be extremely close to one. A partially scattering plate R_1 is located in front of a lens to be measured at twice its focal length. The lens under test forms an image of the scattering plate R_1 at another identical scatter plate R_2 , which is rotated 180° with respect to the first plate, so that there is a point to point coincidence between R_2 and the image of R_1 . A part of the light incident on the scatter plate R_1 passes through it without scattering and arrives at S' . Since this beam touches the lens only at a small region around S' , it is not affected by the errors of the lens surface. This beam acts as the reference beam. Some of the incident light is, however, scattered by R_1 and fills all of the aperture of the lens. This beam picks up the errors of the mirror and is the measured light beam.

Let us consider a ray incident at a point A on the scatter plate R_1 . The directly transmitted ray (solid line in Fig. 3.1) follows the path $AS'A'$ and encounters a scattering center at A' that is identical to the one at A . This ray is scattered at A' and gives rise to a cone of rays, one of which will follow the same path as the ray scattered on the first plate but not on the second. The rays scattered at A (dotted lines) that fill the aperture of the lens under test, arrive at the image A' and pass through R_2 without scattering. We thus have two mutually coherent beams emerging from R_2 . One beam is directly transmitted by R_1 and scattered by R_2 , and the second is scattered by R_1

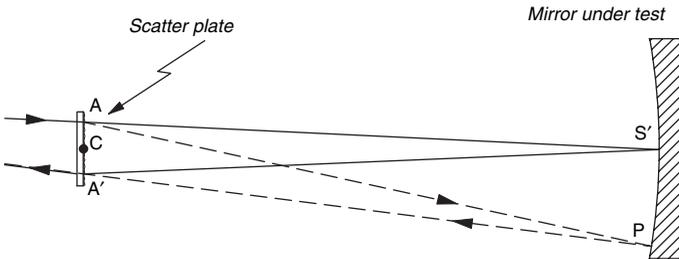


FIGURE 3.2. Burch interferometer testing a concave spherical mirror with a double pass through a scatter plate.

TABLE 3.1. Beams in Burch's scattering interferometer.

After first scatter plate	After second scatter plate	Exiting light beam
Unscattered	Unscattered	Central bright spot
Unscattered	Scattered	Reference beam
Scattered	Unscattered	Beam under test
Scattered	Scattered	Too dim to observe

and transmitted by R_2 . An observer looking at the mirror surface through R_2 will see an interferogram between these two beams. Many rays arrive at the first scatter plate, not only one, so the observed interferogram is the superposition of many identical fringe patterns. If the mirror is free of any error in the region of S' , the interferogram will provide explicit information about the mirror aberrations, as in any separate-beam interferometer.

The light that is directly transmitted by both R_1 and R_2 gives rise to a bright spot located at S' and is quite troublesome for visual observations. The light that is scattered by R_1 and again by R_2 gives a weak background and slightly diminishes the contrast of fringes. The dimensions of the source S should be such that its image S' remains localized within a fringe. If the fringes are quite broad in a certain region of the mirror (this is equivalent to saying that the mirror is almost free of aberrations in this region), the source image S' should be made to lie there. Summarizing, the observer receives the four beams in Table 3.1.

To obtain a permanent record of the interference pattern, a camera lens, placed after the second scatter plate R_2 , is used to form an image of the mirror surface on a photographic film. Each point of the film receives light from a conjugate point on the mirror surface. The interference effect (the light intensity) on the film will, thus, give information about the mirror aberration at the conjugate point.

Figure 3.2 is a schematic simplified diagram of Burch's interferometer as applied to the testing of a concave mirror M . The first scatter plate is assumed to have rotational symmetry, so that the scattering point A is identical to the scattering point A' . This scatter plate has to be at the center of curvature of the mirror, so that the magnification is one. As shown by Su et al. (1984), the optical path difference (OPD) for the two interfering rays is

$$OPD = (AP + PA') - 2AS' \tag{3.1}$$

If the surface under test is spherical, it is possible to see that this optical path difference is zero only for rays passing very close to the center of the scatter plate. However, assuming that the scatter plate size is small enough, if the surface is not spherical and its shape with respect to the sphere is W , the optical path difference is $2W$. Su et al. (1984) showed that if the scatter plate is too large, the contrast decreases, so that its maximum size should be

$$\Delta OPD = \frac{D^2 s^2 r^3}{4} < \frac{\lambda}{2} \tag{3.2}$$

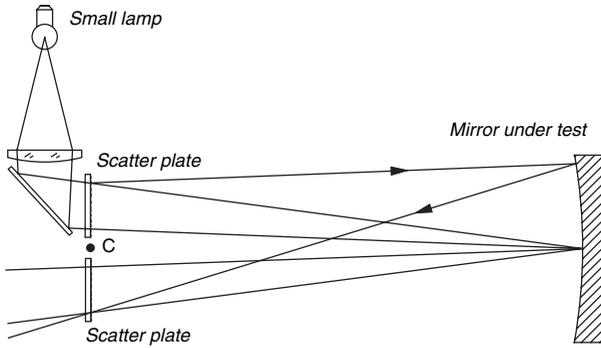


FIGURE 3.3. Burch interferometer testing a concave spherical mirror with two identical scatter plates.

where D is the mirror diameter, r is its radius of curvature, and s is the semidiameter of the scatter plate.

Two practical Burch interferometers for testing concave surfaces with a ratio of the radius of curvature to the diameter larger than about six, in order to be able to test them off-axis, are shown in Figures 3.3 and 3.4. The light source is a small tungsten lamp. If the mirror has a hole at the center as in some telescope mirrors, the image S' has to be off-center. In the interferometer in Figure 3.3, the two scatter plates have to be identical, but one has to be rotated to 180° with respect to the other. The scatter plate for the double pass interferometer in Figure 3.4 is made by fine grinding and then by half polishing the front face of a cube beam splitter. Symmetrically placed with respect to the center of the curvature is a small flat mirror. In order to prevent an

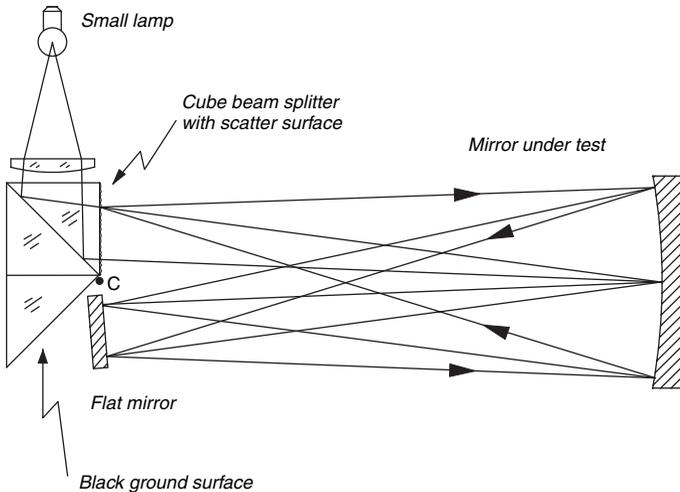


FIGURE 3.4. Burch interferometer testing a concave spherical mirror with a double pass through the scatter plate and symmetrically located with respect to the scatter plate.

unwanted reflection from going to the eye of the observer, a small triangular prism with a ground and black-painted surface is cemented to the cube.

A slight displacement of one of the scatter plates in its own plane, with respect to the image of the other, gives rise to a set of parallel straight-line fringes (tilt), and the mirror defects appear as distortions in the straightness of the fringes. This displacement is produced if the whole interferometer with the two scatter plates is moved perpendicularly to the optical axis. The tilt fringe disappears only when the two scatter plates are symmetrically placed with respect to the optical axis. A slight displacement in the axial direction, toward or away from the surface under test produces the effect of a defocusing term, introducing circular fringes. Since the optical paths for the two interfering light beams are identical, it is impossible to introduce a piston term. In this system, like in any symmetrical system, testing the coma aberration at the center of curvature is canceled out. Only the astigmatism remains. In this case, the fringes will have an elliptical shape. Scott (1969) used this interferometer to test a 91.5-cm, $f/4$ paraboloid, and a 35-cm Gregorian secondary. The details of testing and modifications made in Burch's original design are discussed in the article cited.

Burch's interferometer is quite sensitive to vibrations taking the form of tilts about an axis normal to the line of sight or translations across the line of sight. To make the system insensitive to these vibrations, Shoemaker and Murty (1966) modified the setup by replacing the second scatter plate with a plane mirror, and thus reimaging the first scatter plate point-by-point back on itself. This setup gives double sensitivity for even-order aberrations but cannot detect odd aberrations. An obvious great advantage is that only one scatter plate has to be made. Vibrations in the form of fast variations in the distance between the mirror and the scatter plates would produce vibrations in the focusing term, but they do not represent an important problem. An important problem, however, appears with this arrangement. Since the light is reflected twice on the mirror under test, the surface has to be highly reflective. Uncoated surfaces produce quite dim interferograms if a bright light source is not used.

Since the two interfering beams in the interferometer have the same optical paths, the interferometer is compensated for white light. Thus, the light source can be a small white tungsten lamp. If a high light intensity is desired, mainly when measuring uncoated surfaces in a double pass configuration, a laser is more intense, but spurious fringes or speckle may appear.

Two identical scatter plates R_1 and R_2 or one with rotational symmetry can be made by several possible methods, for example,

- (a) by photographing a speckle pattern. A symmetrical scatter plate can be made with a double exposure of the same speckle pattern, with a 180° rotation of the plate before taking the second exposure, as described by Su et al. (1984a), Smartt and Steel (1985), and North-Morris et al. (2002);
- (b) by taking two replicas from a lightly ground surface (Houston, 1970);
- (c) with microphotography of a rotationally symmetric pattern (Murty, 1963).

3.2.1. Fresnel Zone Plate Interferometer

A similar interferometer using Fresnel zone plates instead of scatter plates has been proposed by Murty (1963) and later by Smartt (1974), Lohmann (1985), Stevens (1988), and Huang et al. (1989). Instead of diffracting the light in a random manner in all directions, like scatter plate, a Fresnel zone plate produces many convergent as well as divergent spherical wavefronts. The advantage is that the Fresnel zone plate has rotational symmetry and thus the configurations in Figures both 3.2 and 3.3 are possible. The Fresnel zone plate can be made by photography of a large drawing and then reducing it to the desired size, or by photography of the interference between a divergent and a plane wavefront (Smartt, 1974).

3.2.2. Burch and Fresnel Zone Plate Interferometers for Aspheric Surfaces

We have seen that the optical path difference for spherical surfaces is not zero if the scattering point in the scatter plate is far from the optical axis. Then, when the scatter plate is large, different radial positions of the scattering point will produce different fringe patterns, reducing the contrast of the observed interferogram. If we could associate different radial positions of the scatter point to different radial positions on the surface under test, we could use with optical path difference variation to our advantage to design a null test for aspherical surfaces. This has been described by Su et al. (1986) and by Huang et al. (1989). They used a diaphragm with a small aperture, placed after the observing scatter plate, as illustrated in Figure 3.5. It can be proved that the optical path difference expressed by Eq. (3.1) remains valid in this case. The presence of the diaphragm associates different scattering points A to different points P on the surface under test. By proper choice of the diaphragm position L, the desired mirror asphericity can be compensated to produce a null test.

3.2.3. Burch and Fresnel Zone Plate Interferometers for Phase Shifting

Patorski and Salbut (2004) have described a scattering interferometer where the phase difference between the wavefront under test and the reference wavefront can be changed as desired. The light source for this instrument, illustrated in Figure 3.6 is a laser. If this laser is unpolarized, a polarizer has to be inserted before the first scattering plate, otherwise, if it is linearly polarized, it is not necessary. Then, a quarter wave phase plate with its slow or fast axis at 45° with respect to the plane of polarization is placed in the light beam to produce an illuminating circularly

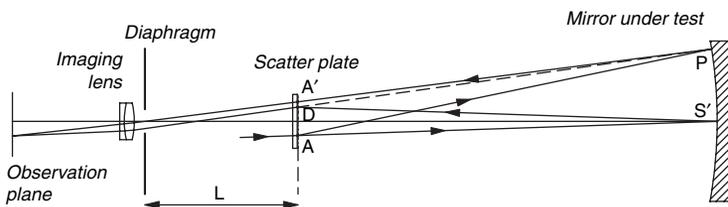


FIGURE 3.5. Burch and Fresnel zone plate interferometer to test an aspheric surface.

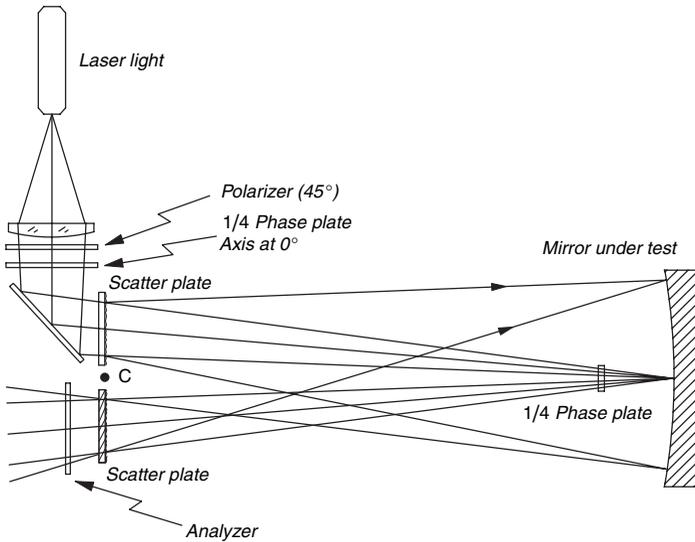


FIGURE 3.6. Burch and Fresnel zone plate interferometer for phase shifting using polarized light.

polarized light beam. Close to the mirror under test, in the path of the reference beam, a small quarter wave plate with an orientation is inserted. This phase plate is traversed twice by the reference beam, so that the sense of the circularly polarized beam is reversed. As a result of this arrangement, the beam under test and the reference beam are both circularly polarized, but with opposite senses. The light beams exiting the analyzer just before the observer are linearly polarized, with a phase difference given by the orientation of this analyzer.

Another phase shifting scattering interferometer (North-Morris and Wyant 2002) is shown in Figure 3.7. The main element in this instrument is a specially made scatter plate. It is made with an etched calcite plate with its optical axis parallel to the

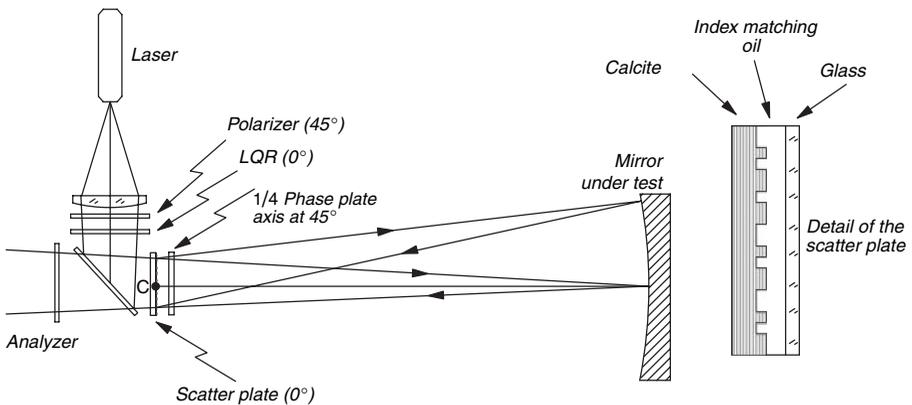


FIGURE 3.7. Burch interferometer for phase shifting using polarized light.

two faces of the plate. A speckle pattern is engraved on one of the two faces with a procedure described in the cited publication. Then, a thin glass cover is placed on top on the engraved calcite plate, with an oil layer between them. The oil has a refractive index matched to the ordinary refractive index of the calcite. In a uniaxial crystal as calcite, the extraordinary ray is polarized in a plane containing the optical axis of the crystal. The result is a scatter plate that scatters the light with its polarizing plane along the optical axis of the calcite but does not produce any scattering for the light linearly polarized in the perpendicular direction to the optical axis of the calcite. Another important optical element is a liquid crystal phase plate (LQR) whose phase delay difference between two perpendicular axes is variable and can be set at any value as desired.

As in the previously described interferometer, the light source from the laser is either linearly polarized or polarized with a polarizer in front of it at an angle of 45° . Then the phase difference between the horizontal and vertical components of the linear polarization is changed to any desired amount by means of the liquid crystal plate (changing it to elliptically polarized light). Next, the scatter plate scatters the polarization components that are along the optical axis of the calcite but not the polarization components that are perpendicular to this axis. The result is that the beam to be measured and the reference beam are polarized in orthogonal planes and with a phase difference given by the LQR. Both light beams pass twice through a quarter wave phase plate oriented at 45° . After this, the two interfering beams will be circularly polarized but in opposite senses. At the end, there is an analyzer just before the observer, which produces two linearly polarized beams with a phase difference given by the orientation of this analyzer.

3.3. BIREFRINGENT BEAM SPLITTERS

An important class of interferometers uses birefringent crystal elements as beam splitters. These interferometers are known as polarization interferometers (Françon and Mallick, 1971). We discuss in this section, three principal types of these beam splitters.

3.3.1. Savart Polariscopes

A Savart polariscopes consists of two identical uniaxial crystal plates with the optic axis cut at 45° to the plate normal (Fig. 3.8). The principal sections (plane containing the optic axis and the plate normal) of the two plates are crossed with each other. The optic axis of the first plate lies in the plane of the page and that of the second plate makes an angle of 45° with it; the dotted double arrow in the figure represents the projection of the optic axis on this plane. An incident ray is split by the first plate into two rays, the ordinary ray O and the extraordinary ray E. Since the second plate is turned through 90° with respect to the first one, the ordinary ray in the first plate becomes extraordinary in the second and vice versa. The ray OE does not lie in the plane of the page, though it emerges parallel to its sister ray EO, the dotted line

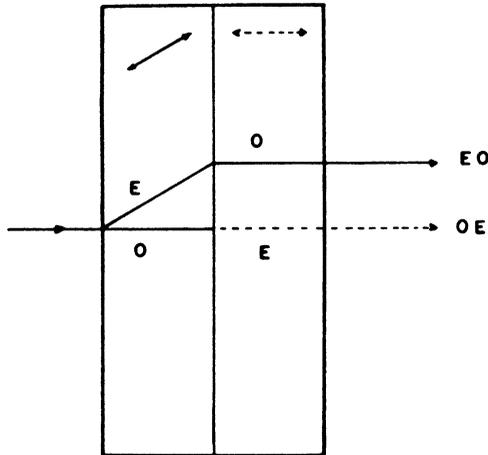


FIGURE 3.8. Beam splitting produced by a Savart polariscope. The figure is drawn for a polariscope made of a positive crystal (e.g., quartz).

represents the projection of the ray path on this plane. The lateral displacements between the two rays, each produced by one of the two component plates, are equal and are in perpendicular directions. The total displacement between the emerging EO and OE rays produced by a Savart polariscope of thickness $2t$ is given by

$$d = \sqrt{2} \frac{n_e^2 - n_o^2}{n_e^2 + n_o^2} t \tag{3.3}$$

where n_o and n_e are the ordinary and extraordinary indices of refraction, respectively. A 1-cm-thick polariscope will produce a lateral displacement of $80 \mu\text{m}$ if it is made of quartz and a 1.5-mm displacement if made of calcite. In Figure 3.8 if the incident ray is inclined to the plate normal, the two emerging rays are still parallel to the original ray, and their relative displacement remains practically unaltered.

The parallel emerging rays interfere in the far field (or in the back focal plane of a positive lens), and the interference pattern is similar to that produced in Young’s experiment with the two mutually coherent sources separated by a distance equal to d . For small angles of incidence, the fringes are equidistant straight lines normal to the direction of displacement. The angular spacing of these fringes is as follows:

$$\text{Angular spacing} = \frac{\lambda}{d}. \tag{3.4}$$

The zero-order fringe corresponds to normal incidence and lies in the center of the field of view. With a Savart polariscope of 1 cm thickness and a lens of 10 cm focal length, the fringe spacing in yellow light is 2 mm for quartz and 0.1 mm for calcite.

The OE and EO rays emerging from the Savart polariscope vibrate in mutually orthogonal directions. To make them interfere, the vibration directions are set

parallel to each other by means of a linear polarizer, the transmission axis of which is oriented at 45° to the orthogonal vibrations. This polarizer is, however, not sufficient for interference to take place. We know that a natural (unpolarized) ray of light is equivalent to two mutually incoherent components of equal amplitude vibrating in perpendicular directions. Thus, the ordinary and the extraordinary rays produced by a crystal have no permanent phase difference between them. To make these rays mutually coherent, a polarizer is placed across the incident beam so that only a single component of the natural light is transmitted onto the crystal. The transmission axis of this polarizer is at 45° to the principal axes of the crystal.

3.3.2. Wollaston Prism

A Wollaston prism (Fig. 3.9) consists of two similar wedges cemented together in such a way that the combination forms a plane parallel plate. The optic axes in the two component wedges are parallel to the external faces and are mutually perpendicular. A Wollaston prism splits an incident ray into two rays traveling in different directions; the lateral displacement between the rays is thus different at different distances from the Wollaston. The angular splitting α is given by the relation

$$\alpha = 2(n_e - n_o)\tan\theta \quad (3.5)$$

where θ is the wedge angle. For most practical purposes, α can be considered to be independent of the angle of incidence. For an angle $\theta = 5^\circ$, the angular splitting is 6 min of arc for a Wollaston prism made of quartz and 2° for one made of calcite.

The path difference between the OE and the EO rays (Fig. 3.10) emerging at a distance x from the axis $y - y'$ of the Wollaston prism is given by

$$\Delta = 2(n_e - n_o)x \tan\theta = \alpha x \quad (3.6)$$

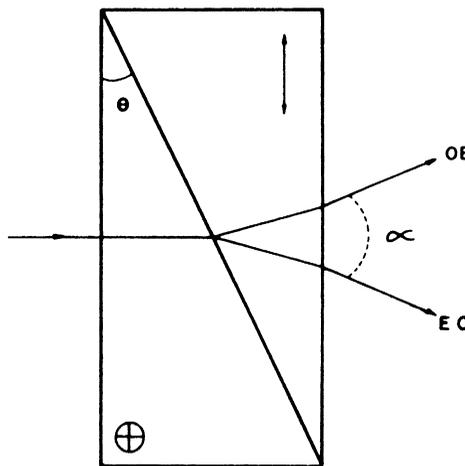


FIGURE 3.9. Beam splitting by a Wollaston prism made of a positive crystal.

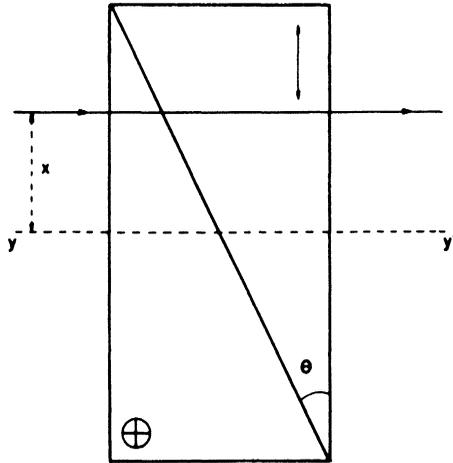


FIGURE 3.10. Path difference produced by a Wollaston prism between the two split-up rays is linearly related to x .

The path difference is zero along the axis, where the thicknesses of the two component wedges are equal, and increases linearly with x . When a Wollaston prism is placed between two suitably oriented polarizers, one observes a system of straight-line fringes parallel to the edges of the component wedges and localized in the interior of the prism. (The fringes are perpendicular to the plane of Figure 3.10). The path difference along the axis being zero, fringes are visible in white light. The fringe spacing is equal to

$$x_0 = \frac{\lambda}{2(n_e - n_o) \tan \theta} \tag{3.7}$$

where $\theta = 5^\circ$, $\lambda = 0.55 \mu\text{m}$, and $(n_e - n_o) = 9 \times 10^{-3}$ (quartz), there are approximately three fringes per millimeter. When the angle θ is very small (a few minutes of arc), the fringes are wide apart and the Wollaston prism can be used as a compensator. In this form, the Wollaston prism is known as a Babinet compensator.

Relation (3.6) for the path difference between OE and EO rays is true for normal incidence (the angular splitting, being small, is neglected for the calculation of Δ). For nonnormal incidence, a term proportional to the square of the angle of incidence is added to the right-hand side of Eq. (3.6). However, this term is negligible, for example, for a quartz prism of 10 mm thickness and for a case in which the angles of incidence remain less than 10° . Some modified Wollaston prisms have been devised that can accept much larger angles of incidence.

3.3.3. Double-Focus Systems

A lens made of a birefringent crystal acts as a beam splitter. A parallel beam of light incident on such a lens will be split into an ordinary beam and an extraordinary beam, which come to focus at two different points (Fig. 3.11). The O and E images are

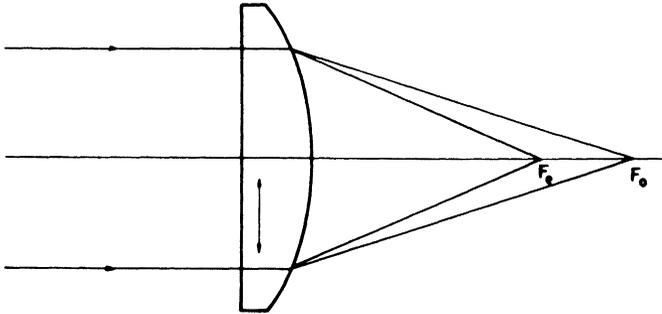


FIGURE 3.11. A birefringent lens splits up an incident beam into an ordinary and an extraordinary beam, which are brought to focus at two different points along the lens axis. The figure is drawn for a lens made of a positive crystal.

displaced along the axis of the beam, in contrast to the case of a Savart polariscope or of a Wollaston prism, where the displacement is normal to the direction of the incident beam. Various types of compound lenses suitable for specific applications have been designed.

3.4. LATERAL SHEARING INTERFEROMETERS

3.4.1. Use of a Savart Polariscope

Lateral shearing interferometers using birefringent beam splitters have been widely used to study the aberrations of an optical system. We describe here an arrangement by Françon and Jordery (1953) in which a Savart polariscope is used to produce a lateral shear of the aberrated wavefront (Fig. 3.12). The lens L (or the mirror) under

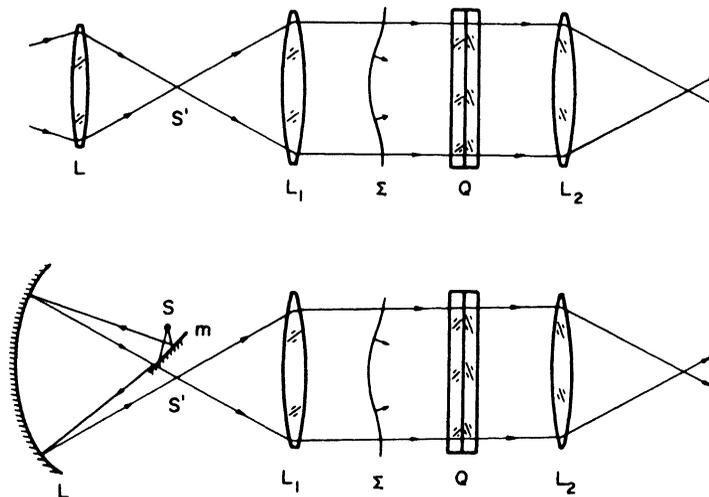


FIGURE 3.12. Interference arrangement employing a Savart polariscope Q for testing the optical system L .

test forms an image S' of a small source S . The distance of L from the source is fixed by the conditions under which the lens is to be tested. The lens L_1 collimates the light coming from S' so that the Savart polariscope Q is traversed by a parallel beam of light. Two linear polarizers (not shown in the figure) are placed before and after the Savart Q . The combination of lenses L_1 and L_2 constitutes a low-power microscope that is focused on the test lens L . If the lens L is perfect, the wavefront Σ is plane, then the ordinary and the extraordinary wavefronts produced by the Savart will have a uniform path difference between them. The eye placed in the focal plane of L_2 will observe a uniform color (or a uniform intensity in the case of monochromatic light) in the entire field of view. In the presence of aberrations, Σ will be deformed and the field of view will appear nonuniform. If the aberrations are large, a system of fringes will be observed. The nature and the magnitude of aberrations can be determined from the observed variations of color (or of intensity).

The far-field fringes of the Savart polariscope Q are located virtually in the plane of source image S' . The source size should be such that S' occupies a small fraction (say one fifth) of a fringe width.

The background color (or intensity) can be chosen by inclining the Savart Q about an axis parallel to the fringes. When the Savart is normal to the optical axis, that is, normal to the incident light, the zero-order fringe coincides with source image S' and the background will be dark (crossed polarizers). By inclining the Savart, S' can be made to coincide with a fringe of any desired color, which will then appear in the background. Instead of producing a uniform ground color in the field of view, we may produce a regular system of rectilinear fringes that are deformed in the region where the wavefront departs from the ideal form. Such fringes can be produced in a plane conjugate to the test lens by placing an additional Savart to the right of L_2 .

To illustrate the principle of the method, we study the aspect of the field of view in the presence of primary spherical aberration. The ground is chosen to be of uniform intensity. The distance parallel to the optical axis between the aberrated wavefront Σ and the ideal wavefront (corresponding to the Gaussian image point) at a height h from the axis is given by

$$z = ah^4 \quad (3.8)$$

where a is a constant depending on the magnitude of the aberration. To determine the aspect of the field of view, we have to calculate the path difference between the two sheared wavefronts Σ_1 and Σ_2 produced by the Savart polariscope. Figure 3.13 represents Σ_1 and Σ_2 as projected on a plane perpendicular to the optical axis of the system (this is the plane of the ideal wavefront); O_1 and O_2 are the centers of Σ_1 and Σ_2 , respectively. The coordinate system is chosen such that the x -axis passes through O_1 and O_2 and the y -axis is the right bisector of $O_1 - O_2$. Now consider a point $m(x, y)$ lying on the ideal plane wavefront; its distance from the aberrated wavefront Σ_1 is given by

$$z_1 = ar_1^4. \quad (3.9)$$

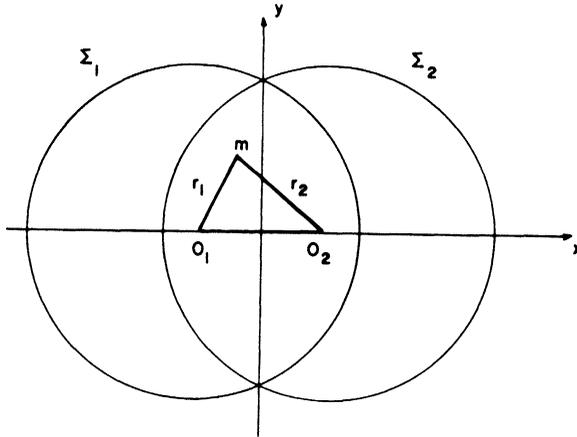


FIGURE 3.13. Representing the projection of the two sheared wavefronts on a plane normal to the optical axis of the system.

Similarly, the distance of point m from Σ_2 is given by

$$z_2 = ar_2^4. \quad (3.10)$$

The separation between Σ_1 and Σ_2 is, therefore,

$$z_1 - z_2 = a(r_1^4 - r_2^4) = 4a \times d \left(x^2 + y^2 + \frac{d^2}{4} \right) \quad (3.11)$$

where d is the shear between Σ_1 and Σ_2 . The lines of equal path difference, $z_1 - z_2$, are represented in Figure 3.14. The form of fringes for other aberrations can be

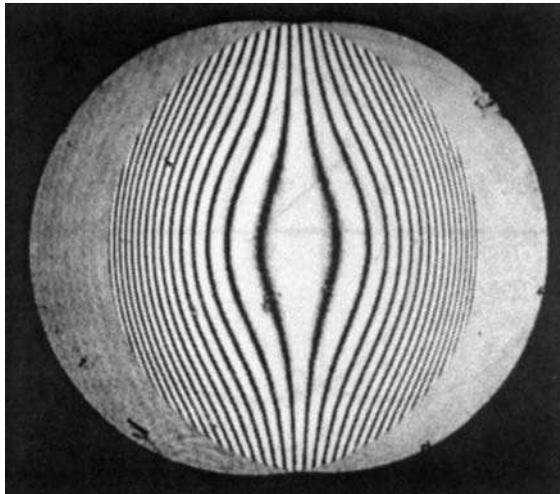


FIGURE 3.14. Lateral shearing interferogram of a wavefront distorted by spherical aberration of third order.

determined similarly. Evidently, when the aberrations are small, no fringes will be seen; there will simply be small variations of intensity in the field of view.

A complete analysis of the lateral shearing interferogram can be carried out by a mathematical operation described by Saunders (1961, 1962) (see Chapter 4). The method yields values of the deviations of the wavefront under test from a close fitting sphere. The reference sphere may be chosen statistically so that the results are the deviations from the best fitting surface.

3.4.2. Use of a Wollaston Prism

In the arrangement represented in Figure 3.12, it is possible to use a Wollaston prism instead of the Savart polariscope. This prism is placed at the source image S' . The background intensity can be changed by translating the Wollaston prism laterally perpendicular to the optical axis. A system of rectilinear fringes can be produced in the background by shifting the Wollaston prism along the optical axis.

The size of the source in the setup of Figure 3.12 is quite limited. It can be increased considerably, however, if the setup is modified so that the light passes twice through the Wollaston prism. Figure 3.15 illustrates such an arrangement. An image of the source S is formed on the Wollaston prism at the point S' , which is near the center of curvature of the mirror M under test. A lens L forms an image of M on the observation screen M' . As usual, two polarizers are needed to complete the system; one may be placed between m and W and the second between W and L . If observations are to be made between parallel polarizers, a single polarizer placed between W and L and covering all of the aperture of W , will suffice. If S' and S'' are symmetrically situated with respect to the central fringe of the Wollaston prism, the path difference between the interfering beams is zero and the background appears uniformly dark/bright with crossed/parallel polarizers. The ground intensity can be varied by displacing W in a direction perpendicular to its fringes. A system of straight fringes will appear on the screen if W is displaced along the axis of the interferometer so that it is no longer located at the center of curvature of M .

Philbert (1958) and Philbert and Garyson (1961) employed this interferometer to control the homogeneity of optical glass (the glass plate is placed close to M and to

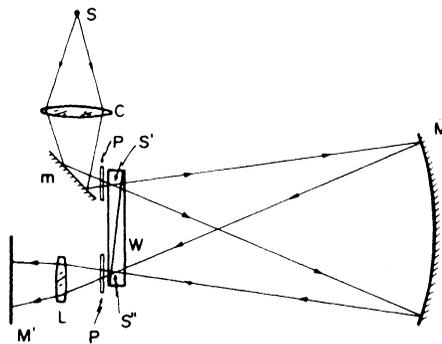


FIGURE 3.15. A double-pass compensated interferometer for testing the mirror M .

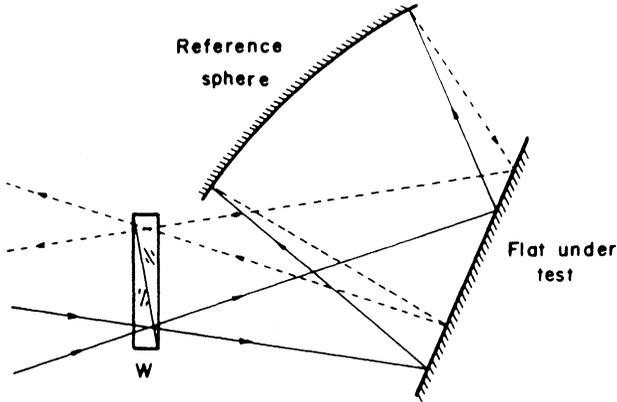


FIGURE 3.16. Setup for testing a flat surface.

test spherical, paraboloidal, and plane mirrors during the process of figuring. To test a paraboloidal surface, the Wollaston prism is placed at the focus and an auxiliary plane mirror is used to send back the parallel beam of light emerging from the paraboloid. A plane surface is tested by the arrangement represented in Figure 3.16. During the final stages of figuring, the deviations from the perfect surface are quite small and, therefore, the interferogram shows only slight variations in intensity. Under these conditions the aspect of the field of view is similar to that observed in Foucault's knife-edge test.

To make the system insensitive to vibrations, Dyson (1963) used a small plane mirror near the Wollaston prism in order to form an image of it back on itself. Then, instead of a large prism, a small one is used since half of it is replaced by the small mirror.

3.5. DOUBLE-FOCUS INTERFEROMETER

Dyson (1957a, 1957b, 1970) devised an interferometer for the testing of optical components in which he employed a birefringent lens as a beam splitter (Fig. 3.17).

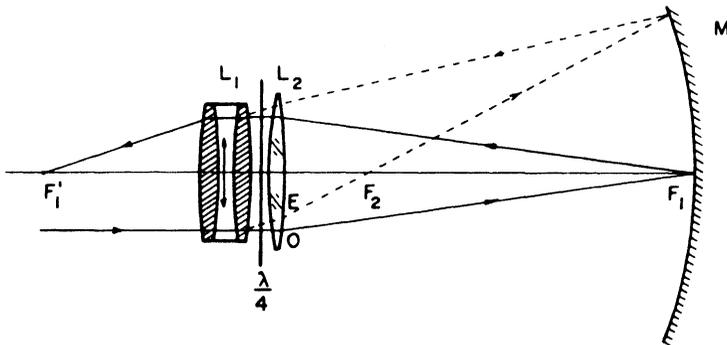


FIGURE 3.17. Dyson's double-focus interferometer.

The birefringent, double-focus lens L_1 is a symmetrical triplet, consisting of a central bio-concave calcite lens and two biconvex glass lenses. The optic axis of calcite lies in the plane of the lens. The triplet is so designed as to have zero power for the ordinary ray and a focal length of a few centimeters for the extraordinary ray. As in the case of Burch's interferometer (Section 3.2), this arrangement gives explicit information about wavefront deformations since a part of the incident light is focused on a small region in the aperture of the system under test and acts as the reference beam.

The system under test in Figure 3.17 is the concave mirror. The center of the triplet lens is located at the center of curvature of the mirror. A lens L_2 with its focus F_1 on the mirror surface is placed just to the right of the triplet. A quarter-wave plate with its principal axes at 45° to the optic axis of the calcite lens is also placed to the right of L_1 . A collimated beam of linearly polarized light is incident from the left. The lens L_1 splits it up into an ordinary beam and an extraordinary beam. The O beam, undeviated by L_1 , is brought to focus at F_1 by the lens L_2 . An image of the source is thus formed at F_1 . On its return journey, the O beam is collimated by the lens L_2 , and since its vibration direction has been rotated through 90° because of the double passage through the quarter-wave plate, it is refracted to F'_1 by the lens L_1 .

At its first passage the extraordinary beam is refracted by both the lenses L_1 and L_2 and converges to F_2 , the focus of the combination L_1L_2 . The beam then expands to fill the whole aperture of the mirror M . Because of the symmetry of the arrangement this beam, too, is brought to focus at F'_1 . A semireflecting surface is placed to the left of the triplet so that the source (or the system of observation) can be placed outside the axis of the interferometer.

An observer receiving the light at F'_1 will see (a) a uniform disk of light, determined according to size by the angular aperture of the lens L_1 (reference field) and (b) the illuminated aperture of the mirror M (test field). These two fields will interfere (there is evidently an analyzer that sets the O and E vibrations parallel to each other), and in the absence of aberrations, the resultant field will be of uniform intensity. If the triplet is slightly displaced laterally, so that its center no longer coincides with the center of curvature of the mirror, the field of view will be crossed with rectilinear fringes. When the triplet is displaced axially, circular fringes are observed. When the mirror has aberrations, these fringes are distorted. The aberrations can be deduced from these distortions in the same way as in any separate-path interferometer.

Dyson's interferometer is applicable to autostigmatic systems, that is, systems in which light diverging from a point in a particular plane is refocused to a point in the same plane to form an inverted image. Systems that are not autostigmatic can be converted to this form by the addition of one or more auxiliary components. To test a lens, for example, the scheme of Figure 3.18 is employed. In Figure 3.18(a), the lens is tested at infinite conjugates and in Figure 3.15b, it is tested at finite conjugates. The focus C coincides with the center of the triplet. It may be noted that the system under test is not operating exactly under its correct conditions as the test beam does not return along its original path. The arrangement gives the sum of the aberrations for two focal positions, one on each side of the desired position. The resultant error is often very small. Because of the aberrations of the triplet lens, optical systems of only moderate aperture $-f/5$, for example—can be tested with this interferometer.

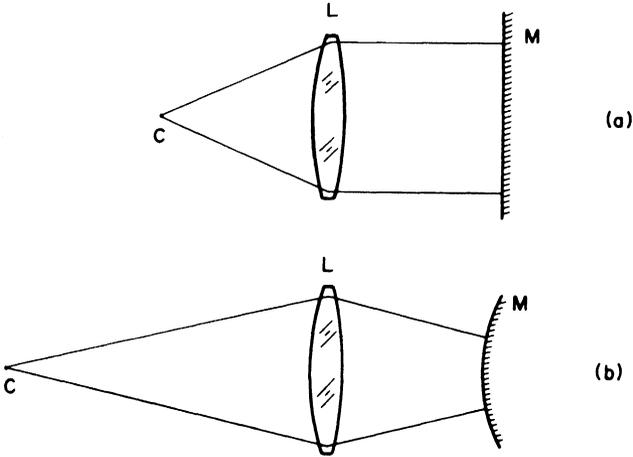


FIGURE 3.18. The lens L under test can be made autostigmatic by the addition of an auxiliary mirror M.

3.6. SAUNDERS'S PRISM INTERFEROMETER

Saunders (1967, 1970) described a lateral shearing interferometer in which the beam divider is made by cementing together the hypotenuse faces of two right-angle prisms, one of which is half silvered (Fig. 3.19). The faces B and B' are made highly reflecting. To obtain the zero-order fringe in the center of the field of view, the distances from the center of the beam-dividing surface to the two reflecting surfaces are made equal. If the two component prisms are identical, the two beams emerging from the face A' are mutually parallel. An angular shear between the beams can be introduced by rotating one prism relative to the other about an axis normal to the semireflecting surface. The direction of shear is approximately parallel to the vertex edges of the prisms. The shear can also be produced by making the angles α and α' of the two component prisms slightly different. This is usually the case when the prisms

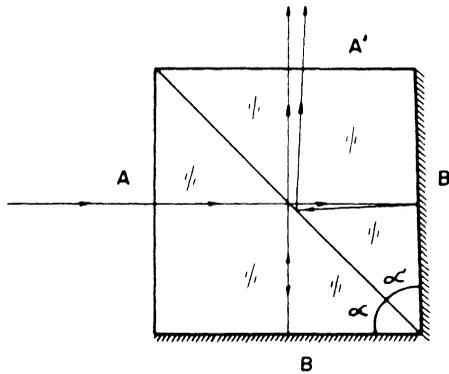


FIGURE 3.19. A beam splitter devised by Saunders.

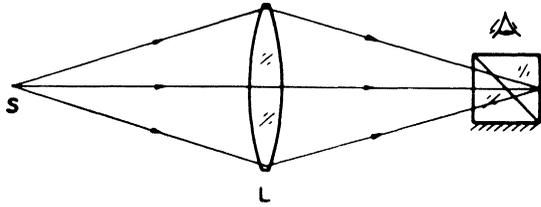


FIGURE 3.20. A lateral shearing interferometer using the beam splitter shown in Figure 3.19.

are not cut from a single large prism but are made separately. The angular shear is then equal to $2(\alpha - \alpha')$, and the direction of shear is perpendicular to the vertex edges.

Figure 3.20 shows an arrangement for testing a lens at finite conjugates. The prism is adjusted so that its back surface is approximately parallel to the image plane and is near it, with the principal ray of light passing near the center of the prism. This adjustment should produce visible fringes. The fringe width is very large when the source image lies on the back surface of the prism. The fringe width can be decreased by moving the prism along the principal ray away from the source image. By translating the prism laterally parallel both to the image plane and to the direction of shear, any chosen fringe can be made to pass through any chosen point of the interferogram. The adjustments of Saunders's prism are similar to those of a Wollaston prism. To obtain high contrast fringes, the source size in the shear direction is kept small. The recommended size of the cube is 10–15 mm.

Saunders (1957) also studied a wavefront-reversing interferometer that employed a modified Kösters double-image prism. Figure 3.21 is a sketch of the arrangement for testing a lens with one conjugate at infinity. The base of the dividing prism is spherical, and its center of curvature S_0 coincides with the image point at which the lens is to be tested. The observer's eye is located at S' and the image of the source is S . In this arrangement, the part of the wavefront lying below the dividing plane of the prism appears to be folded onto the upper half after the second passage through the prism. When the dividing plane cuts through the center of the lens, the even-order aberrations are eliminated. However, when the dividing plane is adjusted to form an angle with the axis of the lens, the even-order terms are retained. Saunders gave different variations of this arrangement for determining different aberrations.

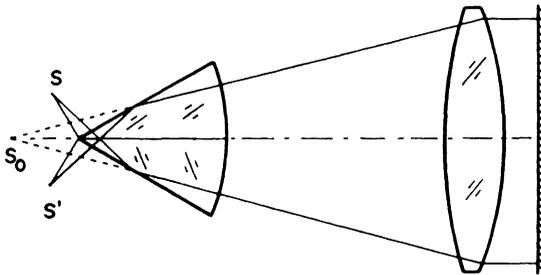


FIGURE 3.21. Saunders's wavefront-reversing interferometer.

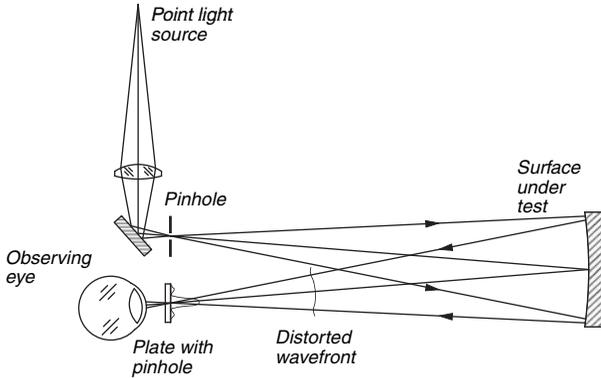


FIGURE 3.22. Point diffraction interferometer.

3.7. POINT DIFFRACTION INTERFEROMETER

Another interesting common path interferometer is the so-called point diffraction interferometer, first described by Linnik in 1933, rediscovered by Smartt and Strong (1972), and more fully developed by Smartt and Steel (1975). The principle of this interferometer is shown in Figure 3.22. The wave to be examined is brought to a focus, usually aberrated, of a point source. At the plane of that image, an absorbing film is placed. This film contains a small pinhole or opaque disk in order to diffract the light and produce a spherical reference wavefront. Figure 3.23(a) shows the amplitude magnitude at the image plane, where the diffracting plate is located and Figure 3.23(b) shows the total amplitude after passing the diffracting plate, which can be considered as the superposition of two images as in Figure 3.23(c).

To produce an interferogram with good contrast, the wave passing through the film and the diffracted spherical wave should have the same amplitude at the

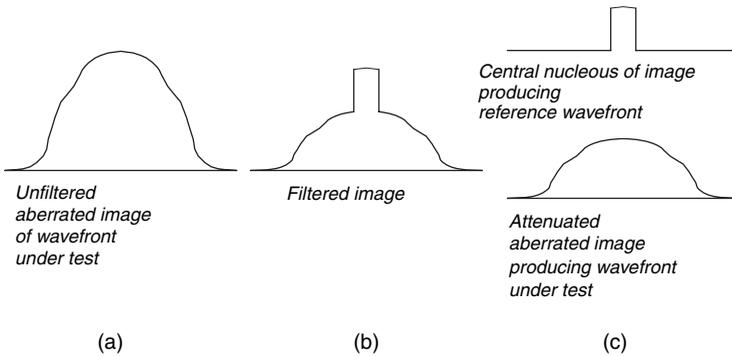


FIGURE 3.23. Modulus of the amplitudes in the image plane, where the diffraction plate is located. (a) Before the diffraction plate, (b) after the diffraction plate, and (c) the two separated components for the reference beam and the beam under test.

observing plane. This is controlled by means of the filter transmittance and the pinhole or disk size. Also, the amplitude of the spherical wave depends on how much of the light in the image falls on the pinhole or disk, and this, in turn, depends on the aberrations of the wave and on the pinhole or disk position. Smartt and Steel (1975) advised using filter transmittances between 0.005 and 0.05, with a most common value of 0.01. The optimum size for the pinhole or disk is about the size of the Airy disk that the original wave would produce if it had no aberrations. To match the amplitudes of the two beams, Wu et al. (1984) used a clear pinhole in a polarizing sheet of vectograph film; rotation of a polarizer behind this sheet changes the amplitude of the beam transmitted by the film but not that of the diffracted beam. The usual tilt and focus shift of the reference wavefront can be produced by displacing the diffracting point laterally and longitudinally, respectively.

A phase shifting point diffraction interferometer has been reported by Miller et al. (2004). The most important component of this interferometer is a polarization point diffraction plate, as illustrated in Figure 3.24. The central zone and the annular part in this plate are both polarizers in orthogonal directions. They are made with a wire grid structure whose construction details are in the cited publication. After the diffraction plate the reference wavefront and the wavefront under measurement have linear polarizations in orthogonal directions. A quarter wave phase plate at 45° with the two orthogonal polarization planes will make the two interfering beams circularly polarized in orthogonal directions. Then, as usual, a rotating analyzer will make the phase difference of the desired value. Another phase shifting point diffraction interferometer has been recently reported by Neal and Wyant (2006) using a birefringent pinhole plate.

The point diffraction interferometer has been used with success to test astronomical telescopes (Speer et al., 1979) and toric surfaces (Marioge et al., 1984). Smartt and Steel (1985) have developed a white-light interference microscope based on point diffraction interference principle.

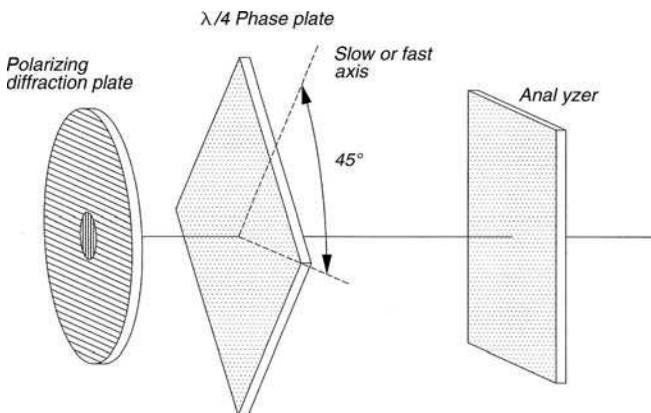


FIGURE 3.24. Main elements for polarization diffraction interferometer for phase shifting interferometer.

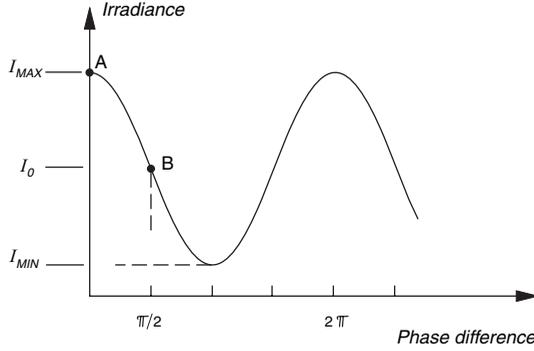


FIGURE 3.25. Irradiance as a function of the phase difference in an interferometer.

3.8. ZERNIKE TESTS WITH COMMON-PATH INTERFEROMETERS

In any two-beam interferometer, the irradiance in the interference pattern is a function of the phase difference between the two beams, as shown in Figure 3.25. If the interferogram has many fringes, the irradiance goes through many maxima and minima of this function. However, if the wavefront is almost perfect and its deformations are smaller than half the wavelength of the light, the phase changes will not produce any variations in the irradiance since the slope at the point A in this figure is zero.

These small wavefront errors may be easily detected if a bias in the phase difference is introduced by any means, so that it has a value equal to $\pi/2$ when the wavefront is perfect (point B). Then, the interferometer sensitivity to small errors is very large. Figure 3.26(a) shows an interferogram with a piston term equal to zero and, Figure 3.26(b) is formed with the same wavefront but a piston term equal to $\pi/2$.

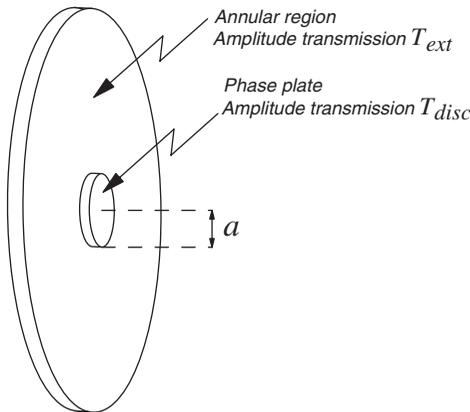


FIGURE 3.26. Schematics of a Zernike diffraction plate.

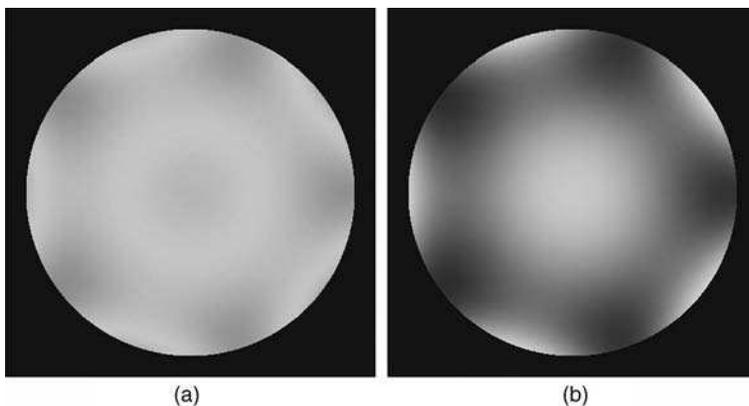


FIGURE 3.27. Two interferograms produced with the same wavefront, with a deformation much smaller than the wavelength of light. (a) Without a piston term and (b) with a piston term equal to $\lambda/4$.

To make a Zernike Point diffraction interferometer a diffracting plate as illustrated in Figure 3.27 is used. The annular region is coated with an amplitude transmission T_{ext} , while the central disc is made with an amplitude transmission T_{disc} , typically equal to one. The central disc is thicker, so that the phase optical path through the center is greater than the phase optical path through the annular region. Let us assume that the optical phase difference is equal to ϕ . Thus, the phase difference between the wavefront under test and the reference wavefront is γ , as given by the expression

$$T_{disc} \exp(i\phi) - T_{ext} = |T_1| \exp(i\gamma) \quad (3.10)$$

where

$$|T_1|^2 = T_{disc}^2 + T_{ext}^2 - 2T_{disc}T_{ext} \cos \phi \quad (3.11)$$

Thus, we may see that the phase difference between the reference wavefront and the wavefront under test is given by γ , as

$$\tan \gamma = \frac{T_{disc} \sin(\phi)}{T_{disc} \cos \phi - T_{ext}} \quad (3.12)$$

Since we need $\gamma = \pi/2$, we require that

$$\cos \phi = \frac{T_{ext}}{T_{disc}} \quad (3.13)$$

We see that ϕ and γ approach the same value if T_{ext} becomes quite small. The Zernike test in the point diffraction interferometer from a physical optics point of view is studied in Chapter 8.

REFERENCES

- Burch J. M., "Scatter Fringes of Equal Thickness," *Nature*, **171**, 889 (1953).
- Burch J. M., "Scatter-Fringe Interferometry," *J. Opt. Soc. Am.*, **52**, 600 (1962).
- Burch J. M., "Interferometry With Scattered Light," in: *Optical Instruments and Techniques*, J. Home Dickson, Ed., Oriol Press, England, 1969, p. 213.
- Dyson J., "Common-Path Interferometer for Testing Purposes," *J. Opt. Soc. Am.*, **47**, 386 (1957a).
- Dyson J., "Interferometers," in: *Concepts of Classical Optics*, John Strong, Ed., W. H. Freeman, San Francisco, 1957b, Appendix B, p. 377.
- Dyson J., "Very Stable Common-Path Interferometers and Applications," *J. Opt. Soc. Am.*, **53**, 690 (1963).
- Dyson J., *Interferometry as a Measuring Tool*, Machinery Publishing Co., Brighton, 1970.
- Françon M. and M. Jordery, "Application des Interférences par Double Réfraction à l'Etude des Aberrations," *Rev. Opt.*, **32**, 601 (1953).
- Hopkins H. H., "Interferometric Methods for the Study of Diffraction Images," *Opt. Acta*, **2**, 23 (1955).
- Houston J. B. Jr., "How to Make and Use a Scatterplate Interferometer," *Opt. Spectra*, **4**, 32 (1970).
- Huang J., N. Ohyama, and T. Honda, "A Null Test of Conic Surfaces in Zone Plate Interferometer," *Opt. Commun.*, **72**, 17 (1989).
- Juneji H., T. Honda, N. Ohyama, and J. Tsujiuchi, "Fringe Scanning Scatter Plate Interferometer," *Opt. Commun.*, **68**, 235 (1988).
- Linnik W., "Simple Interferometer to Test Optical Systems," *Comptes Rendus de l'Académie des Sciences de l'U.R.S.S.* 1, 208 (1933). Abstract in *Z. Instrumentenk.*, **54**, 463 (1934).
- Lohmann A. W., "An Interferometer with a Zone Plate Beam-Splitter," *Opt. Acta*, **32**, 1465 (1985).
- Marioge J.-P., B. Bonino, F. Bridou, P. Fournet et al., "La Fabrication et le Contrôle de Surfaces Toriques," *J. Opt. (Paris)*, **15**, 286 (1984).
- Miller J. E., N. J. Brock, J. B. Hayes, and J. C. Wyant, "Instantaneous Phase-Shift, Point-Diffraction Interferometer," *Proc. SPIE*, **5531**, 264–272 (2004).
- Murty M. V. R. K., "Common Path Interferometer Using Fresnel Zone Plates," *J. Opt. Soc. Am.*, **53**, 568 (1963).
- Neal R. M. and J. C. Wyant, "Polarization Phase-Shifting Point-Diffraction Interferometer," *Appl. Opt.*, **45**, 3463–3476 (2006).
- North-Morris M., J. Van Delden, and J. C. Wyant, "Phase-Shifting Birrefringent Scatterplate Interferometer," *Appl. Opt.*, **41**, 668–677 (2002).
- Patorski K. and L. Salbut, "Simple Polarization Phase-Stepping Scatterplate Interferometry," *Opt. Eng.*, **43**, 393–397 (2004).
- Philbert M., "Applications Métrologiques de la Strioscopie Interférentielles," *Rec. Opt.*, **37**, 598 (1958).
- Philbert M. and M. Garyson, "Réalisation et Contrôle par Strioscopie Interférentielle de Miroirs Plans, Sphériques et Paraboliques," in: *Optical Instruments and Techniques*, K. J. Habell, Ed., Chapman and Hall, London, 1961, p. 352.

- Rubin L. F., "Scatter Plate Interferometry," *Opt. Eng.*, **19**, 815 (1980).
- Saunders J. B., "The Kösters Double-Image Prism," in *Concepts of Classical Optics*, John Strong, Ed., W. H. Freeman, San Francisco, 1957, Appendix C, p. 393.
- Saunders J. B., "Measurement of Wavefronts Without a Reference Standard I: The Wavefront-Shearing Interferometer," *J. Res. Nat. Bur. Stand.*, **65B**, 239 (1961).
- Saunders J. B., "Measurement of Wavefronts Without a Reference Standard II: The Wavefront-Reversing Interferometer," *J. Res. Nat. Bur. Stand.*, **66B**, 29 (1962).
- Saunders J. B., "A Simple, Inexpensive Wavefront Shearing Interferometer," *Appl. Opt.*, **6**, 1581 (1967).
- Saunders J. B., "A Simple Interferometric Method for Workshop Testing of Optics," *Appl. Opt.*, **9**, 1623 (1970).
- Scott R. M., "Scatter Plate Interferometry," *Appl. Opt.*, **8**, 531 (1969).
- Shoemaker A. H. and M. V. R. K. Murty, "Some Further Aspects of Scatter-Fringe Interferometry," *Appl. Opt.*, **5**, 603 (1966).
- Smartt R. N. and J. Strong, "Point-Diffraction Interferometer" (abstract only), *J. Opt. Soc. Am.*, **62**, 737 (1972).
- Smartt R. N., "Zone Plate Interferometer," *Appl. Opt.*, **13**, 1093 (1974).
- Smartt R. N. and W. H. Steel, "Theory and Application of Point-Difference Interferometers," Proceedings of the ICO Conference on Optical Methods in Scientific and Industrial Measurements, Tokyo, 1974; *Jpn. J. Appl. Phys.*, **14**, Suppl. 1, 351 (1975).
- Smartt R. N. and William H. Steel, "Point Diffraction Interferometer," *Appl. Opt.*, **24**, 1402 (1985).
- Speer R. J., M. Crisp, D. Turner, S. Mrowka, and K. Tregidjo, "Grazing Incidence Interferometry: The Use of the Linnik Interferometer for Testing Image-Forming Reflection Systems," *Appl. Opt.*, **18**, 2003 (1979).
- Stevens R. F., "Zone Plate Interferometers," *J. Mod. Optics (Form. Opt. Acta)*, **35**, 75 (1988).
- Su D.-C., T. Honda, and J. Tsujiuchi, "A Simple Method of Producing Accurately Symmetrical Scatter Plates," *Opt. Commun.*, **49**, 161 (1984).
- Su D.-C., N. Ohyama, T. Honda, and J. Tsujiuchi, "A Null Test of Aspherical Surfaces in Scatter Plate Interferometer," *Opt. Commun.*, **58**, 139 (1986).
- Su D.-C., T. Honda, and J. Tsujiuchi, "Some Advantages of Using Scatter Plate Interferometer in Testing Aspheric Surfaces," *Proc. SPIE.*, **813**, 217 (1987).
- Wu S. T., Ch. L. Xu, and Zh. J. Wang, "New Application of Point Diffraction Interferometer—Polarization Fringe Scanning PDI," in: *ICO 13 Conference Digest, Optics in Modern Science and Technology*, H. Ohzu, Ed., Reidel, Dordrecht, 1984, p. 458.

4

Lateral Shear Interferometers

M. Strojnik, G. Paez, and M. Mantravadi

4.1. INTRODUCTION

Lateral shearing interferometry is an important field of interferometry and has been used extensively in diverse applications such as the testing of optical components and systems and the study of flow and diffusion phenomena in gases and liquids. Basically, the method of lateral shearing interferometry consists of duplicating wavefront under study, displacing it laterally by a small amount, and obtaining the interference pattern between the original and the displaced wavefronts.

Figure 4.1 schematically illustrates the principle of shearing interferometry for (a) an approximately planar wavefront and (b) spherical wavefront. When the wavefront is nearly planar, the lateral shear is obtained by displacing the wavefront in its own plane. If the wavefront is nearly spherical, the lateral shear is obtained by sliding the wavefront along itself by rotation about an axis passing through the center of curvature of the spherical wavefront.

There are many physical arrangements that produce lateral shear. The famous Italian optical scientist Ronchi is the first to have introduced laterally sheared wavefronts to test optical components in the first half of the 20th century. He employed diffraction at a set of suitably separated lines to produce zeroth- and first-order beams. Prior to the discovery of lasers in the 1960s, this became a popular technique in optical testing, still bearing the inventor name, *Ronchi test*.

In this chapter, we discuss arrangements that can be obtained by the use of beam dividers, which divide the amplitude of the incident wavefront but do not change the shape of the wavefront. This means that plane surfaces coated with semireflecting material are used as beam dividers. Several arrangements to obtain lateral shear will be described in this chapter mainly to show that with available components, one can easily fashion a workable lateral shearing interferometer in one's laboratory or optical workshop. Lateral shearing interferometry is basically a one-dimensional action. When it is performed in two orthogonal directions, it becomes twice a one-dimensional function. We will also discuss a more general case of a vectorial

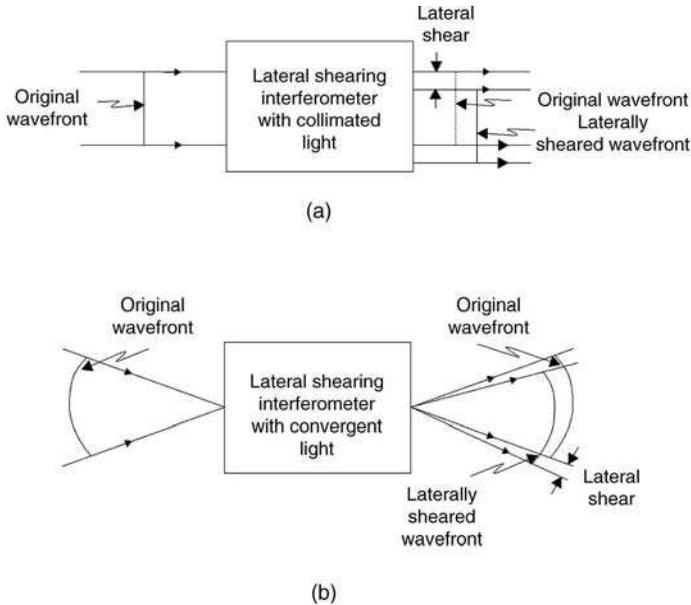


FIGURE 4.1. Schematic diagram illustrating lateral shearing interferometry in (a) collimated light and (b) convergent light.

shearing interferometry, where a two-dimensional action is obtained with a single shear.

Another important consideration in the design of lateral shearing interferometers is the nature of the light source. From the point of view of lateral shearing interferometry, the sources can be classified into two categories: (a) laser sources, such as the helium-neon gas laser giving a 632.8-nm light beam of very high spatial and temporal coherence, and (b) all other sources, such as gas discharge lamps, which are temporally coherent to some extent but not spatially coherent.

4.2. COHERENCE PROPERTIES OF THE LIGHT SOURCE

Figure 4.2 illustrates the arrangement of a lateral shearing interferometer in which a shear takes place for a nearly plane wavefront obtained from the collimating lens. Let the full beam width of the wavefront under test be denoted as d , the amount of lateral shear S , and the focal length of the collimating lens be f . The wavefront will be spatially coherent across its beam diameter when the size of the source is equal to the width of the central diffraction maximum (Airy disk) corresponding to the f -number of the particular collimating lens.

The f -number is the ratio of focal length of the optical system f , divided by the diameter of its aperture d . The diameter of the diffraction disk is $1.22 \lambda f/d$ for a circular aperture. Here, λ is the wavelength of a particular spectral line of the source that is to be used. Thus, the order of magnitude of the size of the pinhole to be

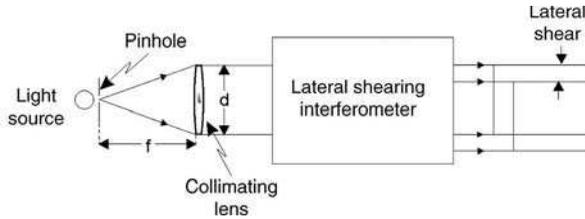


FIGURE 4.2. Schematic diagram indicating the various parameters for the consideration of the size of pinhole to be used in a lateral shearing interferometer.

used over the source to achieve spatial coherence is given by $(\lambda f/d)$. Fortunately in the lateral shearing interferometry, the spatial coherence should be sufficient so that the interference can be observed between parts of the wavefront separated by the distance S (which is less than d). Hence the source (pinhole) size can be $(\lambda f/d)(d/S) = (\lambda/S)$. Thus the pinhole size chosen is some multiple of the diffraction-limited pinhole size.

As an example, let us assume that we are using a mercury discharge lamp as the source of light and that the green line (546.1 nm) is isolated by means of a filter. If we are using a collimating lens of f -number = 5, then, assuming a shear ratio S/d of 0.1, the pinhole must be of the order of 25 μm . This is an extremely small pinhole, and generally very little intensity can be obtained in the fringe pattern. Hence, an intense source such as a high-pressure mercury arc has to be used. It has poor temporal coherence, even after a spectral line suitable for the purpose has been isolated by means of a filter. Use of such sources results in the requirement to compensate the two optical paths in an interferometer to be used as a lateral shearing instrument. This condition (and its implementation) is sometimes referred to as white-light compensation. When white light is used, a lateral shearing interferogram is obtained where the central fringe is achromatic (white) and the other fringes are colored.

Until the gas laser came into general use, all lateral shearing interferometers were designed with white-light compensation. Now it is possible to devise lateral shearing interferometers in which the light paths of the two interfering beams are of unequal length (uncompensated). A laser source having a high degree of spatial and temporal coherence is, however, necessary for this purpose. A helium–neon (He–Ne) laser emitting at 632.8 nm is often used as a source of light for many of these applications. A lateral shearing interferometer designed for white-light compensation can always be used with a laser light source. However, the inverse is not true. A lateral shearing interferometer designed for laser use, and hence probably having unequal optical paths, cannot produce a visible or recordable interference fringe pattern with sources of light of lesser coherence.

4.3. BRIEF THEORY OF LATERAL SHEARING INTERFEROMETRY

The wavefront error $W(x, y)$ is the difference between the actual wavefront and the desirable one or the one required to accomplish the design objectives. Wavefront is a

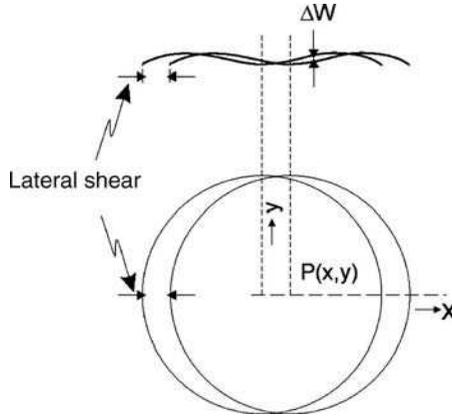


FIGURE 4.3. Schematic diagram (in plane and elevation) illustrating the original and the sheared wavefronts for a circular aperture. The fringes arising due to the lateral shear may only be seen in the area of overlap of the two wavefronts.

locus of all points with the same phase in the three-dimensional space. Therefore, its unit in the MKS system is meter. In interferometry and optical engineering, it is expressed in terms of λ (number of wavelengths of the illuminating source). When the illuminating source is a He-Ne laser emitting at 632.8 nm, the conversion between these two systems is $\lambda = 632.8 \text{ nm} = 632.8 \times 10^{-9} \text{ m}$.

Figure 4.3 shows the original and the laterally sheared wavefronts. For now, we assume that the wavefront is nearly planar so that the wavefront errors may be considered as small deviations from this plane. The wavefront error may be denoted as $W(x, y)$, where (x, y) are the coordinates of an arbitrary point $P(x, y)$. When this wavefront is sheared in the x direction by an amount S , the error at every point on the sheared wavefront is $W(x - S, y)$. The resulting wavefront difference $\Delta W(x, y)$ at $P(x, y)$ between the original and the sheared wavefronts is $W(x, y) - W(x - S, y)$. Thus, in later shearing interferometry, $\Delta W(x, y)$ is the quantity that is found.

When the displacement S is zero, there is no wavefront difference anywhere in the interferometer. Consequently, wavefront difference cannot be seen and error cannot be measured independently of its magnitude. Now, the path wavefront difference $\Delta W(x, y)$ may be expressed in terms of number of wavelengths according to the usual relationship

$$\Delta W(x, y) = n\lambda \tag{4.1}$$

where n is the order of the interference fringe and λ is the wavelength used. The left-hand side of Eq. (4.1) may be multiplied by 1, in a specific form of $(S/\partial x)$. When the displacement S is made increasingly smaller, and is theoretically approaching zero, the change in wavefront difference over change in displacement Δx becomes a partial derivative. Equation (4.1) may be written as

$$\frac{\partial W(x, y)}{\partial x} S = n\lambda \tag{4.2}$$

Thus the information obtained in the lateral shearing interferometer is ray aberration ($\partial W/\partial x$) in angular units. The accuracy of Eq. (4.2) increases as shear S approaches 0. It is believed that sensitivity decreases as shear S approaches 0. Therefore, the experimentalist strives to arrive at a suitable compromise for the optimal value of shear S if Eq. (4.2) is to be used.

Let us now consider some specific situations, employing the simpler form in the case of nonrotationally symmetric aberrations.

4.3.1. Interferograms of Spherical and Flat Wavefronts

Defocusing. The wavefront error for defocusing may be represented as

$$W(x, y) = D(x^2 + y^2). \quad (4.3)$$

The coefficient D represents the magnitude of the aberration, usually given as a number of wavelengths. A slight defocusing of the optical system, designed to produce a perfectly plane wave, will result in the emergence of a wavefront that is either slightly concave or convex spherical wavefront with a very long radius of curvature. Hence in this case,

$$\Delta W(x, y) = 2DxS = n\lambda. \quad (4.4)$$

Equation (4.4) represents a system of straight fringes that are equally spaced and perpendicular to the x direction (direction of shear). This situation is illustrated in Figure 4.4(a). The straight fringes appear in the common area of the overlapping wavefronts. If there is no defocusing ($D = 0$), there are no fringes. The area of wavefront overlap appears to be of uniform intensity, corresponding to the same optical path for both beams.

Tilt. When the wavefront is laterally sheared, normally we assume that the new wavefront is not tilted with respect to the original wavefront. In certain arrangements, however, it is possible to obtain a known amount of tilt between the two wavefronts. In such cases, it is a usual practice to obtain the tilt in the direction orthogonal to that of the lateral shear. The optical path difference associated with this tilt may be represented as a linear function of the y coordinate. Thus, in the case of only tilt, we find

$$\Delta W(x, y) = Ey = n\lambda. \quad (4.5)$$

Here E is the angle of tilt between the original and the sheared wavefronts. Their line of intersection is parallel to the x axis. If defocusing and tilt are simultaneously present, the optical path difference is given by

$$\Delta W(x, y) = 2DxS + Ey = n\lambda. \quad (4.6)$$

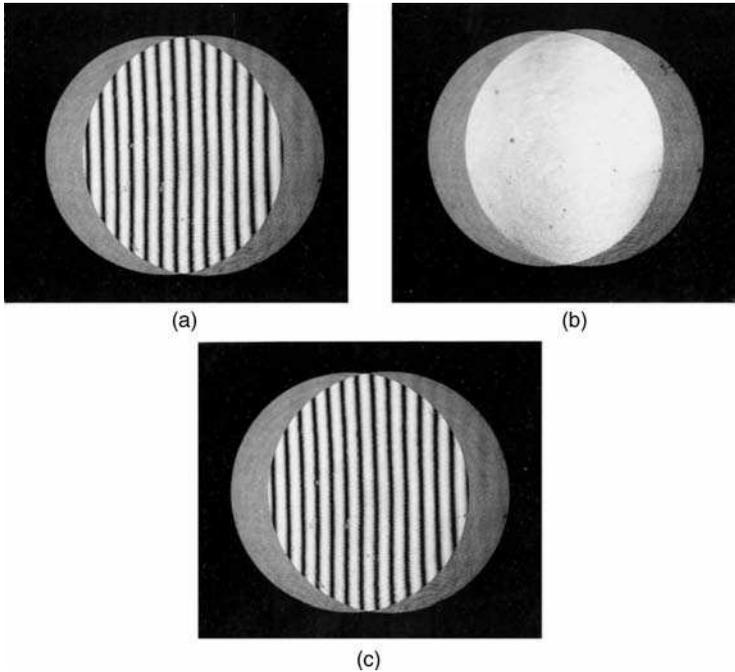


FIGURE 4.4. Lateral shearing interferograms for an aberrationless wavefront (a) inside the focus, (b) at the focus, and (c) outside the focus. In part (b), fringeless pattern is obtained when there is no defocusing. The patterns (a) and (c) are due to slight defocusing in positive and negative direction by the same amount.

This equation represents a system of straight fringes that are parallel to neither the x axis nor the y axis. Only when either the coefficient D or E is zero, they are parallel to either the x axis or the y axis, respectively.

We note the difference between the two situations represented by Eqs. (4.4) and (4.6). When there is no defocusing ($D = 0$), Eq. (4.4.) describes a uniform or fringe-free field, while Eq. (4.6) gives a system of straight fringes parallel to the x axis. When an optical system is being collimated with respect to the point source of light, we go through the region of the focus. Figure 4.4 illustrates the representative lateral shearing interferograms inside the focus (a), at the focus (b), and outside the focus (c), in a lateral shearing interferometer without tilt. When we use a lateral shearing interferometer that can also introduce tilt for the same purpose, the corresponding sequence of interferograms will be as shown in Figure 4.5. In this case, it is possible to detect slight defocusing. It is much easier to detect a change in the direction of fringes than to identify the plane with total absence of fringes.

Therefore, the ability to introduce tilt in addition to lateral shear might be a distinct advantage in certain situations. Later we talk about the use of this feature to accomplish different tasks in optical arrangements.

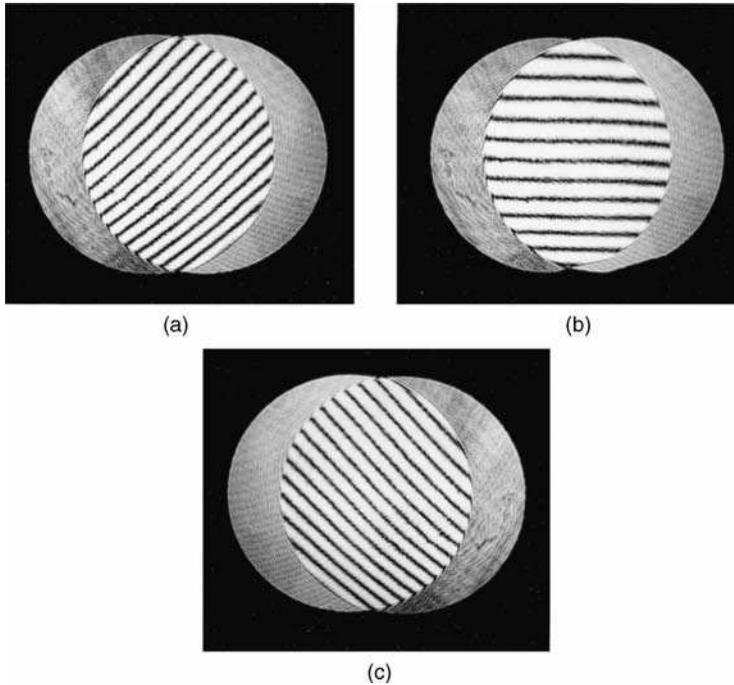


FIGURE 4.5. Lateral shearing interferograms for an aberrationless wavefront (a) inside the focus, (b) at the focus, and (c) outside the focus. In this case, however, a certain amount of tilt orthogonal to the direction of shear is introduced. At the focus, part (b), the fringes are parallel to the direction of shear. Inside and outside the focus they are inclined with respect to the shear direction.

4.3.2. Interferograms of Primary Aberrations upon Lateral Shear

Primary Spherical Aberration. The wavefront error for primary spherical aberration may be expressed as

$$W(x, y) = A(x^2 + y^2)^2. \quad (4.7)$$

Thus, the shearing interferogram may be obtained from the following equation, in the absence of defocusing term:

$$\Delta W(x, y) = 4A(x^2 + y^2)xS = n\lambda. \quad (4.8)$$

The fringe distribution can be determined from the following equation when defocusing is also present:

$$\Delta W(x, y) = [4A(x^2 + y^2)x + 2Dx]S = n\lambda. \quad (4.9)$$

Both Eqs. (4.8) and (4.9) have exponents that add to 3. Consequently, the fringes indicating the presence of the spherical aberration are cubic curves. Figure 4.6 shows

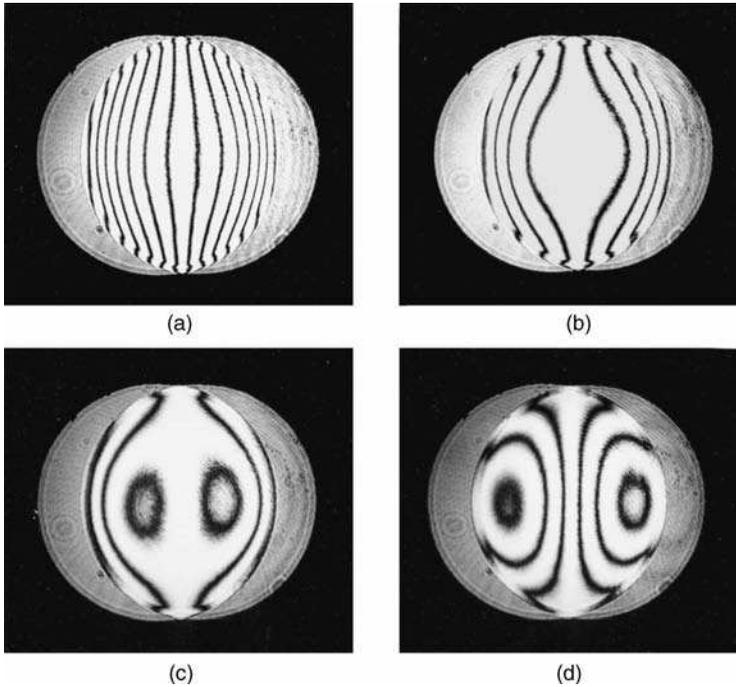


FIGURE 4.6. Typical lateral shearing interferograms of primary spherical aberration due to various amounts of defocusing: (a) inside the focus, (b) at the focus, and (c), (d) outside the focus. The fringe pattern in part (b) occurs when there is no defocusing.

the representative interferograms in the presence of primary spherical aberration in the original wavefront, inside the focus (a), at the focus (b), and outside the focus (c) and (d). In addition to the spherical aberration and defocus, tilt is also present. The equation for the fringes is given by

$$\Delta W(x, y) = [4A(x^2 + y^2)x + 2Dx]S + 2Ey = n\lambda. \tag{4.10}$$

If there are only spherical aberration and tilt present, the equation for the shape of the fringes may be predicted from the following equation:

$$\Delta W(x, y) = 4A(x^2 + y^2)xS + 2Ey = n\lambda.$$

When the primary spherical aberration is very small and there is no defocusing, Eq. (4.10) may be approximated for the central fringe close to the x axis:

$$\Delta W(x, \varepsilon) = 4Ax^3S + 2E\varepsilon = 0. \tag{4.11}$$

This equation for the central fringe gives the characteristic horizontal S-shaped curve by which very small amounts of spherical aberration may be visually identified. Typical fringe pattern is displayed in Figure 4.7.

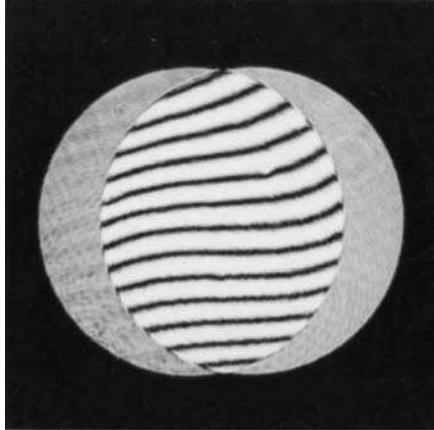


FIGURE 4.7. Typical lateral shearing interferogram of primary spherical aberration when there is a small amount of tilt in the orthogonal direction. Note the characteristic S-shape of the fringes.

Primary Coma. The wavefront error for primary coma may be expressed as

$$W(x, y) = By(x^2 + y^2). \quad (4.12)$$

In view of the unsymmetrical nature of this aberration, the shape of the lateral shear fringes is different, depending on whether the shear is in the x direction, y direction, or in some other direction.

Let us first consider the case in which the shear is in the x direction. Then the fringe shape may be found from the following equation:

$$\Delta W(x, y) = 2BxyS = n\lambda. \quad (4.13)$$

The curves represented by Eq. (4.13) are rectangular hyperbolas with the asymptotes in the x and y directions. The effect of defocusing is the addition of another term ($2DxS$), see Eqs. (4.4)–(4.13):

$$\Delta W(x, y) = 2BxyS + 2DxS = n\lambda.$$

The introduction of defocusing term results in the displacement of the center of the system of rectangular hyperbolas along the y direction. Figure 4.8 shows the representative fringe patterns for these two cases. In part (a), centered fringes are indicative of coma only. In part (b), tilt in the direction orthogonal to shear is added to coma. When tilt is added to coma in some other direction, the center of the hyperbolic fringes will be moved along a line inclined to both the x and y axes.

Next, let us consider the situation in which the shear is in the y direction. Then, if the shear magnitude along the y direction is denoted as T , the shape of the fringes can be found from the following equation:

$$\Delta W(x, y) = B(x^2 + 3y^2)T = n\lambda. \quad (4.14)$$

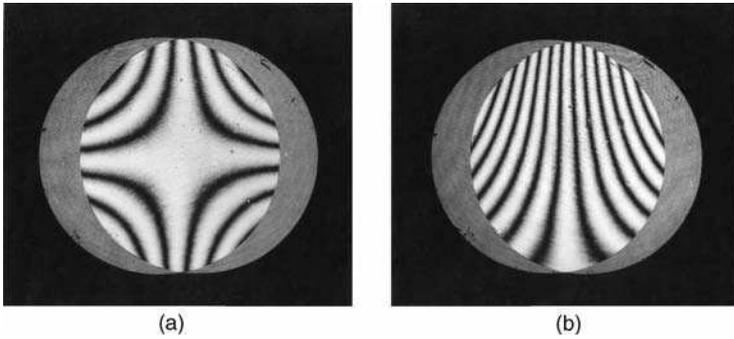


FIGURE 4.8. Representative lateral shearing interferograms of a wavefront with primary coma when the shear direction is chosen along the sagittal (horizontal, or x axis). (a) Centered rectangular hyperbolas are obtained when there is no defocusing. (b) Small defocusing causes displacement of the center of the rectangular hyperbolas (actually, the intersection of the asymptotes) in a direction perpendicular to the direction of shear (y axis).

In this case, the fringes form a system of ellipses with a ratio of major to minor axis of $3^{1/2}$. Also, the major axis is parallel to the x axis. The effects of defocusing and tilt are similar to those discussed earlier. Representative fringe patterns characteristic of coma, generated in the lateral shearing interferometer with the shear parallel to the x axis, are shown in Figures 4.9(a) and 4.9(b) with and without defocus, respectively.

Primary Astigmatism. The wavefront error for primary astigmatism may be expressed as

$$W(x, y) = C(x^2 + y^2). \tag{4.15}$$

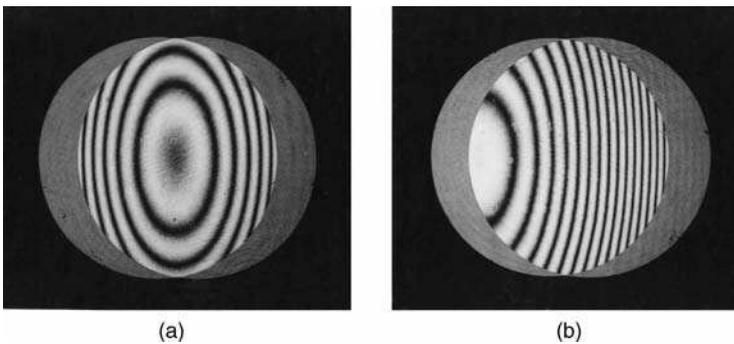


FIGURE 4.9. Representative lateral shearing interferograms of a wavefront with primary coma when the direction of shear is in the sagittal (horizontal, x) direction. (a) Centered elliptical fringes are observed when there is no defocusing. (b) When defocusing is introduced, the center of the fringe pattern moves along the direction (horizontal, x) of shear (x axis).

In the analysis of an astigmatic wavefront, the lateral shearing interferograms provide little useful information. Independently whether the lateral shear is in the x or the y direction, the straight fringes are orthogonal to the direction of shear. Thus, one could easily mistake an astigmatic wavefront for a spherical one. Fortunately, the introduction of defocusing reveals the presence of astigmatism. Then, the fringe pattern produced by lateral shear in the x direction is given by the following expression:

$$\Delta W(x, y) = 2(D + C)xS = n\lambda. \quad (4.16)$$

Similarly, the difference in wavefronts for lateral shear in the y direction becomes

$$\Delta W(x, y) = 2(D - C)yT = n\lambda. \quad (4.17)$$

If the sagittal shear S and the tangential shear T have the same magnitude, there exist two values of D , namely, $D = \pm C$, for which the lateral shearing interferogram fails to exhibit fringes. These two defocus conditions correspond to the positions of the tangential and sagittal foci of the astigmatic wavefront. Another way of detecting astigmatism is by counting the number of fringes for two perpendicular shear directions. Their different numbers indicate the presence of astigmatism, as illustrated in Figure 4.10. In part (a), with shear along the horizontal direction, we count 7 dark vertical fringes. In part (b), 10 dark horizontal fringes, upon application of equal shear in the vertical direction, indicate the presence of an astigmatic wavefront.

Another method of detecting astigmatism involves the use of lateral shear in a general direction. In this case, the system of fringes may be obtained from the following equation:

$$\Delta W(x, y) = 2(D + C)xS + 2(D - C)yT = n\lambda. \quad (4.18)$$

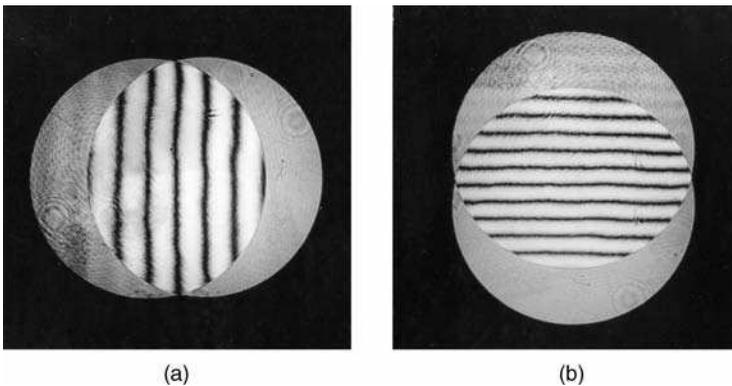


FIGURE 4.10. Representative lateral shearing interferograms arising due to the presence of primary astigmatism. Its existence is confirmed visually by counting different number of straight fringes, at a given focal setting, when the shear direction is (a) sagittal, and (b) tangential.

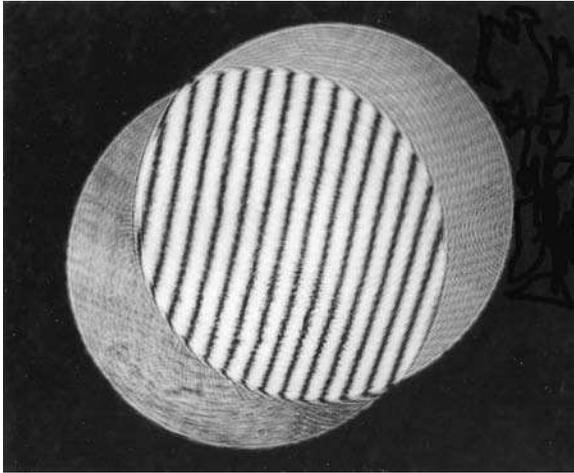


FIGURE 4.11. A representative lateral shearing interferogram due to primary astigmatism and defocus. The direction of shear is along the $y = x$ line (principal diagonal), or halfway between the sagittal and tangential directions. The inclination of straight fringes with respect to the normal to the shear direction (the other diagonal, $y = -x$) confirms the presence of the astigmatism in the wavefront.

Equation (4.18) represents a system of equally spaced straight fringes. Their slope may be found by setting n to zero and solving for y/x .

$$\frac{\delta y}{\delta x} = \frac{(C + D) S}{(C - D) T}$$

In the absence of astigmatism $C = 0$, the equidistant straight fringes are normal to the shear direction, $\delta y/\delta x = -(S/T)$, according to Eq. (4.4). By changing the direction of shear (T/S), we observe that the slope of the fringes, $\delta y/\delta x$, remains normal to the shear direction in the absence of astigmatism.

When there is astigmatism present in the wavefront, the slope of the equidistant straight fringes differs from the direction orthogonal to the shear direction. Figure 4.11 illustrates this aspect of astigmatism in relation to lateral shearing interferometry.

Curvature of Field and Distortion. Curvature of field is a displacement of focus longitudinally, and hence it can be treated as a defocusing situation. Distortion is a linear function of the pupil height. It is, in general, not detected in the lateral shearing interferometers.

Chromatic Aberration. Longitudinal chromatic aberration is a change of focus for different wavelengths. By changing the light source or using different wavelengths from the same source, one can count the number of fringes introduced by spectral

defocusing. This way, we may quantify the amount of longitudinal chromatic aberration. Similar to distortion, the lateral chromatic aberration is a linear function of the pupil height and is not detectable in a lateral shearing interferometer.

The foregoing brief account will help in evaluating the performance of an optical system by simple and quick inspection utilizing lateral shearing interferometry. We discussed the presence of pure aberrations with defocus and tilt to give us an idea of their fringe patterns. A general wavefront will have a mixture of aberrations whose presence is quantified, employing formal mathematical reconstruction techniques.

4.4. EVALUATION OF AN UNKNOWN WAVEFRONT

We shall now see how it is possible to determine the shape of a wavefront from a lateral shearing interferogram of the wavefront under study. One method, proposed by Saunders (1961, 1970), estimates the order of interference at equally spaced points along a diameter as indicated in Figure 4.12. He evaluates the wavefront by setting $W_1 = 0$, $W_2 = \Delta W_1$, $W_3 = \Delta W_1 + \Delta W_2$, and so on, obtaining the wavefront by summation of the lateral shear measurements ΔW_i . This procedure was extended to two dimensions by Saunders and Bruning (1968), and later by Rimmer (1972), and Nyssonen and Jerke (1973).

A more formal procedure assumes that the unknown wavefront $W(x, y)$ is a smooth function that may be represented by a so-called aberration polynomial, as we did in the preceding section. The wavefront difference function $\Delta W(x, y)$ is then formulated in terms of its coefficients. Values of $\Delta W_i(x_i, y_i)$ are found from measurements of the fringe positions. From those, the coefficients of the wavefront (aberration polynomial) are computed upon fitting. Malacara (1965a), Murty and Malacara (1965), and Dutton et al. (1968) developed this method in one dimension to find the wavefront shape along a diameter parallel to the shear. Malacara and Mende (1968) applied the method for the evaluation of aberrations produced by surfaces

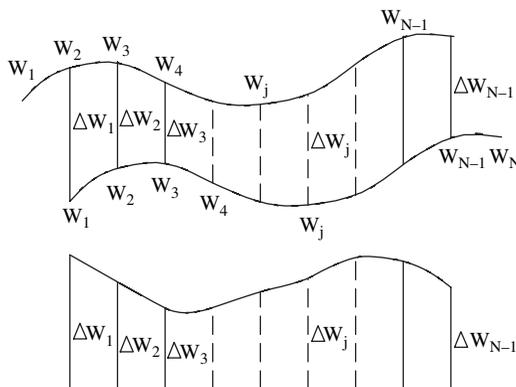


FIGURE 4.12. Original and displaced wavefront and their difference to illustrate the Saunders' procedure to find a wavefront from its lateral shearing interferogram.

with rotational symmetry. In general, the polynomial method is very good, especially after being extended to two dimensions by Rimmer and Wyant (1975) in the following manner.

The wavefront, denoted by $W(x, y)$, may be represented by a two-dimensional polynomial of degree k , of the form

$$W(x, y) = \sum_{n=0}^k \sum_{m=0}^n B_{nm} x^m y^{n-m}. \quad (4.19)$$

It contains $N = (k+1)(k+2)/2$ terms. If we wish to reconstruct the whole wavefront, we perform two shears, represented by S and T , along two mutually perpendicular directions, generating two sheared interferograms. For shear along the x direction, we obtain

$$W(x + S, y) = \sum_{n=0}^k \sum_{m=0}^n B_{nm} (x + S)^m y^{n-m}. \quad (4.20)$$

Similarly, for shear along the y direction, we find

$$W(x, y + T) = \sum_{n=0}^k \sum_{m=0}^n B_{nm} x^m (y + T)^{n-m}. \quad (4.21)$$

We may use the binomial theorem,

$$(x + S) = \sum_{j=0}^m \binom{m}{j} x^{m-j} S^j. \quad (4.22)$$

The binomial factor represents a quotient of factorials (here denoted by the exclamation mark !).

$$\binom{m}{j} = \frac{m!}{(m-j)!j!} \quad (4.23)$$

Using expression (4.22), Eq. (4.20) may be expanded into a polynomial

$$W(x + S, y) = \sum_{n=0}^k \sum_{m=0}^n \sum_{j=0}^m B_{nm} \binom{m}{j} x^{m-j} y^{n-m} S^j. \quad (4.24)$$

With expression (4.22), Eq. (4.23) may be expanded into a polynomial

$$W(x, y + T) = \sum_{n=0}^k \sum_{m=0}^n \sum_{j=0}^{n-m} B_{nm} \binom{n-m}{j} x^m y^{n-m-j} T^j. \quad (4.25)$$

These functions become equal to $W(x, y)$ when $j = 0$. We rearrange the sums to obtain the following two shearing interferogram equations:

$$\Delta W_S = W(x + S, y) - W(x, y) = \sum_{n=0}^{k-1} \sum_{m=0}^n C_{nm} x^m y^{n-m}. \quad (4.26)$$

Similarly, we find

$$\Delta W_T = W(x, y + T) - W(x, y) = \sum_{n=0}^{k-1} \sum_{m=0}^n D_{nm} x^m y^{n-m}. \quad (4.27)$$

Then we get

$$C_{nm} = \sum_{j=1}^{k-n} \binom{j+m}{j} S^j B_{j+n, j+m}. \quad (4.28)$$

Likewise, we find

$$D_{nm} = \sum_{j=1}^{k-n} \binom{j+n-m}{j} T^j B_{j+n, m}. \quad (4.29)$$

Rimmer and Wyant developed this result in 1975. The values of C_{nm} and D_{nm} may be obtained from the interferograms by means of a least-square fitting of the measured values of ΔW_S and ΔW_T to functions (4.26) and (4.27), respectively. There are $M = k(k+1)/2$ coefficients C_{nm} and D_{nm} . From them we have to determine the N wavefront coefficients B_{nm} . Expression (4.28) represents a system of M equations with M unknowns, where the unknowns are all B_{nm} coefficients with the exception of B_{n0} . Likewise, expression (4.29) represents a system of M equations with M unknowns, where these unknowns are all B_{nm} coefficients with the exception of B_{nn} . If the wavefront has rotational symmetry, $B_{n0} = B_{nn} = 0$ for all values of n , it becomes sufficient to use either expression (4.28) or (4.29) and, hence, only one interferogram is necessary. If m is different from n and also different from zero, the value of B_{nm} is found from each expression. An average of these values is appropriately taken because their difference may be explained by rounding errors in computations.

Many alternative approaches to computing the wavefront from the interferogram, generated by a lateral shearing interferometer, have been devised. This includes, for example, Gorshkov and Lysenko (1980) and others. A most interesting one is based on the wavefront expansion in terms of Zernike polynomials, first described by Rimmer and Wyant (1975) and later revised by Korwan (1983). The wavefront difference function $\Delta W(x, y)$ as well as the wavefront $W(x, y)$ is expressed as a linear combination of Zernike polynomials. Shen et al. (1997) analytically updated this work. The Zernike polynomial coefficients of the wavefront under test are

expressed in terms of the Zernike polynomial coefficients of the shearing interferograms.

Leibbrandt et al. (1996) designed a double-grating lateral shearing interferometer based on the Michelson configuration, where Zernike polynomials are used for a high accuracy wavefront reconstruction.

A number of research groups adapted the Fourier techniques after they were demonstrated as a viable tool of phase reconstruction in the early eighties. Malacara and Servin (1995) and Malacara et al. (1999) outlined the steps for the wavefront reconstruction. Liang et al. (2006a, 2006b) worked out the details and provided specific examples that are applicable to reconstruction of a wavefront even when shears are large. The wavelet transforms also proved useful for wavefront retrieval in lateral shearing interferometry (Xu et al., 2002).

4.5. LATERAL SHEARING INTERFEROMETERS IN COLLIMATED LIGHT (WHITE LIGHT COMPENSATED)

We first consider interferometric arrangements with incoherent light sources, requiring white-light compensation. Here again, we have two shear geometries, namely, lateral shear in a collimated beam and lateral shear in a convergent beam. We note, however, that it is possible to convert one into the other by the use of a well-corrected lens.

4.5.1. Arrangements Based on the Jamin Interferometer

Figures 4.13–4.15 illustrate three modifications to the Jamin interferometer to serve as lateral shearing interferometers (Murty, 1964b). The extended light source is replaced by a pinhole. The lens to be tested acts as the collimator and imprints its aberrations on the nearly collimated beam, entering the interferometer. Each glass plate on either side has the dual function of splitting the beams and reflecting it. They have to be of high quality and identical. If they are parallel to each other, the

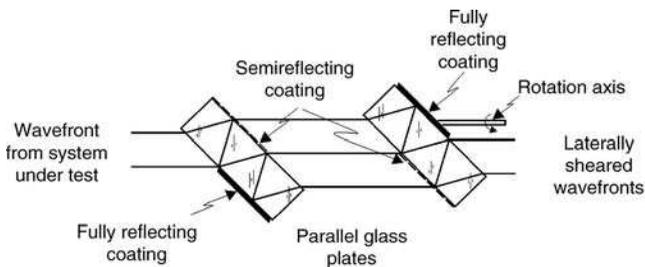


FIGURE 4.13. The layout of the principal components employed to modify the Jamin interferometer into a lateral shearing interferometer. The lateral shear may be introduced by rotating the glass plate on the right about the optical axis of the incident beam. With the zero angle of rotation of the plate, the interferometer becomes the traditional Jamin interferometer.

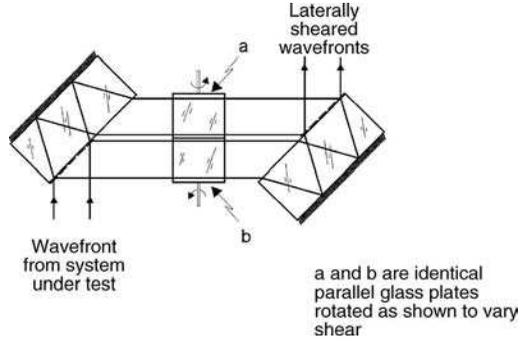


FIGURE 4.14. In another modification of the Jamin interferometer into a lateral shearing interferometer, two identical glass blocks are introduced into each beam. Lateral shear is achieved by rotating each block about an axis passing through both blocks by the same angle, but in opposite directions.

superimposed light beams emerge resulting in the traditional Jamin interferometer. In the Jamin interferometer, two glass blocks are positioned parallel to each other. Their alignment may be accomplished by observing the zero-order fringe with a broad source of light. The Jamin plates should be large enough to accommodate the beams displaced by lateral shear.

Two methods are available to introduce lateral shear into Jamin interferometer. In the first one, the glass parallel plate on the right is rotated about the optical axis of the incident beam, as indicated in Figure 4.13. With this arrangement, the shear is generated in the direction perpendicular to the plane of the paper.

The second method of producing shear in the Jamin interferometer is illustrated in Figure 4.14. Two identical parallel plates of glass are placed in each interferometer arm so that they may be rotated about the same axis, normal to the beam, by the same angle, and in opposite directions. This may be accomplished by the use of a simple gear arrangement. In an alternate setup, a fixed tilt is introduced in the orthogonal direction by making the end Jamin plates slightly wedge shaped (by a few seconds of arc).

A modification of a cyclical form of the Jamin interferometer into the lateral shearing interferometer is indicated in Figure 4.15. In this layout, only one plane parallel plate is employed. The right-angle prism or a set of two plane mirrors at an angle of 90° to each other are used to fold the light path. A transparent block is placed into one interferometer arm. It is rotated about an axis that passes through and is normal to both beams in order to introduce lateral shear. Two beams pass through the block from either side: one beam is lifted up and the other pushed down in the plane of the paper so that lateral shear is obtained as a sum of two beam displacements. This lateral shearing interferometer seems to be very convenient for testing lenses of small aperture. It is also possible to introduce tilt; it may be introduced in the orthogonal direction by incorporating slightly imperfect right-angle prism or slightly misaligned two-mirror assembly. The tilt depends on the error in the 90° angle of deviation in both cases. This arrangement may be used for measuring the accuracy of the apex-angle of the right-angle prism, the mirror alignment in the 90° mirror assembly, and

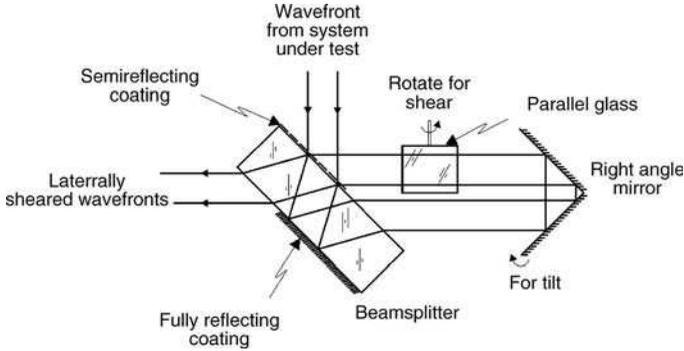


FIGURE 4.15. A modification of a cyclical form of the Jamin interferometer into the lateral shearing interferometer. A transparent block is rotated about an axis that passes through and is normal to both beams to introduce lateral shear. Two beams pass through the block from either side: one beam is lifted up and the other pushed down in the plane of the paper so that lateral shear is obtained as a sum of two beam displacements.

most importantly, detecting errors in the right angles in a corner cube prism (see also Scholl, 1995).

4.5.2. Arrangements Based on the Michelson Interferometer

The Michelson interferometer is compensated for white light, especially when adjusted for the zero-order interference position. If we place right-angle prisms or cube-corner prisms, it is possible to obtain lateral shear (Kelsal, 1959). The Michelson interferometer with the right-angle mirrors (or prisms) in each arm is presented in Figure 4.16. For simplicity, we assume that the right-angle prisms are identical in size and material. When their virtual images, as viewed in the beam splitter, are exactly superimposed on each other, no shear is generated between beams. If one prism is displaced laterally by some amount, the wavefront is sheared by twice this amount.

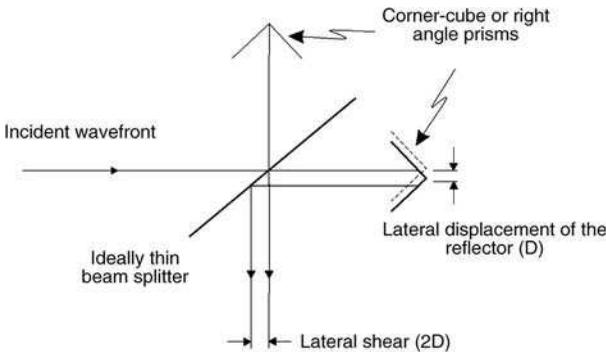


FIGURE 4.16. A modification of an equal-path Michelson interferometer into the lateral shearing interferometer. The end reflectors are either right-angle prisms (mirror assembly) or cube-corner prisms, one of them displaced in a direction transverse to the beam.

The beam divider plate must be large enough to accommodate the displacement of one beam. Similarly, at least one of the right-angle prisms must be oversized. Tilt may be implemented by tilting the right-angle prism (mirror assembly) perpendicularly to the direction of the shear. With a corner cube, only pure lateral shear is obtained because tilt may not be implemented by their nature. Tilt only may be introduced in the traditional Michelson interferometer by tilting one of the plane mirrors.

Lee et al. (1997) use the Michelson configuration in a two-dimensional shearing interferometer to measure a long radius of curvature. Pfeil et al. (1998) use a beam divider cube with a tilted end-mirror to test fast cylindrical gradient-index lenses. Albertazzi and Fantin (2002) use the Michelson configuration with the on-the-pixel processing to evaluate small incidence differences on the charged couple device (CCD). Instead of moving or tilting a reflecting component sideways, the camera is displaced sideways by one or more pixels. This may be done physically or electronically. Schreiber (2005) invented a variation of the Michelson shearing interferometer by replacing the retro reflector with a return sphere to measure tilt.

4.5.3. Arrangements Based on a Cyclic Interferometer

A cyclic interferometer is referred to as the layout where two beams travel in opposite directions, encountering exactly the same components until they emerge to form interference pattern. The Jamin interferometer presented in Figure 4.15 is one familiar example of a cyclic interferometer. We note that beams pass through the same path but they travel in the opposite directions.

A representative cyclic interferometer is a triangular-path interferometer (Hariharan and Sen, 1960). Figure 4.17 depicts a typical layout to obtain lateral shear in collimated light in this compact instrument, with minimal systematic errors. Again, we have two ways of generating lateral shear. Employing the first one, a transparent block is introduced into the path and rotated as in folded Jamin interferometer (see also Fig. 4.15). In the second method, illustrated in Figure 4.18,

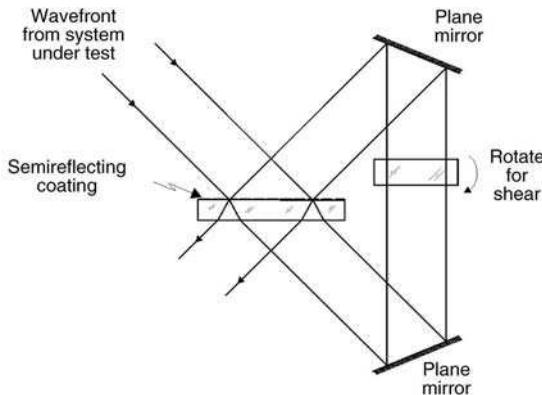


FIGURE 4.17. A representative cyclic triangular-path interferometer may be modified to obtain lateral shear in collimated light. A transparent slab is introduced into the beams. When it is rotated by an angle about the axis normal to the plane of paper and beams, they are displaced laterally in opposite directions.

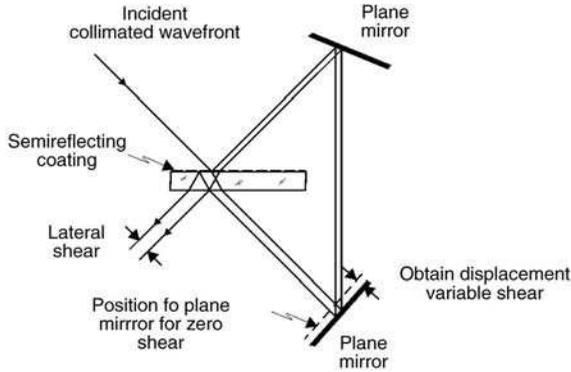


FIGURE 4.18. A modification of a representative cyclic triangular-path interferometer to obtain lateral shear in collimated light. The beam is displaced laterally when one of the plane mirrors is translated in the plane of the paper, along the beam direction.

beams will be displaced laterally when one of the plane mirrors is translated along the beam direction, in the plane of the paper.

In a symmetrical implementation of a cyclic interferometer, we may introduce a plane parallel plate of glass, as presented in Figures 4.15 and 4.17. In either case, the rotation of the glass block or plate produces lateral shear. This is also a relatively easy instrument to construct. It is insensitive to vibration and other environmental effects, including those in a controlled laboratory conditions. The simple elegance of the plane parallel plate was exposed by Waddell et al. (1994) to assess the symmetry of concave mirrors whose surface was formed with a stretchable plastic membrane.

Another compact arrangement of generating shear in a cyclic interferometer is laid-out in Figure 4.19. A $112.5^\circ, 90^\circ, 112.5^\circ$ angle pentaprism may be used as a beam splitter mirrors assembly by placing suitable coatings on the key surfaces. A 90° pentaprism is cut along its axis of symmetry. One of the cut surfaces is covered with a semireflecting coating to act as a beam splitter. They are maintained in optical contact, using index-matching liquid or oil. Translation of one half-pentaprism along

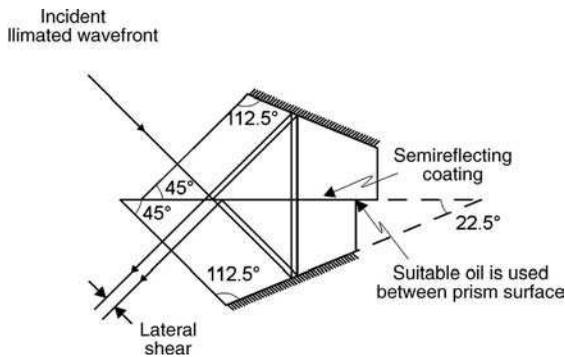


FIGURE 4.19. A modification of a pentaprism-based triangular-path interferometer into a lateral shearing interferometer. One half-pentaprism is moved along the beam-splitting surface to produce lateral shear.

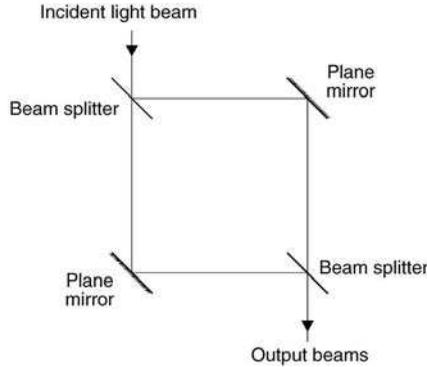


FIGURE 4.20. Simplified diagram of the principal components in the Mach-Zehnder interferometer using thin beam splitters.

the beam-dividing surface generates variable shear of the output wavefronts. Successful engineering of this device results in the heart of interferometer immune to vibrations. It is most suitable for testing small aperture optical systems due to its weight and size limitations.

Kanjilal et al. (1984) and Kanjilal and Puntambekar (1984) described another cyclic shearing interferometer that may be used with convergent light.

4.5.4. Arrangements Based on the Mach-Zehnder Interferometer

A schematic diagram of the principal optical components in the Mach-Zehnder interferometer is displayed for reference in Figure 4.20. It includes two beam dividers and two plane reflectors (mirrors). Two glass plane parallel plates of the same thickness and material are inserted in each interferometer arm to generate lateral shear in collimated light, as schematically indicated in Figure 4.21.

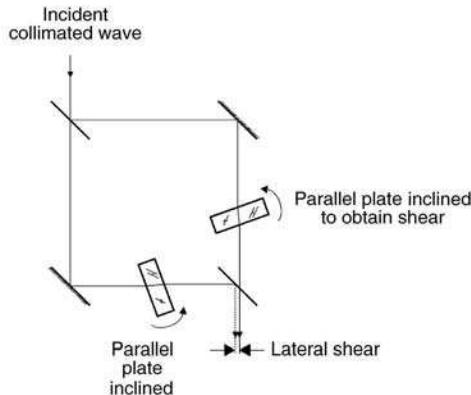


FIGURE 4.21. A basic Mach-Zehnder interferometer may be modified to obtain lateral shear in collimated light. Two identical glass plane parallel plates are inserted with an angle of inclination into each interferometer arm to generate lateral shear in collimated light. To control the amount of shear, the angle of inclination of the plates with respect to the incident beam is changed.

Each plate displaces the wavefront by the same amount and in opposite direction. The Mach-Zehnder interferometer may easily be modified for use as a lateral shearing interferometer in collimated light (Paez et al. 2000). Some researchers find it a cumbersome device to adjust, especially when all the optical elements are separately mounted and each has its own tilting screws. The Mach-Zehnder interferometer is often considered very useful for obtaining lateral shear in a convergent beam.

4.6. LATERAL SHEARING INTERFEROMETERS IN CONVERGENT LIGHT (WHITE LIGHT COMPENSATED)

4.6.1. Arrangements Based on the Michelson Interferometer

The basic block diagram relating the input-output relationship in a lateral shearing interferometer in convergent light is presented again in Figure 4.22. The nearly spherical wavefront is converging to its center of curvature where the lateral shearing interferometer is placed. Due to the space constraints, the size of the lateral shearing interferometer used in convergent light is rather small.

The output of the laterally shearing interferometer includes two beams with the amplitude decreased (by at least factor of 4) and the same size of the radius of curvature as the input beam (but opposite in sign). The diverging beams are tilted with respect to each other. There is an angle between the radius of curvature to the central point on one wavefront and the radius of curvature connecting the center of the “displaced” wavefront. If we think of the wavefront as forming a part of a spherical surface, then the sheared wavefront slides over this spherical surface. The quantity of interest is the difference between the sphere and the actual wavefront, which we are accustomed to call the wavefront aberration function, $W(x, y)$.

A converging beam of light from the optical system under test enters a simple Michelson interferometer, having a thin beam splitter, as sketched in Figure 4.23. When two plane mirrors are positioned symmetrically with respect to the beam splitter (equal path condition) and perfect optical components are used, a broad zero-order fringe presents itself in the interference plane. When the plane mirrors are used, the split incident beams converge on a point on the mirror surface. This is the

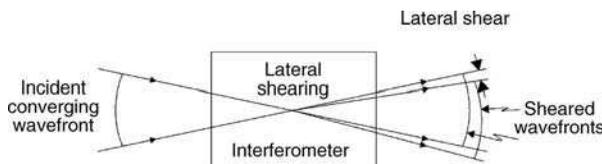


FIGURE 4.22. The basic block diagram relating the input-output relationship in a lateral shearing interferometer in convergent light. The nearly spherical wavefront is converging to its center of curvature where the lateral shearing interferometer is placed.

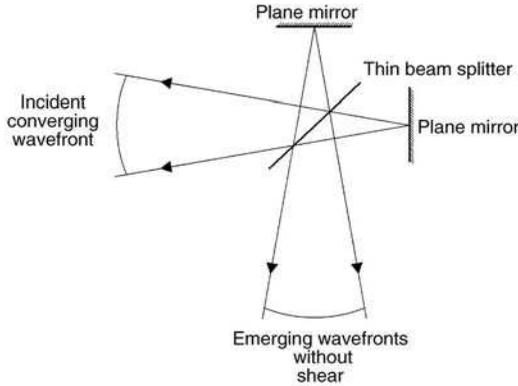


FIGURE 4.23. A converging beam of light from the optical system under test entering a simple Michelson interferometer with a thin beam splitter. When two plane mirrors are positioned symmetrically with respect to the beam splitter, a broad zero-order fringe presents itself in the interference plane. When the plane mirrors are used, the split incident beams converge on a point on the mirror surface.

layout of the basic Michelson interferometer: the emerging wavefronts are superimposed without any lateral shear.

The lateral shear may be introduced by rotating one of the mirrors by a small angle about an axis that coincides with the wavefront center of curvature. This is illustrated in Figure 4.24. The wavefront aberration is made visible in the interferometric pattern. There are no fringes in the absence of aberration. If the centers of curvature fall ever so slightly outside the plane reflectors, defocusing is introduced. Even in the absence of aberrations, straight fringes indicative of defocus condition in the shearing interferometer are displayed. It is not possible to obtain tilt in the orthogonal direction.

A sturdy and compact version of this interferometer is made with two nearly identical right-angle prisms and cementing them together along partially transmitting

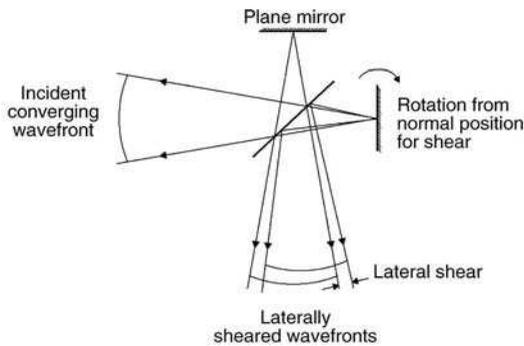


FIGURE 4.24. A modification of a Michelson interferometer into a lateral shearing interferometer. Tilting one of the plane mirrors about an axis passing through the point of convergence of the wavefront on the mirror surface introduces lateral shear.

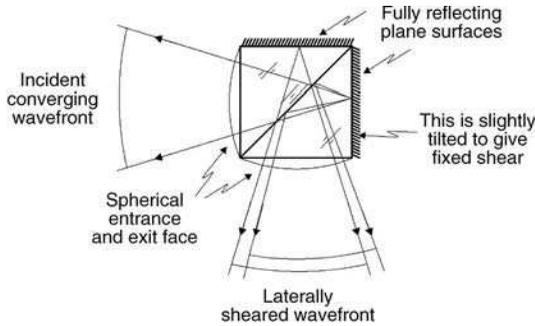


FIGURE 4.25. A modification of a Michelson interferometer into a lateral shearing interferometer with a cemented cube. The plane surfaces are deformed to introduce a fixed amount of shear. The entrance and the exit faces may be given convex spherical shapes so that the rays enter and exit normally.

diagonal. A fixed amount of shear may be implemented with one slightly tilted totally reflecting surface, depicted in Figure 4.25. Lenouvel and Lenouvel (1938) were probably the first to devise the practical solid version of this interferometer.

Later, Murty (1969), Saunders (1970), and several others found simplicity in this interferometer very useful in the testing of optical systems. One can vary the amount of shear using similar right-angle prisms in a cube arrangement (Murty, 1970). An index matching liquid or suitable oil inserted between the faces of the hypotenuse, permits rotation of one prism. The axis of rotation passes through both centers of curvature of split wavefronts, as illustrated in Figure 4.26. The entrance and exit faces may be made spherical so that the rays enter and exit almost without deviation. This implementation of the lateral shearing interferometer has been used to characterize cryogenic laser fusion targets (Tarvin et al., 1979).

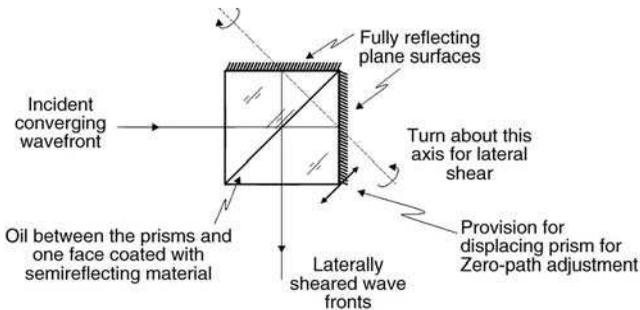


FIGURE 4.26. A layout of a lateral shearing interferometer based on the Michelson interferometer. The lateral shear may be varied by rotating one or both prisms about an axis passing through the centers of curvature of the incident beams.

4.6.2. Arrangements Based on the Mach–Zehnder Interferometer

Let us consider a simplified Mach–Zehnder interferometer with idealized thin beam dividers as indicated in Figure 4.27. A converging wavefront is focused on the plane mirrors, placed symmetrically with respect to infinitely thin beam splitter. If one of these mirrors is rotated about an axis, passing through the geometrical focal point, normal to the plane of the paper, a lateral shear is established between the beams emerging from the interferometer. Tilt, orthogonal to the direction of shear, may be introduced upon rotation of the second beam divider about an axis normal to the plane of paper (also illustrated in Fig. 4.27). The introduction of this tilt has been offered as an important feature of this interferometer. It might be advantageous to bring both beams to focus on the second beam splitter. In this case, the roles of the plane mirror and the beam divider are interchanged: rotation of one mirror introduces tilt, and the rotation of the second beam splitter results in introduction of shear. When one is using actual components with finite thicknesses, compensation plates need to be inserted judiciously.

We found it relatively simple to align a shearing interferometer incorporating tilt, based on the Mach–Zehnder interferometer (Paez and Strojnik, 2000; Paez and Strojnik, 2001). Many researchers, however, believe that it is difficult to adjust quickly an equal path interferometer, incorporating as a minimum two beam dividers and two plane mirrors, each separately mounted on a suitable base and capable of all possible adjustments. Some examples of pre-aligned and pre-adjusted devices based on the Mach–Zehnder configuration are presented next. Most of them use solid glass polygon prisms to generate a fixed amount of lateral shear. Saunders (1965) devised almost a literal solidification of the layout in Figure 4.27. The space between the beam dividers and the plane mirrors is filled with glass, as seen in Figure 4.28. The convergent beam focuses on surfaces parallel to each other but not parallel to the beam divider. The fabrication of this compact component could start with the

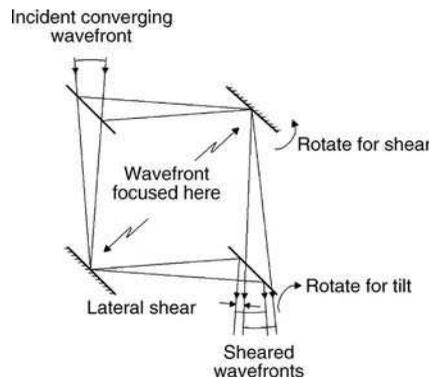


FIGURE 4.27. A modification of a simplified Mach–Zehnder interferometer into a lateral shearing interferometer, in convergent light. The converging wavefront comes to a geometrical focus on the surface of plane mirrors. Either mirror is rotated by a small angle to generate shear. Tilt may be introduced by rotating the second beam splitter by a small amount.

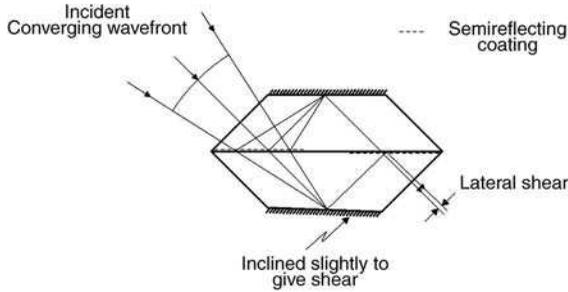


FIGURE 4.28. A solidification of the lateral shearing interferometer sketched in Figure 4.27. The space between the beam dividers and the plane mirrors is filled with glass. The convergent beam focuses on surfaces parallel to each other but not parallel to the beam divider. Top and bottom surfaces, slightly inclined with respect to the beam splitter surfaces to produce fixed shear, are coated to function as mirrors. A fixed amount of tilt about the direction orthogonal to shear is possible with slightly wedged cement.

formation of hexagonal a prism, cut in two prisms halving 90° angles. The cut surfaces are coated to act as beam splitters and then cemented together. Top and bottom surfaces, slightly inclined with respect to the beam splitter surfaces to produce fixed shear, are coated to function as mirrors. A fixed amount of tilt about the direction orthogonal to shear is possible with slightly wedged cement.

A shearing interferometer in Mach–Zehnder configuration with mirrors tilted may be used for flame visualization. Philipp et al. (1993) developed a technique to assess flame dynamics using cube beam splitters and collimated light.

A possible improvement to the lateral shearing interferometer featuring variable shear and based on the Mach–Zehnder layout is indicated in Figure 4.29. Here two rhomboidal prisms are cemented in a symmetrical fashion along the semireflecting beam-dividing surface. The top and bottom surfaces are mirrored. The incident converging beams are focused on the second beam divider, implemented as a cube beam splitter. By rotating it about the axis through the center of curvature, a limited amount of shear may be introduced. The angle-of-incidence effects limit the amount of acceptable rotation. The angle of rotation and size of the cube determine the

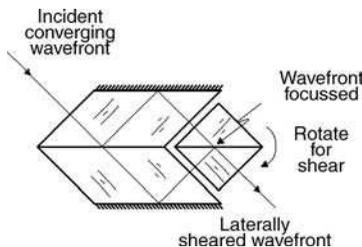


FIGURE 4.29. A possible improvement to the lateral shearing interferometer featuring variable shear and based on the Mach–Zehnder layout. Beams converge on the cemented half-transmitting diagonal of the cube beam splitter. By rotating it about the axis through the center of curvature, a limited amount of shear may be introduced. A fixed amount of tilt may be incorporated into the wavefront by inclining the reflecting surfaces.

amount of displacement. A fixed amount of tilt may be incorporated into the wavefront by inclining the reflecting surfaces.

This configuration may also be seen as a version of the shearing interferometer sketched in Figure 4.28, with the cube beam splitter cut off from the hexagonal prism assembly and slightly displaced to the right. By converging the beam on the second beam splitter, the function of shear and the perpendicular tilt corresponds to the second case, discussed with respect to Figure 4.27.

Another improvement to the octagon in Figure 4.28, with planar surfaces replaced by sections of a sphere, is indicated in Figure 4.30. It employs two identical plano-convex spherical lenses whose center thickness is equal to half its radius of curvature. A small on-axis section of each lens is flattened, polished, and coated with a reflecting material. Each mirrored surface is slightly inclined with respect to the planar half-transmitting surface to introduce a fixed shear. The planar half-transmitting surfaces are cemented together to function as a beam divider. A beam of large numerical aperture may be tested without introducing its own errors because entrance and exit faces are spherical in form.

Saunders (1964a) proposed another modification of the Mach–Zehnder interferometer into a lateral shearing interferometer. The principal components, depicted in Figure 4.31, include two identical prisms, possibly cut along the principal diagonal from a single four-sided (kite-like) prism with angles $\alpha = 120^\circ$, $\beta = 100^\circ$, $\gamma = 40^\circ$. The main diagonal surface is coated to divide the beam into two, functioning as a beam splitter. The wavefront converges on the inclined mirrored faces. The lateral shear may be implemented by rotating (one of) the prisms about the line joining the two centers of curvature of the two prisms in opposite directions. This layout may be simplified to perform fixed shear by cementing prisms at an appropriate angle.

The systematic aberrations inherent to this device may be completely eliminated by cementing plano-convex lenses of suitable radii of curvature on the entrance and exit faces. (See discussion about Fig. 4.30.) Tilt may be implemented by suitable

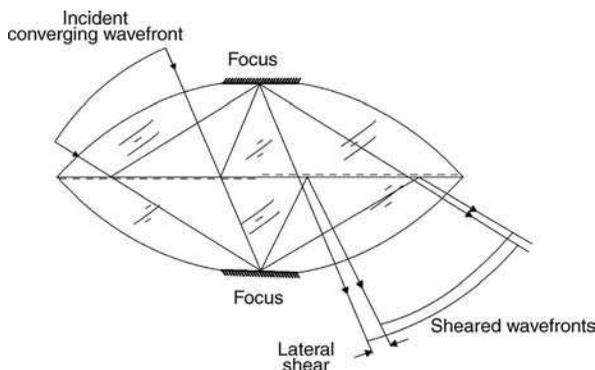


FIGURE 4.30. An improvement to the octagon in Figure 4.28, with planar surfaces replaced by sections of a sphere. Two identical plano-convex spherical lenses whose center thickness is equal to half its radius of curvature are coated along the planar part and cemented to function as a beam divider. A small on-axis section of each lens is flattened, polished, and coated with a reflecting material. Each mirrored surface is slightly inclined with respect to the planar half-transmitting surface to introduce a fixed shear.

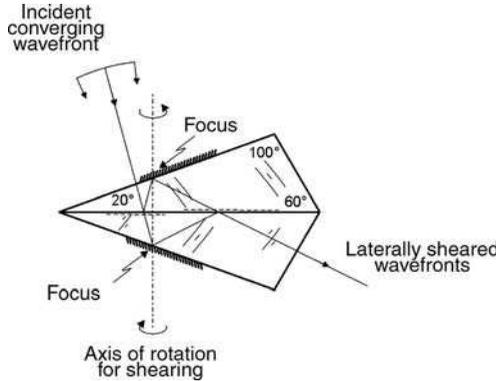


FIGURE 4.31. Another modification of the Mach–Zehnder interferometer, into a lateral shearing interferometer, using two obtuse-angle prisms. The main diagonal surface is coated to divide the incident wavefront into two, functioning as a beam splitter. The wavefront converges on the inclined mirrored faces. The lateral shear may be implemented by rotating prisms relative to each other. This layout may be simplified to provide fixed shear by cementing prisms at an appropriate angle.

wedging of the cement and by coating only partially the beam divider, covering only half of each prism-base surface.

4.7. LATERAL SHEARING INTERFEROMETERS USING LASERS

As we pointed out earlier, any white-light compensated lateral shearing interferometer works with a laser source. It is possible to devise interferometers that are simpler in constructions and use when requirement for use with white light is eliminated. The most convenient (and cost-effective) laser is the helium–neon (He–Ne) laser emitting a few milliwatts of power at the 632.8 nm line.

A plane parallel plate, devised by Murty (1964a), is undoubtedly one of the most elegant and simplest lateral shearing interferometers for qualitative work in a laboratory. Figure 4.32 shows that the coherent laser radiation is focused by a microscope objective on a pinhole to clean up the beam. The lens (group) under test, having the same f -number as the microscope objective, collimates the divergent light beam. It is incident on an inclined plane parallel plate, normally used without any coating on either surface. The light is reflected from the front and the back of the plate due to Fresnel reflections. The plate thickness introduces the beam displacement, that is, a lateral shear. The lateral displacement S for a plate of thickness t , refractive index N , and for the beam angle of incidence i is given by Malacara (1965a).

$$\frac{S}{t} = \sin 2i(N^2 - \sin^2 i)^{1/2} \tag{4.30}$$

Figure 4.33 presents the graph of S/t versus angle of incidence i , for glass with the index of refraction $N = 1.515$ at 632.8 nm. The angle of incidence is measured from

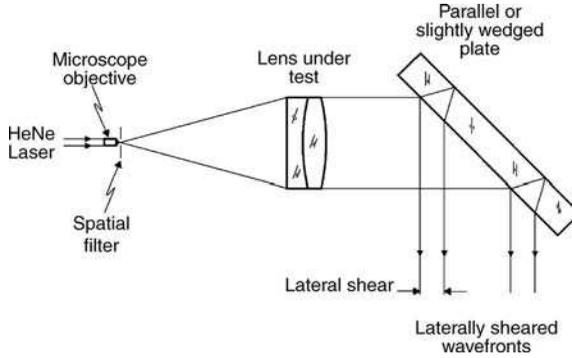


FIGURE 4.32. Schematic diagram of a lateral shearing interferometer incorporating a laser source and a plane parallel plate of glass.

the normal to the glass surface, as in the Snell law. The maximum value of the ratio S/t of approximately 0.8 corresponds to an angle of incidence of 50° . Therefore, a 45° angle of incidence is quite convenient to use in a practical setup. A slight wedge is often introduced into a plate to guide ghost reflections from the main beams.

The intensity of the peak of the fringe pattern increases by coating the front and back surfaces without changing the fringe visibility. Then the internal reflection has enough intensity to exhibit faint secondary sheared beams and interference patterns. Therefore, many practitioners find it advantageous to use an uncoated plate.

Figure 4.34 shows the layout with a plane parallel plate to test a large-diameter concave mirror. A suitable null correcting system may be inserted when the mirror is not spherical. A fixed amount of tilt orthogonal to the shear direction is produced by introducing a small wedge into the shearing plate. The imaginary line, where the two planar surfaces of the wedge intersect, is parallel to the plane of the paper.

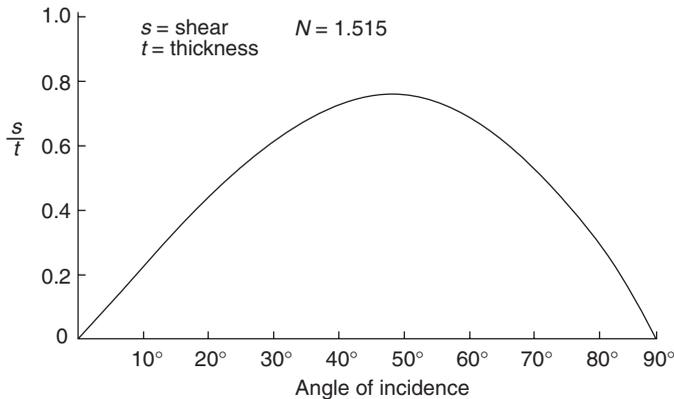


FIGURE 4.33. A graph of the ratio of displacement over plate thickness (S/t) as a function of angle of incidence, for a typical borosilicate crown glass. Examining the plot we note that any angle of incidence up to a maximum of about 50° is convenient.

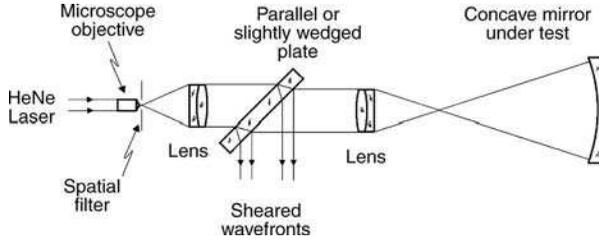


FIGURE 4.34. Optical layout of a plane parallel plate (interferometer) for testing large concave mirrors. In addition to a coherent source, two excellent lenses and a good one are needed for performing this test.

Another variation (Hariharan, 1975) of the basic parallel plate interferometer uses two plates whose separation may be adjusted to expand the range of the possible shear. In Figure 4.35, one of the plates is mounted on a movable platform so that the air gap is variable and, hence, variable shear is obtained without plate rotation. The inside surfaces of these plates are uncoated so that about 4% reflection occurs there. The outer surfaces are treated with high quality antireflection coatings to minimize reflections. The plate mounting may be equipped with additional degree of freedom so that it can be rotated orthogonally to introduce tilt. The physical separation of two plates makes this system sensitive to disturbances and fragile in handling. The stability of the fringe system depends on the quality of mechanical mounts, their design, fabrication, and environmental conditions.

The lateral shearing interferograms displayed in Figures 4.4–4.11 were generated employing a laser shearing interferometer laid out in Figure 4.32. An uncoated glass plate with a laser source results in a sufficiently intense fringe pattern, projected on a ground glass in a dimly illuminated room. An exposure time of only a fraction of a second is sufficient to photograph the fringes. With the currently available CCD detector arrays, the experimental work is even simpler.

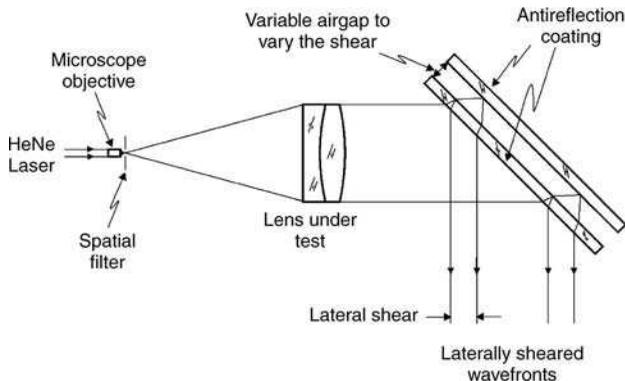


FIGURE 4.35. Modification of the plane parallel plate interferometer by incorporating two separate glass plates. The reflections from the inner surface are utilized to generate the expanded amount of lateral shear. The lateral shear is variable by moving the back plate to change the width of the air gap.

4.7.1. Other Applications of the Plane Parallel Plate Interferometer

One of the most popular uses of the plane parallel plate interferometer is in checking beam collimation. Returning to Figure 4.32, we note that the wavefront is either slightly divergent or convergent if the pinhole is not located at the focus of the lens under test. Then, the interference pattern of a high-quality collimating lens exhibits straight fringes, which is indicated in Figures 4.4(a) and 4.4(c), rather than uniform field of Figure 4.4(b).

The common area of two sheared apertures will be free of fringes, as noted in Figure 4.4(b), when the pinhole center coincides with the focal point of the collimating lens. The best collimation is found when the interference pattern changes from that shown in part (a) to (b) and, finally to (c), upon slight displacement of the lens along the optical axis.

The configuration of exact collimation may be even more accurately determined when a wedged plate is used for alignment. The plate is first used normal to the beam emerging from the collimating lens. The reflected pattern displays the Fizeau fringes indicating the direction of the wedge. In this position the shear is zero, so slight decollimation is not of significance. The plane parallel plate is then rotated in its own plane. The Fizeau fringes rotate along, always indicating the direction of the wedge. In the last step, the plate is tilted at an angle with respect to the incident beam so that the angle of incidence is about 45° . If the pinhole center is slightly outside the focus, inclined fringes will be observed. By moving the high-quality lens longitudinally along the optical axis, the pinhole will be centered on the focal point when horizontal fringes are displayed. Parts (a), (b), and (c) of Figure 4.5 display the sequence of fringe patterns as the pinhole passes through the focus. The addition of tilt introduced by the wedged plate allows the experimenter to follow the fringe rotation rather than to identify fringe free field, which may be difficult to identify with precision.

Determination of inhomogeneity of solid transparent samples is an additional application of the plane parallel plate or wedge plate interferometer. The sample has to be prepared in the form of a parallel piece (of glass) and sandwiched between two very good plane parallel plates using optical contact. This optical assembly, indicated as a solid block in Figure 4.36, is placed between the collimating lens and the

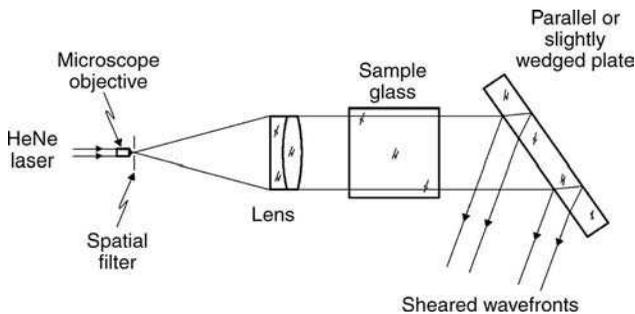


FIGURE 4.36. An optical layout for evaluating the degree of inhomogeneity of a transparent (glass) sample using a plane parallel plate interferometer.

shearing plate. This layout is reminiscent of the shearing modification of the Jamin interferometer (see Figs. 4.14 and 4.15). When a high-quality plane wave passes through the inhomogeneous sample, the distortion is imprinted on it. The laterally sheared wavefronts produce fringe patterns corresponding to the change in the index of refraction. The variation of the refractive index across the beam (along the shear direction) inside the block is determined upon integration. The changes of the index of refraction along the optical axis are averaged. The block needs to be reoriented to find changes along all three dimensions.

The same region between the high-quality collimating lens and the shear plate may be used for flow studies, diffusion studies, and other research into position and time-varying phenomena.

A number of researchers studied the applications of the plane parallel plate to alignment and collimation, starting with Dickey and Harder (1978), Grindel (1986), and Sirohi and Kothiyal (1987a, 1987b). Lens parameters, such as the focal length or the refractive index, may also be measured in this setup (Kasana and Rosenbruch, 1983a, 1983b; Murty and Shukla, 1983).

The evaluation of infrared materials has been done with this kind of interferometer. Venkata and Juyal (1987) adapted the plane parallel interferometer to the testing in infrared by using a CO₂ laser, a plane parallel plate made out of zinc sulfide (ZnS), and a phosphor screen to observe the fringes. The fabrication of ZnS plate to the high flatness requirements is a formidable achievement in itself.

The surface imperfections of a large concave mirror may be tested using a plane parallel plate, with or without a wedge, in a layout indicated in Figure 4.37 (Malacara, 1965a). The first of the two well-corrected lenses collimates the laser beam. The shearing plate is inclined at about 45° inside the collimated light.

This layout may also be used to measure the radius of curvature of a spherical surface. In the second position, the right well-corrected lens brings the collimated beam to focus on the vertex of the concave mirror. Then the lens is moved left along the optical axis to its first position so that its focus coincides with the center of curvature of the concave mirror. In both settings, the wavefront incident on the (wedged) plane parallel plate is planar. In the interference plane, one sees either a blank field or a field containing horizontal fringes depending on the absence or presence of tilt in the shearing device (i.e., whether a plane parallel plate or a wedged plate is used). The lens displacement along the optical axis between these two positions (first and second), equal to radius of curvature of the concave mirror R ,

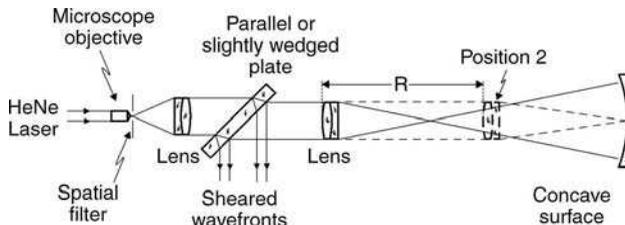


FIGURE 4.37. Optical layout of a (wedged) plane parallel plate interferometer to measure the radius of curvature of a concave spherical surface.

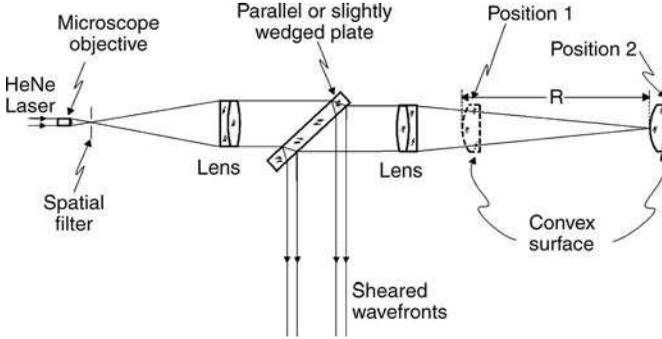


FIGURE 4.38. Optical layout of a (wedged) plane parallel plate interferometer for measuring the radius of curvature of a convex spherical surface.

must be accurately measured by some other method to avoid fringes that arise from defocusing.

Figure 4.38 presents the necessary changes for the measurement of the radius of curvature of a convex surface. In this case, the longest radius of curvature that can be measured is slightly smaller than the focal length of the second lens. All components in the interferometric arrangements have to be of high quality so/as not to introduce their aberrations into test results. Likewise, two collimating lenses in these applications have to be very good. Nowadays with the use of computers, their designs are analyzed and optimized as classroom examples. Malacara (1965b), among others, proposed several early designs.

Another application of the interferometer is for the determination of the refractive index of nearly parallel plates of glass or liquids contained in a parallel-sided glass cell. The basic arrangement is similar to the one shown in Figure 4.34, where we use two well-corrected lenses to obtain the sharp and well-corrected focal point. The second lens is adjusted for the positions corresponding to the two retroreflecting situations from the two faces of the parallel plate as shown in Figure 4.39. The difference between the two positions is t/N , where t is the thickness of the plate and N is the refractive index. If the thickness of the plate is independently measured, the

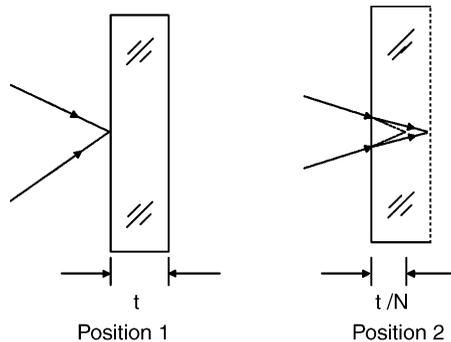


FIGURE 4.39. Schematic for the determination of the refractive index of a parallel plate.

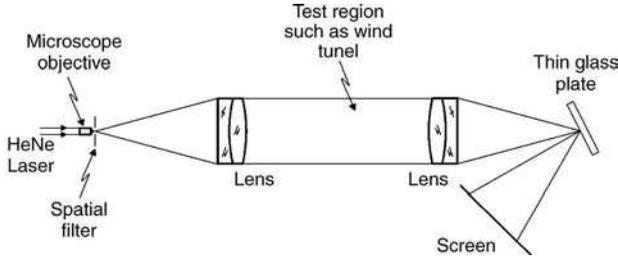


FIGURE 4.40. Lateral shearing interferometer utilizing a thin parallel plate at the point of convergence.

refractive index is determined easily. It is possible to determine the refractive index to five significant digits. Some inaccuracies may arise from the aberrations of the parallel plate and also from the temperature dependence of the index of refraction.

A very small, thin parallel plate (Tanner, 1965) may be used at the focal spot of the system as illustrated in Figure 4.40. Laser light is collimated by the first lens, and after leaving sufficient space in the collimated region, the light is again focused by the second lens. At a position very close to the focus, a thin glass plate is inserted at about 45° angle of incidence. As can be easily seen from Figure 4.41, it is not possible to obtain pure lateral shear because of the longitudinal separation between the two images, reflected from the two surfaces of the plate. Even for a well-corrected optical system, the lateral shear fringes are slightly curved as shown in Figure 4.42. Ideally, this system requires an extremely thin plate of glass with a wedge between the surfaces. Alternatively, the system may be made in the form of an air wedge. Even then only a very limited amount of lateral shear is generated. It is often the system of choice in large aperture wind tunnel applications, homogeneity measurements, and so forth, shown in Figure 4.43.

Griffin (2001) proposed the use of liquid crystal to fill the gap between two glass plates in a plane parallel configuration. This introduces a region of phase-shifted

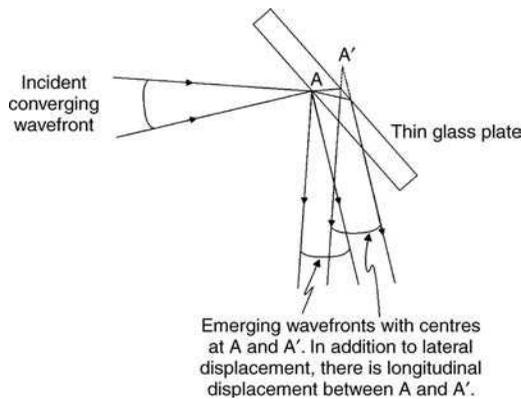


FIGURE 4.41. Ray diagram indicating that pure lateral shear cannot be obtained when the thin glass plate is located at the focus.

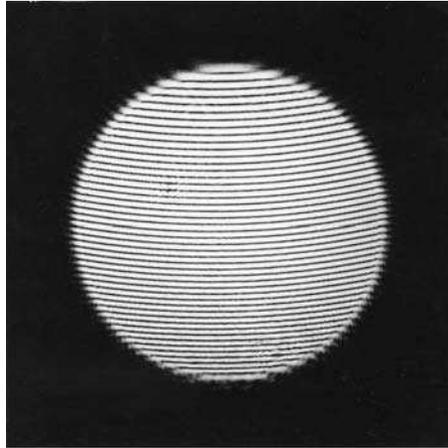


FIGURE 4.42. Typical appearance of the lateral shear fringes in interferometer shown in Figure 4.40.

shear that may be easily controlled. Chang et al. (2002) raised concern that the amount of phase shift is actually not completely repeatable. It possibly offers the source of error upon introduction of unknown phase shift.

Suhara (2002) used a wedge with gap to measure the distribution of the index of refraction in plastic lenses, by additionally performing computational tomography. One very important application of the lateral shearing interferometer incorporating a wedge is in vivo measurement of tears and artificial tears in human eye. Lechna-Marczynska et al. (1999) were the first to use a simple wedge to perform such measurements. Dubra et al. (2004, 2005) recently took up this work. They used double lateral shearing interferometer, also referred to as a three-dimensional shearing interferometer to profile tear surface.

When the wedge moves in the direction along one of its prismatic surfaces, the optical path of the incident beam changes, that is, the wedge thickness as seen by the ray changes. This feature may be used to introduce the phase change into the wavefront by moving the wedge. The controlled introduction of the phase change

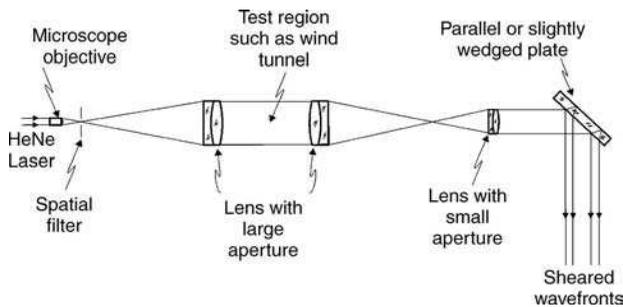


FIGURE 4.43. Schematic arrangement of the parallel plate interferometer for large aperture wind-tunnel applications.

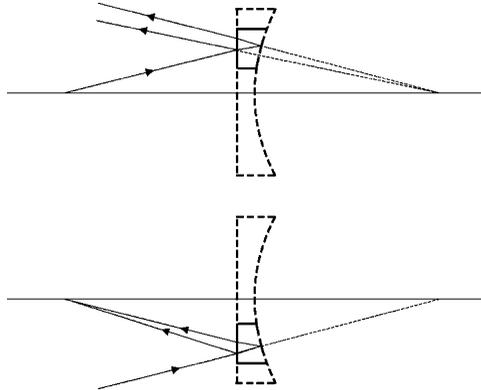


FIGURE 4.44. In a modification of the parallel plate interferometer for convergent or divergent wavefronts, an off-axis section of plano-concave divergent lens is used instead of the plate.

facilitates the phase reconstruction without the need for an expensive piezoelectric-mounted mirror. Lee et al. (2003) added an additional degree of freedom to the prisms in the vectorial-shearing interferometer by moving one prism with respect to the other along their common hypotenuse. Song et al. (2004a, 2004b) simplified this interferometer to introduce a controlled amount of phase by moving a wedged mirror.

More recently, Mehta et al. (2005a) proposed a lateral shearing interferometer incorporating a slightly wedged plate to design a versatile system to measure distance with an extended range of measurements and improved resolution. Mehta et al. (2005b) used the same wedge plate setup with two-wavelength fringe projection to develop a three-dimensional profilometer.

The plane parallel plate interferometer may be modified to obtain lateral shear in the divergent or convergent beam (Malacara et al., 1975). As sketched in Figure 4.44, an off-axis divergent lens section replaces the plane parallel plate. These modifications introduce a certain amount of radial shear in addition to the desired lateral shear, limiting its applications.

Schwider (1980) modified the plane parallel plate interferometer to make possible the fringe generation with white light. His method, called superposition fringe, is based on the chromatic compensation by means of a Fabry-Perot interferometer, placed in the collimated beam entering the shearing plate. This shearing plate is really formed by two plates, like in Figure 4.35.

4.8. OTHER TYPES OF LATERAL SHEARING INTERFEROMETERS

The shearing interferometers considered so far employed the beam-dividing surfaces to divide amplitude. This was accomplished by covering the surface with a semireflecting coating or by relying on reflection at the air-glass boundary.

There exist other types of interferometers to generate lateral displacement where additional optical principles are employed. Interestingly, the first type of shearing

interferometer was based on this alternate technology. We treat these briefly here because some of them are discussed in more detail in other chapters.

4.8.1. Lateral Shearing Interferometers Based on Diffraction

We consider a convergent beam incident normally to the transmission-type diffraction grating with a periodicity d . The beam converges on the point on the grating plane. Let the cone half-apex angle of the converging beam be α . Using the formula for the diffraction grating, we may give the direction of the central ray in the first-order beam, θ , also known as the angle of diffraction, by

$$\sin \theta = \frac{\lambda}{d}. \quad (4.31)$$

With proper choice of grating period d , related to the other experimental parameters, the zeroth- and the first-order beams overlap, as sketched in Figure 4.45. For diffraction gratings with zero-one transmission profile, higher order beams are also present. Their peak beam irradiance decreases with the increasing diffraction order, diminishing visibility of their fringe patterns.

By tailoring their transmission profile, it is possible to manufacture gratings that transmit only the zeroth- and the first-order beams. The aberrations of the original wavefront are preserved in the zero-order and first-order beams for moderate diffraction angles. It is considered convenient that both order beams remain separated so that only two-beam interference pattern is generated (rather than three or more). To satisfy these requirements, the following condition needs to be met:

$$\theta \geq \alpha. \quad (4.32)$$

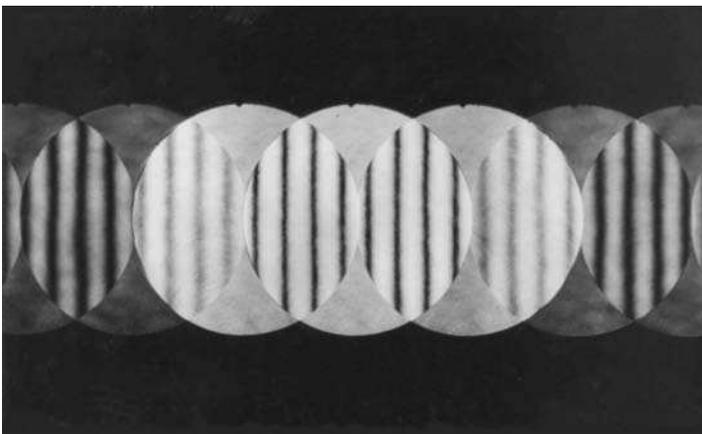


FIGURE 4.45. Typical appearance of a Ronchi interferogram when a grating with proper spacing is chosen for the Ronchi “grating”. Two first order beams just touch each other and pass through the center of the zeroth order beam.

This expression may be rewritten in terms of grating distance d as

$$d \leq 2\lambda(f\text{-number}). \quad (4.33)$$

Under optimal conditions, the beam circumferences barely come in contact, as illustrated in Figure 4.45. These interferograms are named after their inventor, Ronchi, who also lent his name to the optimized interferometric configuration. Then the equality is applicable in expression (4.33). The choice of grating period depends on the f -number of the optical system. For example, by choosing a system with f -number = 5 and the wavelength $\lambda = 0.5 \mu\text{m}$, the value of $5 \mu\text{m}$ results for the grating period d . Hence a grating having 200 rulings per millimeter must be used in this application.

The grating interferometer is also white-light compensated. In principle, ordinary light sources may be employed provided that a pinhole of proper diameter is used to limit the source size. It is quite convenient to use the grating interferometer with a laser source.

The existence of laser source makes it possible to fabricate gratings with several periodicities. This is easily accomplished by recording interference fringes between plane waves incident from different directions. The angle between the plane wavefronts is changed to obtain several periodicities. This type of interferometer is due to Ronchi (1923). Among his many publications, an extensive review of the subject (Ronchi, 1964) provides an excellent summary of his work.

A lateral shearing interferometer incorporating two gratings may be used to optically address the spatial light modulator, which in turn shapes the output beam of the Nd:YAG continuous laser beam (Scholl, 1980). Wattellier et al. (2002) describe the setup to measure and shape the laser beam in the spatial frequency plane.

In order to avoid multiple beam interference, a convenient amount of lateral shear is equal to about half the diameter of the beam. In all other shearing applications, this is actually considered a relatively large amount of shear. However, with a single grating, it is not possible to obtain a lesser amount of lateral shear without confusion of three-beam overlap and interference in the spatial domain.

This could be perceived as a restriction only when working in the spatial domain. With the wide-spread employment of the Fourier techniques in optics, especially for the spatial filtering, overlapping beams in the space are easily separated in the Fourier domain. While noise in multiple beam interference adds somewhat to the computational load, Velghe et al. (2005a, 2005b, 2005c) demonstrated that multiple beams may be allowed to overlap in shearing configurations.

Wyant (1973) devised a double-frequency grating to implement a small amount of shear. This is accomplished by using a grating with two distinct frequencies recorded on it. The lower frequency is chosen so that the zeroth- and first-order beams are physically separated. The higher frequency part of the grating gives rise to two first-order beams that are sheared with respect to the other first-order beams. The multiple beam geometry and their overlap are indicated in Figure 4.46.

Shear in tangential and sagittal direction are generated simultaneously by inserting two gratings: the original one and the second one that is identical and orthogonal to the first one. Figure 4.47 presents the geometrical layout of the central

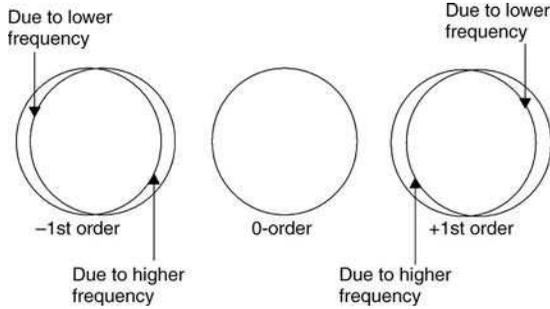


FIGURE 4.46. Typical geometry of the original and the diffracted apertures when a double frequency grating is used.

zeroth-order, flanked by eight first-order beams, two of each in shearing configuration. Figure 4.48 is a photo of a typical lateral shearing pattern of a converging wavefront when two identical double-frequency gratings are orthogonal to each other and in contact. It includes a uniform zero-order beam and four sheared interferograms. The horizontal first-order beams are sheared in horizontal direction, the vertical ones along the vertical direction. The two orthogonal double-frequency gratings may be recorded on the same photographic plate.

Rimmer and Wyant (1975) developed a technique to obtain variable shear with two crossed gratings of the same frequency. A small rotation of one grating gives rise to shear that changes with the magnitude of rotation angle and rotation arm. It is a first-order approximation to the rotationally shearing interferometer (see also Scholl, 1996; Strojnik and Paez, 2003). Figure 4.49 shows the distribution of beam apertures

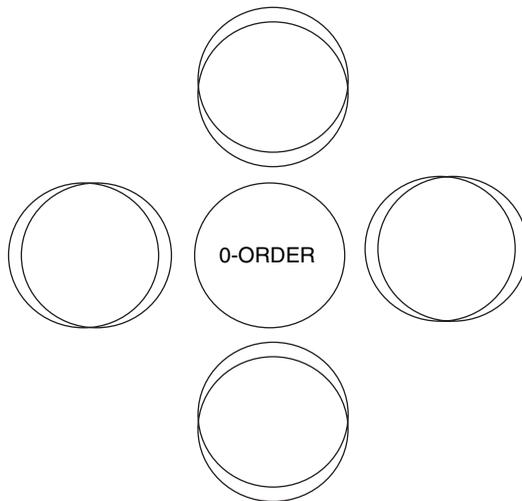


FIGURE 4.47. Multiple-beam geometry and beam overlap obtained with two double-frequency gratings. Lateral shear of the first-order beams is obtained simultaneously in the sagittal and tangential directions. The zero-order beam does not participate in the formation of the interferometric pattern.

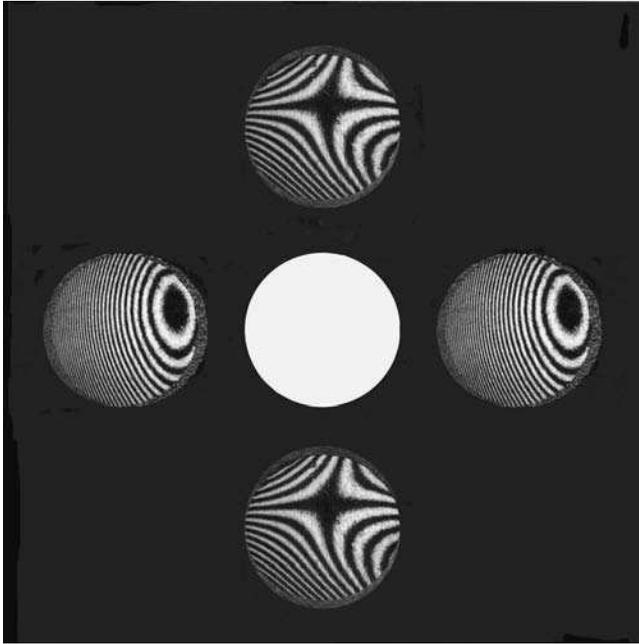


FIGURE 4.48. A photo of the typical set of interferograms obtained with a double-frequency grating. The zero-order beam presents a uniform field and two interferograms are repeated (from Wyant, 1973).

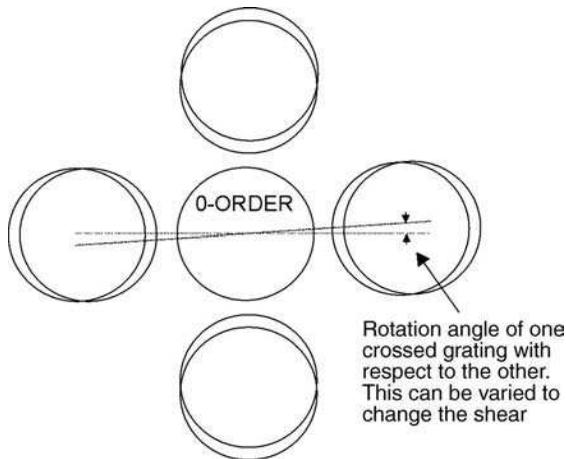


FIGURE 4.49. Beam distribution and beam overlap in the interference plane produced by two crossed gratings of the same frequency. They are in contact with each other to produce simultaneous lateral shear in tangential and sagittal directions. One grating may be rotated with respect to the other to vary the amount of shear (from Wyant, 1973).

that may be compared to those of Figure 4.47. This system is helpful for implementing variable shear. This method could be applied this method in a null test of an aspheric surface.

Joenathan et al. (1984) developed a variation of the double-frequency grating interferometer using two slightly displaced off-axis zone plates.

Schwider (1984) described another lateral shearing interferometer with a continuously varying amount of lateral shear. The shear is obtained with two diffraction gratings in a collimated beam, one after the other, forming an angle between them. The interference takes place between the beam with zero-order in the first grating and first-order in the second grating and the beam with first-order in the first grating and zero-order in the second grating.

Since 1990s, the dimensions of the lithographic masks and the intensity distributions over them have been pushed to the limit due to their increasingly small size and the absence of adequate testing equipment in the deep ultraviolet (UV) region at 193 nm. Lateral shearing interferometers with Ronchi gratings have been demonstrated to perform well even in this spectral region. Hegeman et al. (2001) describe an improved Ronchi setup with gratings modified to higher order beams. Fütterer et al. (2002) implemented the double grating interferometer working at 197 nm for phase shift mask measurement. A few years later, Fütterer and Schwider (2005) developed detailed error analysis to confirm the performance of this instrument.

At about the same time, a group in Japan applied this technology to evaluate the quality of the projection lens in extreme ultraviolet spectral region (EUV). Zhu et al. (2004) estimated the measurement precision of 0.1 nm RMS at 13.5 nm wavelength. These authors estimated a year later (Zhu et al., 2005) that equally high performance may be achieved using the lateral shearing interferometer with a grating as for the point diffraction interferometer. Liu et al. (2005) develop a novel technique to calibrate a double grating lateral shearing interferometer for EUV optics.

4.8.2. Lateral Shearing Interferometers Based on Polarization

A birefringent material gives rise to two orthogonally polarized beams when a beam of unpolarized light is incident on it. These orthogonally polarized beams do not form an interference pattern because they are mutually incoherent. The situation changes for the polarized light when the resulting orthogonally polarized beams are mutually coherent. They produce observable fringes upon interfering.

There exist many types of polarizing prisms that generate two orthogonally polarized beams from a plane polarized incident beam. The Wollaston prism is among the popular ones for use in the lateral shearing interferometers. (Interferometers incorporating such double-image prisms are treated in detail in Chapter 3 of the present volume). Murty and Shukla (1980), and Komissaruk and Mende (1981) described examples of this kind of interferometer using a liquid crystal wedge as a polarizing element.

Saxena (1979) described an interesting interferometer employing a Babinet compensator as indicated in Figure 4.50. This interferometer is similar to the one described in Chapter 3 on common-path interferometers, incorporating a Wollaston

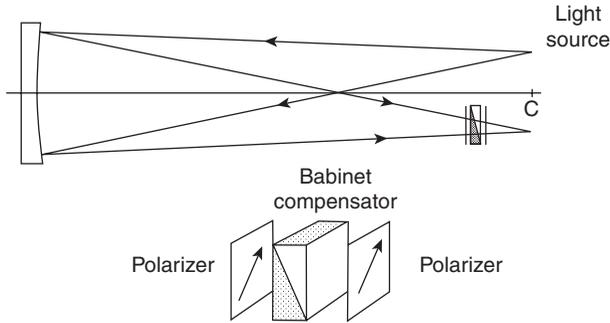


FIGURE 4.50. Interferometer incorporating a Babinet compensator to produce sheared beams.

compensator in double pass. Here, a Babinet compensator introduces the amount of lateral shear S , given by

$$S = 2R(N_e - N_o) \tan \alpha. \tag{4.34}$$

Here R is the radius of curvature of the concave surface under test, N_e and N_o are the extraordinary and ordinary refractive indices, respectively, and α is the compensator wedge angle. A typical Babinet compensator may be made from quartz ($N_e = 1.553305$ and $N_o = 1.544195$) and with wedge angle $\alpha = 7.5^\circ$.

If the Babinet compensator between two Polaroids is observed with a microscope, a pattern of straight, parallel, and equidistant fringes is observed, resembling a Ronchi ruling. The two important differences are that (a) the profile of the pattern is sinusoidal and (b) any two contiguous bright lines have relative phase difference of 180° . A slit separation in the Ronchi ruling, producing the same interference pattern, is equal to the separation of two lines with the same phase. For the compensator parameters listed above, the slit separation is equal to 110 lines per inch.

Saxena and Jayarajan (1981) and Saxena and Lancelot (1982) further improved this interferometer by incorporating two crossed Babinet compensators, doubling its sensitivity. The sensitivity of the interferometer to azimuthal variations in the compensator orientation is reduced, at the same time.

The layout of two crossed Babinet compensators is illustrated in Figure 4.51. Two compensators produce mutually orthogonal lateral shears. Thus, the condition for the fringe maximum (white fringe) is described by the following relations:

$$\frac{\partial W}{\partial x} S + \frac{\partial W}{\partial y} T = n\lambda \tag{4.35}$$

Here W is the wavefront deformation, including defocusing and tilts, and S and T are two orthogonal shears.

Two compensators may be placed in close contact to each other, but they may also be separated by a small distance d . In this case, the defocusing term will be

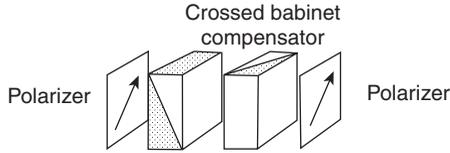


FIGURE 4.51. The layout of two crossed Babinet compensators. Two compensators produce mutually orthogonal lateral shears.

different for each one of two orthogonal lateral shears. Then, if the defocusing is referred to the point in the middle between the compensators, the expression for the fringe pattern becomes

$$\left(\frac{\partial W}{\partial x} + \frac{d}{2R^2}x\right)S + \left(\frac{\partial W}{\partial y} - \frac{d}{2R^2}y\right)T = n\lambda. \quad (4.36)$$

In this case, the fringes for spherical aberration become S-shaped, similar to those displayed in Figure 4.7.

A polarization phase shifter incorporating a birefringent plane parallel plate may be used to measure an aspherical surface. Weijers et al. (1998) present a simple calculation for ray propagation through uniaxial, birefringent element. Its inherent aberrations are derived. The method is used to record directly phase-shifted interferograms as needed for the interferogram evaluation and phase reconstruction.

Begbie et al. (2001) performed a rigorous analysis using Mueller matrices to evaluate the state of polarization necessary to eliminate fringe ambiguity.

4.9. VECTORIAL SHEARING INTERFEROMETER

The vectorial shearing interferometer is a lateral shearing interferometer where a displacement may be chosen along the most favorable direction. It generates a direct derivative of the wavefront, easily integrable with the method of line integration, developed and optimized by Paez and Strojnik (1997, 1998). It is applicable to testing and reconstruction of decentered wavefronts with very high fringe densities. See Paez and Strojnik (1999) and Paez and Scholl (2000).

Its implementation for the testing applications may be based on the Mach-Zehnder configuration by incorporating the displacement shearing system. In the original implementation of this device, it is composed of a pair of wedge prisms that modify the optical path difference and the tilt of the sheared wavefront with respect to that of the reference wavefront. The variable shear and tilt may be implemented along any direction, by choosing appropriate displacements Δx and Δy . The number of fringes and their orientation may be controlled with the shear direction and its magnitude. A gradient of the phase function is obtained in any direction with the knowledge of the displacements in the x and y directions.

4.9.1. Shearing Interferometry

The interest in detecting fainter sources and smaller details has, in the recent years, fostered the development of large diameter optical systems, including those with segmented and diluted primaries. These configurations usually require the fabrication and figuring of the relatively large diameter off-axis optical elements. Testing is very difficult because the common established interferometric techniques, in general, require the existence of a well-formed reference component. Thus, the optical designs that are to be fabricated require the availability of appropriate test components and supplies.

In the shearing interferometers, the amplitude of the wavefront under test is split into two parts: one half is left unchanged, while a key parameter is modified in the other half. The most significant feature of the shearing interferometers is that they do not require a reference wavefront, because the wavefront under test is compared with itself. Those portions of the wavefronts that maintain the same polarization are recombined in the detection plane, where they form an interference pattern. There exist three basic types of shearing interferometers: the lateral, the radial, and the rotational. The wavefront inversion interferometer has been shown to be a special case of the rotational-shearing one, with a shearing angle of 180° .

We have recently developed an interferogram interpretation method that allows phase reconstruction even from very high-density fringe patterns. The input to the algorithm is the vectorial derivative of the phase function. Therefore, the shearing interferometry is a particularly useful way of recording the input intensity data, avoiding the noise increase, exhibited when taking derivative of noisy data. In a (displacement) shearing interferometer, the phase derivative along the shear direction is the measured quantity. One of its salient characteristics is that the derivative is taken along a single direction. We developed a new type of interferometer, where a derivative is taken along an arbitrary direction, under the operator's control. The known incremental displacements in the two perpendicular directions, Δx and Δy , allow the determination of a two-dimensional derivative of the phase function along the direction. The wavefront its replica displaced along vector Δr , and a representative interferogram, are illustrated in Figure 4.52, for the coma.

The possibility of taking a derivative along a specific direction is most beneficial when testing asymmetrical components, where the fringe density in one direction is extremely high, forming the moiré patterns where no phase has been recorded. This is particularly useful when employing the method of phase reconstruction upon the line-integration of its gradient.

With the possibility of choosing the path of the line integral, this may always be selected along a favorable direction. In fact, upon employing a line integral, a two-dimensional problem has been changed into that of a single dimension. As will be shown later on, one of the outstanding features of the vectorial shearing interferometer is that the fringe density is controllable with the amount of shear.

A general wavefront may be expressed in coordinate system erected in the exit pupil. Upon applying the symmetry considerations relevant to the testing of symmetrical optical systems, the aberration polynomial in polar form may be written as a

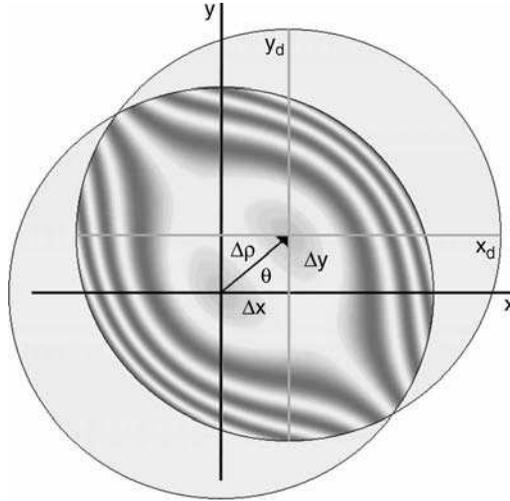


FIGURE 4.52. The original, the modified wavefronts, and the interferogram in the vectorial shearing interferometer with the displacement or shear performed along an arbitrary direction.

simple polynomial, expanded up to the fourth power in the radial coordinate. On the right side, we identify the terms in the wavefront aberration function with the primary aberrations.

$$\begin{aligned}
 W(\rho, \phi) = F & & \text{Piston term} \\
 + E\rho(\sin \phi + \cos \phi) & \text{ Tilts } (x, y) \\
 + D\rho^2 & \text{ Defocus} \\
 + C\rho^2(1 + 2 \cos^2 \phi) & \text{ Astigmatism} \\
 + B\rho^3 \cos \phi & \text{ Coma} \\
 + A\rho^4 & \text{ Spherical aberration}
 \end{aligned} \tag{4.37}$$

The vectorial shearing interferometer is a generalized displacement (position) or linear shearing interferometer along an arbitrary direction, with respect to the coordinates defined by the detector pixel arrangement. The magnitude and direction of shear is kept under the operator control. The tilt introduced in one arm of the interferometer upon the prism rotation is similarly maintained under the operator control. However, the two are mutually dependent in the initial embodiment.

4.9.2. Directional Shearing Interferometer

In a lateral shearing interferometer, the wavefront displacement or shear is performed along a single direction, referred to as x or y . The region of the wavefront overlap is reduced upon increasing the shearing distance. We modified a Mach-Zehnder arrangement to implement our vectorial shearing interferometer. We cannot affirm

though this is either the only way or the best method of doing so. Paez et al. (2000) designed and built this interferometer. García-Torales et al. (2001) analyzed it later on. Most recently, García-Torales et al. (2006) and Ramirez et al. (2007) have applied it to the testing of asphericity of spheres.

As may be seen in Figure 4.52, the region of interference is the common area of overlap of the two wavefronts. This figure additionally illustrates the displacement vector $\Delta\rho$, the net directional displacement of the sheared wavefront with respect to the position of the original wavefront, as defined by its center. The displacement vector $\Delta\rho$ subtends an angle θ relative to the x axis. We can select the size of the area of overlap of the two wavefronts by the proper choice for the incremental displacements Δx and Δy in the shearing vector.

Figure 4.53 shows the schematic diagram of the experimental setup for the implementation of the vectorial shearing interferometer in the Mach–Zehnder configuration, including the shearing and the compensator systems. This setup might be used in transmission mode to test a lens, placed inside the beam just before the first beam splitter. The shearing system, consisting of two thin prisms, displaces the wavefront according to the angles of rotation of each prism (Strojnik et al., 2000, García-Torales et al., 2002). In the detector plane, the two beams are detected forming an interference pattern.

A compensator system, in the form of a pair of fixed prisms, identical to those that constitute the rotator system may be introduced into the reference arm. Due to the use of the compensator system in the reference arm, the interferometer is white-light compensated.

The new position of the center of the displaced wavefront is a vector sum of individual displacements produced by the independent rotation of each prism. The wavefront orientation remains the same. When both prisms are fabricated from the same material and they rotate in the air, the rotation of each individual prism generates the same angular deviation. In the image plane, the beam angular deviation is transformed into distance. The total deviation of the ray traversing

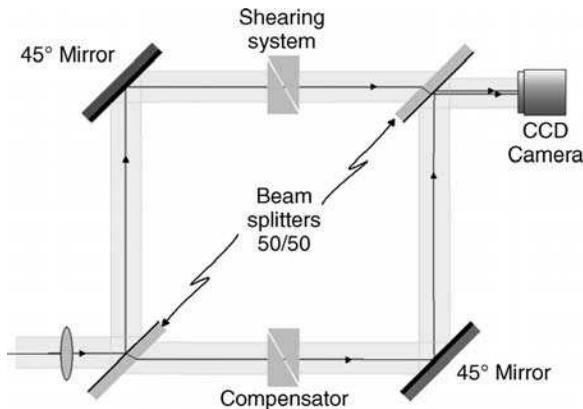


FIGURE 4.53. The experimental setup with the vectorial shearing interferometer for testing a positive lens in transmission.

the shearing system is a function of the individual angular deviations (Sandoval et al., 2001).

Additionally, the shearing system introduces the tilt in the wavefront with respect to the y - z plane, the x - z plane, or both. Once the prism materials and optical parameters are chosen, the tilt may be controlled with the amount and direction of rotation of the prisms. With the tilt change, the shear is also modified, as both are controlled by the action of the beam rotator. When only the prism separation is changed without modifying the relative angle between them, lateral shear in an arbitrary direction may be changed in magnitude without introducing any tilt. This configuration has been found most helpful in flame characterization studies (Paez and Strojnik, 2006).

A sagittal shift might correspond to the displacement in the x direction and the tangential one to that in the y direction. We may write the optical path difference, OPD, using the difference of the wavefront aberration polynomials,

$$\Delta\text{OPD} = W(\rho + \Delta\rho, \theta + \Delta\theta) - W(\rho, \theta). \quad (4.38)$$

The shearing radius vector $\Delta\rho$ and the shearing angle θ are found upon the change of coordinates

$$\rho = (\Delta x^2 + \Delta y^2)^{1/2}, \quad \theta = \arctan\left(\frac{\Delta y}{\Delta x}\right). \quad (4.39)$$

Next, we show the interferometric patterns simulated for a vectorial shearing interferometer and compare them with the familiar patterns obtained in the traditional interferometers.

4.9.3. Simulated Interferometric Patterns

We use Eqs. (4.37) and (4.38) to simulate interferometric patterns that arise from the presence of the primary aberrations. In each case, 4 waves of defocus are added. The resultant simulated patterns are shown for 2.5 waves of spherical aberration in Figure 4.54, for 3 waves of astigmatism in Figure 4.55, and for 7 waves of coma in Figure 4.56. Figure 4.57 depicts a mixture of aberrations. It is included to facilitate the comparison with the experimental results.

The simulation results of Figures 4.54 through 4.57 are presented in an array of three columns by four rows. The intensity distributions are given inside a normalized square with corners at ± 1 , whose center coincides with the central point on the original wavefront.

The regions where the (finite) pupils do not overlap are not indicated, including only the essential information in the figure. The columns show the simulated interferograms for the following positions, from left to right: (a) inside focus (with -4 waves of defocus), (b) in the focus, and (c) outside focus (with $+4$ waves of defocus). The first row presents the interferometric pattern simulated for a Mach-Zehnder interferometer to help us recollect its familiar form. The second row illustrates the patterns obtained in a linear shearing interferometer along the

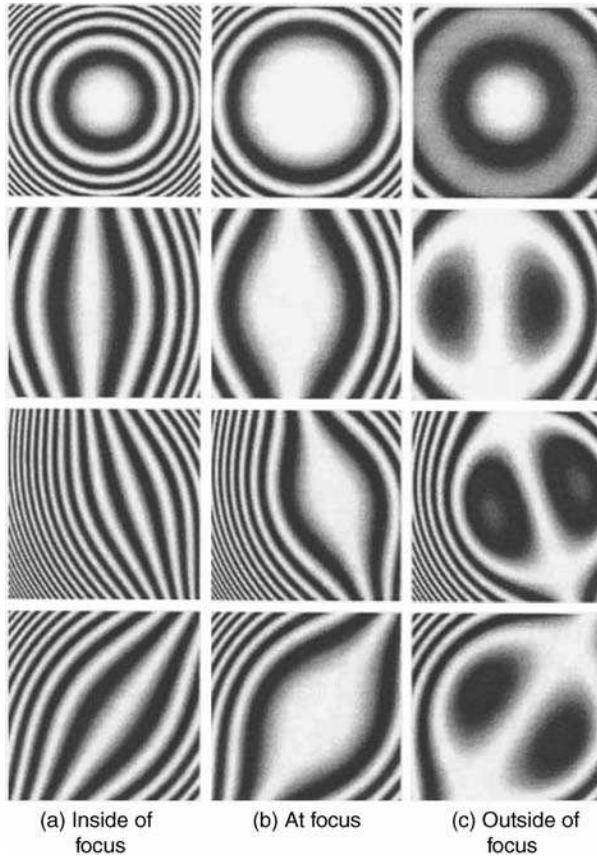


FIGURE 4.54. Simulated intensity patterns of the spherical aberration (2.5 waves) and defocus ((a)-4 waves and (c) +4 waves) are given for the Mach-Zehnder interferometer; the linear shearing interferometer; twice in a vectorial shearing interferometer; first with a large radius vector pointing to the third quadrant and second with a small radius vector pointing to the fourth quadrant.

x direction to facilitate comparison with the familiar results. For this case, the displacement radius vector is $\rho = 0.3$ and the displacement angle is $\theta = \pi$ rad.

The last two rows show the results for the general vectorial displacements of $\rho = 0.632$ and the displacement angle $\theta = 0.322$ rad (the third row), and $\rho = 0.361$ and the displacement angle $\theta = 5.695$ rad (the fourth row). The center of the sheared wavefront is translated by the magnitude of the displacement radius vector ρ along the direction θ with respect to the origin at the center of the square pupil. We discuss each figure in more detail.

Figure 4.54 shows the simulated interferometric patterns when ± 4 waves of defocus and no defocus are added to 2.5 waves of spherical aberration. The second row presenting the intensity patterns of lateral shearing interferometer is in agreement with the patterns presented at the beginning of this chapter. It confirms the general reliability of our simulations. The fringe patterns outside focus appear as

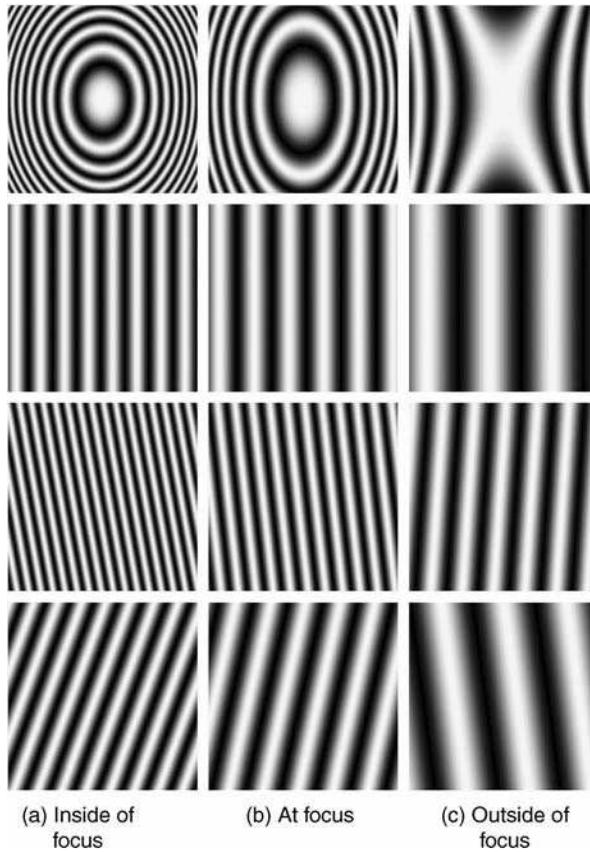


FIGURE 4.55. Simulated intensity patterns of the astigmatism (3 waves) and defocus ((a)-4 waves and (c) +4 waves) are given for the Mach-Zehnder interferometer; the linear shearing interferometer; twice in a vectorial shearing interferometer; first with a larger radius vector pointing to the third quadrant and second with a smaller radius vector pointing to the fourth quadrant.

those generated by the coma in a Mach-Zehnder interferometer. The fringes due to the spherical aberration detected outside the focus appear similar to those in a linear shearing interferometer, except that they are tilted with respect to those generated upon a one-dimensional shear and their origin is displaced.

The coma pattern of the shearing setup is flattened out as the amount of defocus is decreased in all shearing cases. Additionally, the fringe density increases with the shearing radius vector. This feature may be employed to vary the instrument sensitivity by adjusting the prism rotation angles. We can define an axis of symmetry to be vertical in the linearly sheared interferogram in the second row. Thus, the fringe patterns, or the line of symmetry, are rotated in the case of a vectorial shearing interferometer with respect to this line of symmetry. The amount of rotation is related to the shearing angle θ . The position of the fringe center also remains under the operator's control: the fringes are displaced in the direction opposite to that of the

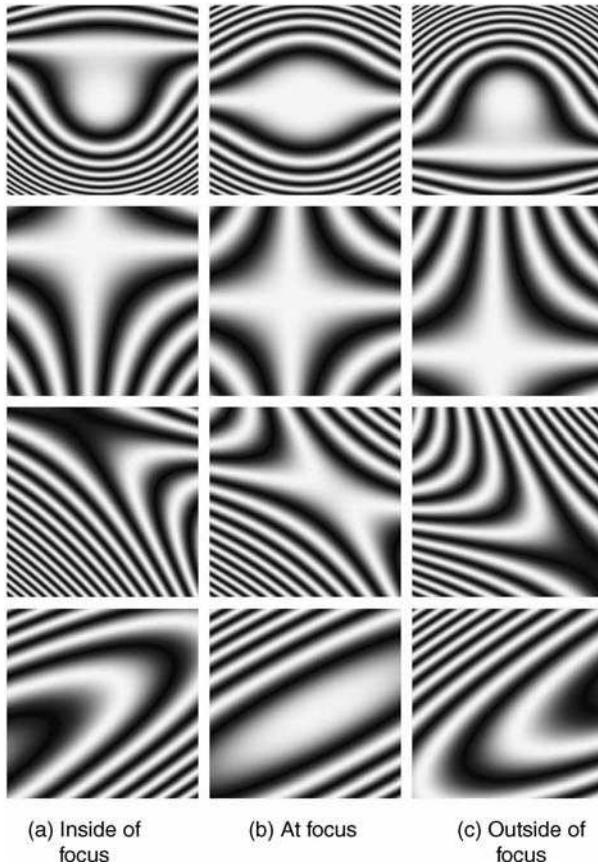


FIGURE 4.56. Simulated intensity patterns of the coma (7 waves) and defocus (± 4 waves) are given for the Mach-Zehnder interferometer; the linear shearing interferometer; twice in a vectorial shearing interferometer; first with a larger radius vector pointing to the third quadrant and second with a smaller radius vector pointing to the fourth quadrant.

displacement vector $\Delta\rho$. The angle θ remains the same when the increments along the x and the y directions maintain the same ratio, even when the net displacement $\Delta\rho$ changes.

Figure 4.55 displays the simulated interferometric patterns when ± 4 waves of defocus and no defocus are added to 3 waves of astigmatism. In the linear shearing interferometer, the presence of astigmatism is recognized by the appearance of the straight fringes normal to the shear direction.

The fringes obtained upon the application of a general vectorial shear are inclined with respect to those of the linear shear. Their inclination, though, depends on the amount of defocus. The fringe direction appears to be reversed for a certain outside focus position, where the defocus and the astigmatism compensate each other. The number of fringes decreases with defocus. The number of fringes increases with the displacement radius vector for any defocus condition.

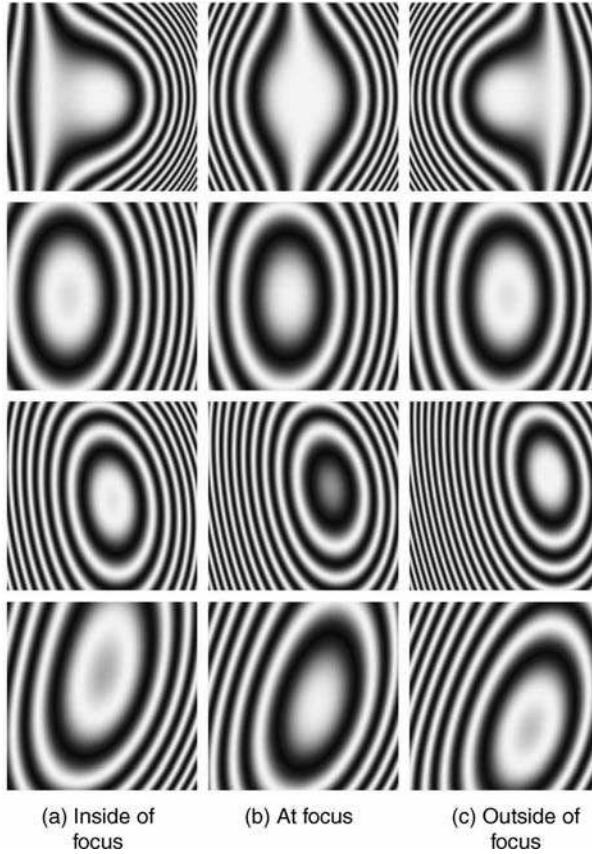


FIGURE 4.57. Simulated interferometric patterns produced by the sum of the spherical aberration, the astigmatism, the coma, and defocus (± 4 waves) are given for the Mach-Zehnder interferometer; the linear shearing interferometer, twice in a vectorial shearing interferometer; first with a larger radius vector pointing to the third quadrant and second with a smaller radius vector pointing to the fourth quadrant.

Figure 4.56 depicts the simulated interferometric patterns when ± 4 waves of defocus and no defocus are added to 7 waves of coma. The shape of the patterns seen in the second row is astigmatism-like, corresponding to those recorded in the linear shearing interferometer. Defocusing is observed to move the astigmatism-like pattern up and down, perpendicular the shear direction.

For the two generalized shearing cases in the last two rows, we observe the deformed astigmatism-like patterns corresponding to the two familiar astigmatism patterns: the X-shaped one and the O-shaped one. The great advantage of the variety of available patterns is that one can choose an intensity distribution that is more easily amenable to the phase reconstruction. When moving from inside the focus to the outside position, the patterns appear to move along two diagonals.

Figure 4.57 illustrates the simulated interferometric patterns produced by the sum of the spherical aberration, the astigmatism, the coma, and defocus (± 4 waves). All the interferometric patterns generated with the vectorial shearing interferometer exhibit the fringe movement out of the center of the reference beam. The fringe pattern tends to rotate in the direction opposite to that of the displacement vector. The fringe density increases with the magnitude of the shearing vector.

4.9.4. Experimental Results

Here, we illustrate the feasibility of the identification of the type and amount of aberration. Figures 4.53 and 4.58 illustrate two of a number of available experimental setups that we used to record the interferograms with the vectorial shearing interferometer in the modified Mach–Zehnder configuration. The first arrangement is used for testing a wavefront in transmission, and the second one in reflection. Even in the laboratory environment, it is possible to achieve very small, differential, displacement as well as a small, but finite one.

A photograph of a pair of beam director prisms is demonstrated in Figure 4.59. The thin prisms are mounted in their rotary holders to control the orientation of each prism and their separation. A photograph of the essential parts of the compact vectorial shearing interferometer is included in Figure 4.60. It presents the wavefront director in one arm and the optical path compensator in the other one. The source could be incident from the left beam, horizontally. Then fringes are formed on the right (see Figs. 4.53 and 4.58). A high resolution CCD camera on the right records images.

An example of the experimental setup with the vectorial shearing interferometer used for testing a positive lens upon transmission has been seen in Figure 4.53. The experimentally obtained interferometric patterns of the lens with the (a) small and

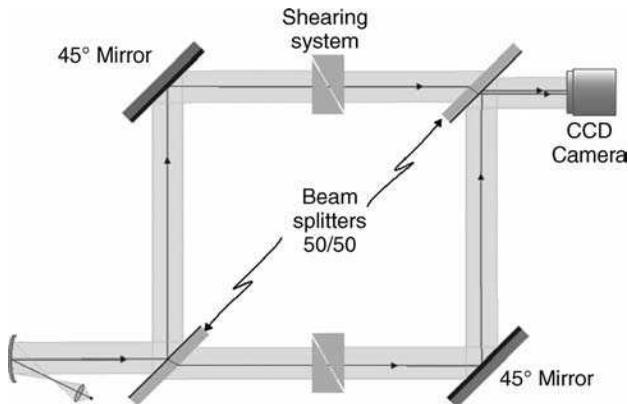


FIGURE 4.58. The experimental setup with the vectorial shearing interferometer for testing a (parabolic) mirror in reflection.



FIGURE 4.59. A photograph of a pair of beam director prisms. The thin prisms are mounted in their rotary holders to control the orientation of each prism and their separation.

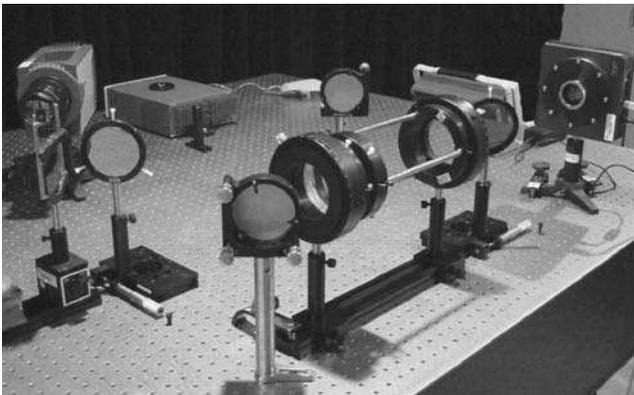


FIGURE 4.60. A photograph of the essential parts of the compact vectorial shearing interferometer illustrates the wavefront director in one arm. The source beam could be incident from the left side, horizontally. Then fringes are formed on the right (and above) of the beam divider (see also Figures 4.53 and 4.58). A high resolution CCD camera records the fringe pattern.

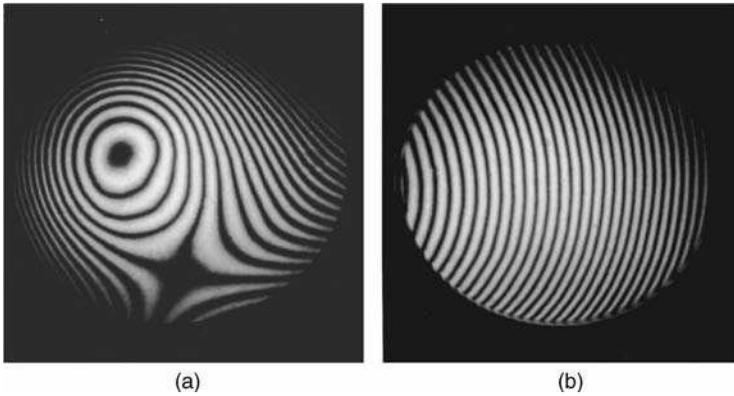


FIGURE 4.61. The experimentally recorded interferometric patterns of the lens for (a) the infinitesimally small and (b) the small, but finite shearing distance obtained with the vectorial shearing interferometer used in transmission.

(b) large shearing distance, respectively, are shown in Figure 4.61. In spite of the use of the transmissive elements, this configuration is not affected by the ghost reflections. The prisms in both the shearing and compensator systems deviate the undesirable reflections out of the beam path of the wavefront under test.

Figure 4.58 has presented the experimental setup with the vectorial shearing interferometer used for testing a parabolic mirror in reflection. The experimentally obtained intensity patterns of the parabolic mirror for the (a) small and (b) large shearing distance, respectively, are depicted in Figure 4.62.

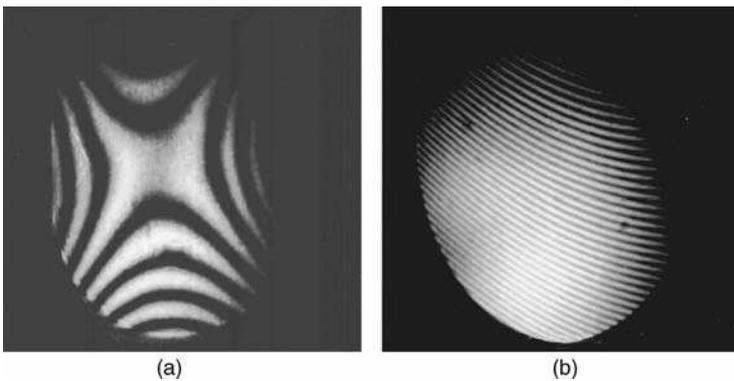


FIGURE 4.62. The experimentally obtained interferometric patterns of the lens for (a) the infinitesimally small and (b) small, but finite shearing distance obtained with the vectorial shearing interferometer used in reflection.

4.9.5. Similarities and Differences With Other Interferometers

We conceptualized, predicted, and demonstrated the performance of a new type of shearing interferometer, with the significant capabilities of controlling the amount and the direction of the displacement shear. The vectorial shearing interferometer may be based on the Mach–Zehnder configuration by incorporating the displacement shearing system. The shearing system is composed of a pair of wedge prisms that modify with a high degree of precision the optical path difference and the tilt between the sheared wavefront and the reference wavefront.

The use of small apex-angle prisms decreases the number of ghost beams within the interferometer and the amount of straight light incident on the detector plane. While it is believed that some limited version of the generalized linear shearing interferometer may be implemented with a single prism, we prefer the implementation with two prisms, due to the enhanced control and the range of wavefront positioning. Due to the knowledge of the displacement increments along two perpendicular axes, the measured quantity is indeed a directional derivative of the phase function. This allows complete phase recovery in two dimensions. The variable shear and tilt, that may be implemented along any other direction, have the particular beneficial characteristics that the number of fringes and their orientation may be controlled with the shear direction and its magnitude.

Xu et al. (2005) proposed an implementation of a compact vectorial shearing interferometer that allows phase shift by moving prisms along the beam normal. This represents a significant simplification of its layout. With only four thin prisms, shearing is accomplished along an arbitrary direction, and phase stepping is achieved without incorporating a movable mirror on a piezoelectric mount. The range of displacement is also large.

All the interferometric patterns generated with the vectorial shearing interferometer exhibit the fringe movement out of the center of the reference beam. We may define the angle of rotation of the fringe pattern to be the angle between the normal to the fringes when the shearing angle θ is zero, and when the shearing angle θ has a finite value. Its algebraic sign is opposite to that of the shearing angle θ . Fringe density increases as the magnitude of the shearing vector is incremented.

Figure 4.63 illustrates that the fringe orientation is under the control of operator. This may be accomplished upon changing the orientation of both prisms simultaneously, without changing their separation or relative orientation. If part (a) is the reference pattern, the fringes are displaced (b) up, (c) down, (d) left, and (e) right using a pair of prism with an index of refraction n , of 1.517, and the apex-angle $\varepsilon = 5^\circ$.

Figure 4.64 presents results of a study between the fringe density and the relative prism orientation. Fringe density increases as the difference angle between prism orientation increases from (a) 0.5° , (b) 0.7° , (c) 1.3° , (d) 1.5° , and (e) 1.6° . Prisms with index of refraction $N = 1.517$ and apex-angle $\varepsilon = 5^\circ$ are used.

The versatility and adaptability of the vectorial shearing interferometer with variable shear and tilt are best appreciated upon examining the experimental results depicted in Figure 4.65. The interferometric patterns are obtained in transmission with a commercial interferometer (WYCO), on the left, and the vectorial shearing

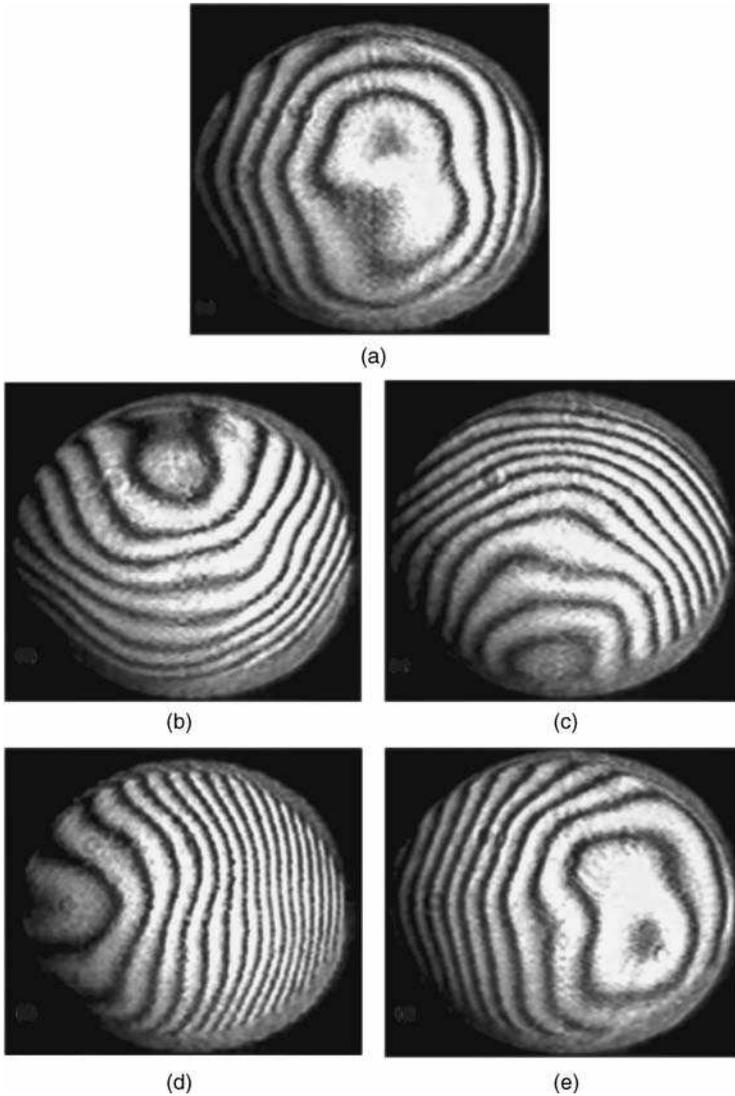


FIGURE 4.63. Fringe orientation is selectable by changing the orientation of both prisms simultaneously, without changing their separation or relative orientation. Using part (a) as the reference pattern, the fringes are displaced up (b), down (c), left (d), and right (e). We are using a pair of prisms with an index of refraction N , of 1.517, and the apex-angle $\varepsilon = 5^\circ$.

interferometer, on the right part (a) display, well-corrected convex lens with a focal length $f = 30$ cm and diameter $D = 5$ cm. In part (b), the difference of the interferograms of a low-quality lens with focal length $f = 25$ cm and diameter $D = 7$ cm favors the one on the right. In both cases, the interferogram on the right recorded with the vectorial shearing interferometer is obtained with favorable fringe density.

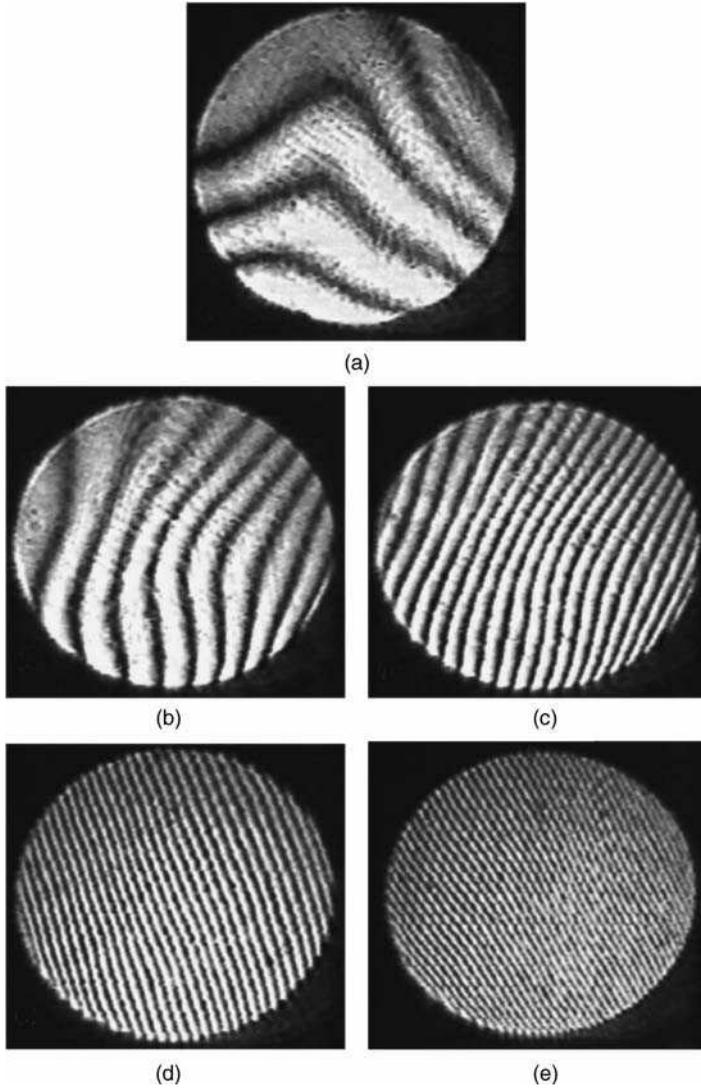
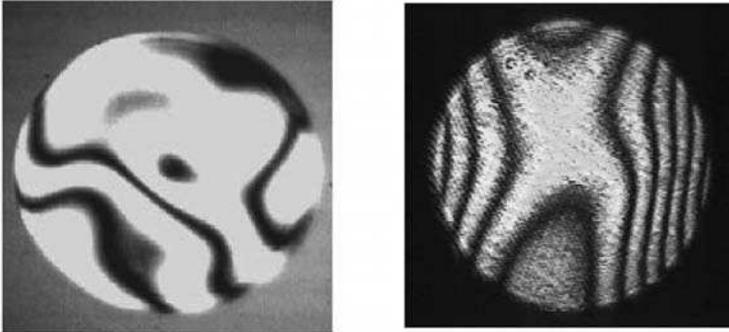
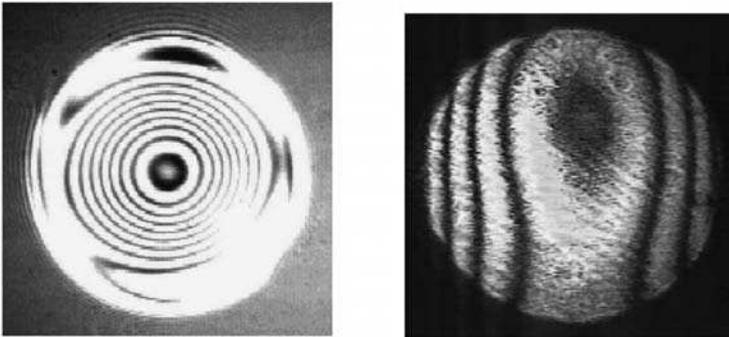


FIGURE 4.64. Fringe density increases as the difference angle between prism orientation increases from (a) 0.5° , (b) 0.7° , (c) 1.3° , (d) 1.5° , and (e) 1.6° . Prisms with index of refraction $N = 1.517$ and apex-angle $\varepsilon = 5^\circ$ are used.



(a) A well-corrected convex lens with a focal length $f = 30$ cm and diameter $D = 5$ cm.



(b) A low-quality lens with focal length $f = 25$ cm and diameter $D = 7$ cm.

FIGURE 4.65. The interferometric patterns obtained in transmission with a commercial interferometer (WYCO) on the left, and the vectorial shearing interferometer on the right. (a) Interferograms of a well-corrected convex lens with a focal length $f = 30$ cm and diameter $D = 5$ cm. (b) Interferograms of a low-quality lens with focal length $f = 25$ cm and diameter $D = 7$ cm.

REFERENCES

- Albertazzi A., Jr. and A V. Fantin, "Digital Complex Holography Using a Shearing Interferometer: Principles and Early Results," *Proc. SPIE.*, **4777**, 57–68 (2002).
- Begbie M. L., W. Sibbett, and M. J. Padgett, "Mueller Matrix Error Correction for a Fringe-Free Interferometry System," *Appl. Opt.*, **40**, 3205–3210 (2001).
- Chang K.-H., A.-C. Hsu, and J. L. Chern, "Comment on Phase-Shifting Shearing Interferometer," *Opt. Lett.*, **27**, 509–510 (2002).
- Dickey F. M. and T. M. Harder, "Shearing Plate Optical Alignment," *Opt. Eng.*, **17**, 295–298 (1978).
- Dubra A., C. Paterson, and C. Dainty, "Study of the Tear Topography Dynamics Using a Lateral Shearing Interferometer," *Opt. Exp.*, **12**, 6278–6288 (2004).
- Dubra A., C. Paterson, and C. Dainty, "Double Lateral Shearing Interferometer for the Quantitative Measurement of Tear Film Topography," *Appl. Opt.*, **44**, 1191–1199 (2005).

- Dutton D., A. Cornejo, and M. Latta, "Semiautomatic Method for Interpreting Shearing Interferograms," *Appl. Opt.*, **7**, 125–132 (1968).
- Fütterer G., M. Lano, N. Lindlein, and J. Schwider, "Lateral Shearing Interferometer for Phase-Shift Mask Measurement at 193 nm," *Proc. SPIE*, **4691**, 541–551 (2002).
- Fütterer G. and J. Schwider, "Uncertainty Analysis for Phase Measurement on Psm with a 193 Nm Common-path Shearing Interferometer," *Proc. SPIE*, **5858**, 321–333 (2005).
- García-Torales G., G. Paez, and M. Strojnik, "Simulations and Experimental Results with a Vectorial Shearing Interferometer," *Opt. Eng.*, **40** (5) 767–773 (2001).
- García-Torales G., M. Strojnik, and G. Paez, "Risley Prisms to Control Wavefront Tilt and Displacement in a Vectorial Shearing Interferometer," *Appl. Opt.*, **41**, 1380–1384 (2002).
- García-Torales G., G. Paez., M. Strojnik, J. Villa, J. L. Flores, and A. González-Alvarez, "Experimental Intensity Patterns-Obtained from a 2D Shearing Interferometer with Adaptable Sensitivity," *Opt. Comm.*, **257**, 16–26 (2006).
- Gorshkov V. A. and V. G. Lysenko, "Study of Aspherical Wavefronts on a Lateral Shearing Interferometer," *Sov. J. Opt. Technol.*, **47**, 689 (1980).
- Griffin D. W., "Phase-Shifting Shearing Interferometer," *Opt. Lett.*, **26**, 140–141 (2001).
- Grindel M. W., "Testing Collimation Using Shearing Interferometry," *Proc. SPIE*, **680**, 44–46 (1986).
- Hariharan P., "Simple Laser Interferometer with Variable Shear and Tilt," *Appl. Opt.*, **14**, 1056–1056 (1975).
- Hariharan P., "Lateral and Radial Shearing Interferometers: A Comparison," *Appl. Opt.*, **27**, 3594–3596 (1975).
- Hariharan P. and D. Sen, "Cyclic Shearing Interferometer," *J. Sci. Instrum.*, **37**, 374 (1960).
- Hegeman P., X. Christmann, M. Visser, and J. Braat, "Experimental Study of a Shearing Interferometer Concept for At-Wavelength Characterization of Extreme-Ultraviolet Optics," *Appl. Opt.*, **40**, 4526–4533 (2001).
- Joenathan C., R. K. Mohanty, and R. S. Sirohi, "Lateral Shear Interferometry with Holo Shear Lens," *Opt. Commun.*, **52**, 153–156 (1984).
- Kanjilal A. K., P. N. Puntambekar, and D. Sen, "Compact Cyclic Shearing Interferometer: Part One," *Opt. Laser Technol.*, **16**, 261–264 (1984).
- Kanjilal A. K. and P. N. Puntambekar, "Compact Cyclic Shearing Interferometer: Part Two," *Opt. Laser Technol.*, **16**, 311–314 (1984).
- Kasana R. S. and K. J. Rosenbruch, "Determination of the Refractive Index of a Lens Using the Murty Shearing Interferometer," *Appl. Opt.*, **22**, 3526–3531 (1983a).
- Kasana R. S. and K. J. Rosenbruch, "The Use of a Plane Parallel Glass Plate for Determining the Lens Parameters," *Opt. Commun.*, **46**, 69–73 (1983b).
- Kelsal D., Thesis, *University of London*, 1959, *Proc. Phys. Soc.*, **73**, 465 (1959).
- Komissaruk V. A. and N. P. Mende, "A Polarization Interferometer with Simplified Double-Refracting Prisms," *Opt. Laser Technol.*, **13**, 151–154 (1981).
- Korwan D., "Lateral Shearing Interferogram Analysis," *Proc. SPIE*, **429**, 194–198 (1983).
- Lechna M. M., T. J. Licznarski, and H. T. Kasprzak, "Interferometry for In-vivo Testing of Artificial Tears on the Surface of the Cornea," *Proc. SPIE*, **3820**, 386–389 (1999).
- Lee Y. W., H. M. Cho, D. J. Shin, and I. W. Lee, "Concollimated Bidirectional Shearing Interferometer for Measuring a Long Radius of Curvature," *Appl. Opt.*, **36**, 5317–5320 (1997).

- Lee H.-H., J.-H. You and S.-H. Park, "Phase-Shifting Lateral Shearing Interferometer with Two Pairs of Wedge Plates," *Opt. Lett.*, **28**, 2243–2245 (2003).
- Leibbrandt G. W. R., G. Harbers, and P. J. Kunst, "Wavefront Analysis with High Accuracy by Use of a Double-Grating Lateral Shearing Interferometer," *Appl. Opt.*, **35**, 6151–6161 (1996).
- Liang P., J. Ding, Z. Jin, C.-S. Guo, and H.-T. Wang, "Two-Dimensional Wavefront Reconstruction from Lateral Shearing Interferograms," *Opt. Exp.*, **14**, 625–634 (2006a).
- Liang P., J. Ding, J. Zhou, and Hui-Tian Wang, "Wavefront Reconstruction for Lateral Shearing Interferometers with Big Shear," *Proc. SPIE*, **6027**, 60270Q (2006b).
- Liu Z., M. Okada, K. Sugisaki, M. Ishii, Y. Zhu, K. Ohtaki, J. Saito, A. Suzuki, M. Hasegawa, C. Ouchi, S. Kato, M. Niibe, and K. Murakami, "Double-Grating Lateral Shearing Interferometer for EUV Optics at-Wavelength Measurement," *Proc. SPIE*, **5752**, 663–672 (2005).
- Lonouvel L. and F. Lonouvel, "Etude des Faisceau Convergentes (A study on convergent beams)," *Rev. Opt. Theor. Instrum.*, **17**, 350 (1938).
- Malacara D., *Testing of Optical Surfaces*, Ph. D. Thesis, Institute of Optics, University of Rochester, New York, 1965a.
- Malacara D., "Two Lenses to Collimate Red Laser Light," *Appl. Opt.*, **4**, 1652–1654 (1965b).
- Malacara D. and M. Mende, "Lateral Shearing Interferometry of Wavefronts Having Rotational Symmetry," *Opt. Acta.*, **15**, 59–63 (1968).
- Malacara D. A. Cornejo, and M. V. R. K. Murty, "A Shearing Interferometer for Convergent or Divergent Beams," *Bol. Inst. Tonantzintla*, **1**, 233–239 (1975).
- Malacara D. and M. Servin, "Frequency Analysis of Lateral Shearing Interferometers," *Proc. SPIE*, **2544**, 218–222 (1995).
- Malacara-Hernández D., G. Paez, D. Malacara-Doblado, and J. García-Márquez, "Wavefront Retrieval from Lateral Shearing Interferograms with Fourier Techniques," *Proc. SPIE*, **2744**, 290–294 (1999).
- Mehta D. S., M. D. Singh, P. Singh, M. S. Faridi, S. Mirza, and Ch. Shakher, "Distance Measurement with Extended Range Using Lateral Shearing Interferometry and Fourier Transform Fringe Analysis," *Opt. Eng.*, **44**, 063602–10 (2005a).
- Mehta D. S., S. K. Dubey, M. M. Hossain, and Ch. Shakher, "Simple Multifrequency and Phase-Shifting Fringe-Projection System Based on Two-Wavelength Lateral Shearing Interferometry for Three-Dimensional Profilometry," *Appl. Opt.*, **44**, 7515–7521 (2005b).
- Murty M. V. R. K., "The Use of a Single Plane Parallel Plate as a Lateral Shearing Interferometer with a Visible Gas Laser Source," *Appl. Opt.*, **3**, 531–534 (1964a).
- Murty M. V. R. K., "Some Modifications of the Jamin Interferometer Useful in Optical Testing," *Appl. Opt.*, **3**, 535–538 (1964b).
- Murty M. V. R. K., "Fabrication of Fixed Shear Cube Type Shearing Interferometer," *Bull. Opt. Soc. India*, **3**, 55 (1969).
- Murty M. V. R. K., "A Compact Lateral Shearing Interferometer Based on the Michelson Interferometer," *Appl. Opt.*, **9**, 1146–1148 (1970).
- Murty M. V. R. K. and D. Malacara, "Some Applications of the Gas Laser as a Source of Light for the Testing of Optical Systems," Proceedings of the Conference on Photographic and Spectroscopic Optics, Tokyo and Kyoto, 1964, *Jpn. J. Appl. Phys.*, **4**, Suppl. 1, 106–111 (1965).

- Murty M. V. R. K. and R. P. Shukla, "Liquid Crystal Wedge as a Polarizing Element and its Use in Shearing Interferometry," *Opt. Eng.*, **19**, 113–115 (1980).
- Murty M. V. R. K. and R. P. Shukla, "Parallel Plate Interferometer for the Precise Measurement of Radius of Curvature of a Test Plate & Focal Length of a Lens Systems," *Ind. J. Pure Appl. Phys.*, **21**, 587 (1983).
- Nyssonen D. and J. M. Jerke, "Lens Testing with a Simple Wavefront Shearing Interferometer," *Appl. Opt.*, **12**, 2061–2082 (1973).
- Philipp H., H. Fuchs, E. Winklhofer, and Georg Pretzler, "Flame Diagnostics by Light Sheet Imaging and by Shearing Interferometry," *Opt. Eng.*, **32**, 1025–1032 (1993).
- Paez G. and M. Strojnik, "Fringe Analysis and Phase Reconstruction from Modulated Intensity Patterns," *Opt. Lett.*, **22**(22), 1669–1971 (1997).
- Paez G. and M. Strojnik, "Convergent, Recursive Phase Reconstruction from Noisy, Modulated Intensity Patterns Using Synthetic Interferograms," *Opt. Lett.*, **23**, 406–408 (1998).
- Paez G. and M. S. Scholl, "Phase-Shifted Interferometry Without Phase Unwrapping: Reconstruction of a Decentered Wavefront," *J. Opt. Soc. Am. A*, **16**, 475–480 (1999).
- Paez G. and M. Strojnik, "Phase Reconstruction from Underdetected Intensity Pattern(s)," *J. Opt. Soc. Am. A*, **16**, 46–52 (2000).
- Paez G. and M. S. Scholl, "Features of the Vectorial Shearing Interferometer," *Proc. SPIE*, **4231**, 328–339 (2000).
- Paez G., M. Strojnik, and G. García -Torales, "Vectorial Shearing Interferometer," *Appl. Opt.*, **39**(28), 5172–5178 (2000).
- Paez G. and M. Strojnik, "Applications of Vectorial Shearing Interferometer," Fringe 91, *The 4th international Workshop on Automatic Processing of Fringe Patterns*, W. Osten, Ed., Bremen, Germany, September 17–19, 2001, p.97.
- Paez G. and M. Strojnik, "Flame Characterization with Vectorial Shearing Interferometer," *Unusual Optical Systems II*, Proc. SPIE 6719, International Society for Optical Engineering, Bellingham, Washington, USA, 2006.
- Pfeil A. V., B. Messerschmidt, V. Blmel, U. Possner, and T. Possner, "Making Fast Cylindrical Gradient-Index Lenses Diffraction Limited by Using a Wavefront-Correction Element," *Appl. Opt.*, **37**, 5211–5215 (1998).
- Rimmer M. P., *Method for Evaluating Lateral Shearing Interferograms*, Itek Corp. Internal Report No. 725802-1, 1972.
- Rimmer M. P. and J. C. Wyant, "Evaluation of Large Aberrations Using a Lateral-Shear Interferometer Having Variable Shear," *Appl. Opt.*, **14**, 142–150 (1975).
- Ronchi V., *Ann. Sc. Norm. Super. Pisa*, **15**(1923).
- Ronchi V., "Forty Years of History of a Grating Interferometer," *Appl. Opt.*, **3**, 437–451 (1964).
- Sandoval J., G. Paez, and M. Strojnik, "Opto-Mechanical Design of a Prism Rotator," *Proc. SPIE*, **4486**, 170–180 (2001).
- Saunders J. B., "Measurement of Wavefronts Without a Reference Standard I: The Wavefront Shearing Interferometer," *J. Res. Nat. Bur. Stand.*, **65B**, 239 (1961).
- Saunders J. B., "Wavefront Shearing Prism Interferometer," *J. Res. Nat. Bur. Stand.*, **68C**, 155 (1964a).
- Saunders J. B., "Interferometer Test of the 26-Inch Refractor at Leander McCormick Observatory," *Astron. J.*, **69**, 449 (1964b).

- Saunders J. B., "Some Applications of the Wavefront Shearing Interferometer," Proceeding of the Conference on Photographic and Spectroscopic Optics, Tokyo and Kyoto, 1964, *Jpn. J. Appl. Phys.*, **4**, Suppl. 1, 99 (1965).
- Saunders J. B., "A Simple Interferometric Method for Workshop Testing of Optics," *Appl. Opt.*, **9**, 1623–1629 (1970).
- Saunders J. B. and J. Bruning, New Interferometric Test and its Applications to the 84-Inch Reflecting Telescope at Kitt Peak National Observatory, *Astron. J.*, **73**, 415 (1968).
- Saxena A. K., "Quantitative Test for Concave Aspheric Surfaces Using a Babinet Compensator," *Appl. Opt.*, **18**, 2897–2895 (1979).
- Saxena A. K. and A. P. Jayarajan, "Testing Concave Aspheric Surfaces: Use of Two Crossed Babinet Compensators," *Appl. Opt.*, **20**, 724–725 (1981).
- Saxena A. K. and J. P. Lancelot, "Theoretical Fringe Profile with Crossed Babinet Compensators in Testing Concave Aspheric Surfaces," *Appl. Opt.*, **21**, 4030–4032 (1982).
- Scholl M. S., "Ray Trace Through a Corner Cube Retro reflector with Complex Reflection Coefficients," *J. Opt. Soc. Am. A*, **12** (7), 1589–1592 (1995).
- Scholl M. S., "Measured Spatial Properties of the CW Nd-YAG Laser Beam," *Appl. Opt.*, **19** (21), 3655–3659 (1980).
- Scholl M. S., "Signal Detection by an Extra-Solar-System Planet Detected by a Rotating Rotationally-Shearing Interferometer," *J. Opt. Soc. Am. A*, **13**, 1584–1592 (1996).
- Schreiber H., "Measuring Wavefront Tilt Using Shearing Interferometry," *Proc. SPIE*, **5965**, 596–607 (2005).
- Schwider J., "Superposition Fringe Shear Interferometry," *Appl. Opt.*, **19**, 4233–4240 (1980).
- Schwider J., "Continuous Lateral Shearing Interferometer," *Appl. Opt.*, **23**, 4403–4409 (1984).
- Shen W., M.-W. Cheng, D.-S. Wan, "Zernike Polynomial Fitting of Lateral Shearing Interferometry," *Opt. Eng.*, **36**, 905–913 (1997).
- Sirohi R. S. and M. P. Kothiyal, "Double Wedge Plate Shearing Interferometer for Collimations Test," *Appl. Opt.*, **26**, 4054–4055 (1987a).
- Sirohi R. S. and M. P. Kothiyal, "On Collimation of a Laser Beam," *Proc. SPIE*, **813**, 205 (1987b).
- Song J. B., Y. W. Lee, I. W. Lee, and J. Lee, "Simple Phase Shifting Method in a Wedge Lateral Shearing Interferometer," *Appl. Opt.*, **43**, 3989–3992 (2004).
- Strojnik M., G. García-Torales, and G. Paez, "Fabricación de Prismas de Cuña para su Aplicación en el Sistema Director del Frente de Onda del Interferómetro de Desplazamiento Vectorial" (Fabrication of Thin Prisms for the Wavefront Director in the Rotational Shearing Interferometer), Informe Técnico 12, Centro de Investigaciones en Optica, León, México, 2000.
- Strojnik M. and G. Paez, "Comparison of Linear and Rotationally Shearing Interferometric Layouts for Extra Solar Planet Detection from Space," *Appl. Opt.*, **42**, 5897–5905 (2003).
- Suhara H., "Interferometric Measurement of the Refractive-Index Distribution in Plastic Lenses by Use of Computed Tomography," *Appl. Opt.*, **41**, 5317–5325 (2002).
- Tanner L. H., "Some Laser Interferometers for Use in Fluid Mechanics," *K. Sci. Instrum.*, **42**, 834 (1965).
- Tarvin J. A., R. D. Sigler, and G. E. Busch., "Wavefront Shearing Interferometer for Cryogenic Laser-Fusion Targets," *Appl. Opt.*, **18**, 2971 (1979).
- Weijers A. L., H. Van Brug, and H. J. Frankena, "Polarization Phase Stepping with a Savart Element," *Appl. Opt.*, **37**, 5150–5155 (1998).

- Velghe S., J. Primot, N. Guineau, R. Haidar, M. Cohen, and B. Wattellier, "Accurate and Highly Resolving Quadric-Wave Lateral Shearing Interferometer, from Visible to IR," *Proc. SPIE*, **5776**, 134–143 (2005).
- Velghe S., J. Primot, N. Guineau, M. Cohen, and B. Wattellier, "Visible and Infrared Wavefront Metrology by Quadric-Wave Lateral Shearing Interferometry," *Proc. SPIE*, **5965**, 596–512 (2005).
- Velghe S., J. Primot, N. Guineau, M. Cohen, and B. Wattellier, "Wavefront Reconstruction from Multidirectional Phase Derivatives Generated by Multilateral Shearing Interferometers," *Opt. Lett.*, **30**, 245–247 (2005).
- Venkata B. and D. P. Juyal, "A 10-mm CO₂ Laser Interferometer Using Shearing Technique," *J. Optics (India)*, **16**, 31 (1987).
- Waddell P., M. Stickland, S. McKay, and L. S. Mair, "Checking the Symmetry of Stretchable Plastic Membrane Concave Mirrors Using a Lateral Shearing Interferometer," *Proc. SPIE*, **2248**, 141 (1994).
- Wattellier B., C. Sauteret, J.-C. Chanteloup, and A. Migus, "Beam-Focus Shaping by Use of Programmable Phase-Only Filters: Application to an Ultra long Focal Line," *Opt. Lett.*, **27**, 213–215 (2002).
- Wyant J. C., "Double Frequency Grating Lateral Shear Interferometer," *Appl. Opt.*, **12**, 2057 (1973).
- Xu X., Q. Lu, and Z. Liu, "Optical Phase Retrieval by Lateral Shear Interferometer Based on Wavelet Transform," *Proc. SPIE*, **4929**, 532–537 (2002).
- Xu R., H. Liu, Z. Luan, and L. Liu, "A Phase-Shifting Vectorial-Shearing Interferometer with Wedge Plate Phase-Shifter," *J. Opt. A: Pure Appl. Opt.*, **7**, 617-23 (2005).
- Zhu Y., K. Sugisaki, C. Ouchi, M. Hasegawa, M. Niibe, A. Suzuki, and K. Murakami, "Lateral Shearing Interferometer for EUVL: Theoretical Analysis and Experiment," *Proc. SPIE*, **5374**, 824–832 (2004).
- Zhu Y., K. Sugisaki, M. Okada, K. Ohtaki, Z. Liu, M. Ishii, J. Kawakami, J. Saito, K. Murakami, C. Ouchi, M. Hasegawa, S. Kato, T. Hasegawa, A. Suzuki, and M. Niibe, "Experimental Comparison of Absolute PDI and Lateral Shearing Interferometer," *Proc. SPIE*, **5752**, 1192–1199 (2005).
- Xu R., H. Liu, Z. Luan, and L. Liu, "A Phase-Shifting Vectorial-Shearing Interferometer with Wedge Plate Phase-Shifter," *J. Opt. A: Pure Appl. Opt.*, **7**, 617 23 (2005).

5

Radial, Rotational, and Reversal Shear Interferometer

D. Malacara

5.1. INTRODUCTION

The two interfering wavefronts in a two beam interferometer can have several possible relative orientations and sizes with many different configurations. As pointed out by Steel (1983) and by Walmsley and Malacara (1995), the first order properties of these two beam interferometers can be completely described by

- (a) the light source position and the location, orientation, and size of the two images of this light source as seen from the observation plane;
- (b) the image plane position and the location, orientation, and size of the two images of this observation screen as seen from the light source position; and
- (c) the optical path difference between the two optical paths.

These parameters completely define the lateral shear, tilt, and lead. The shear is the relative lateral displacement of the two interfering wavefronts and shift is their longitudinal separation. The tilt is the lateral separation of the two images of the light source and the lead is their longitudinal separation. A fifth independent parameter is the optical path difference (OPD), which is a function of the traversed refractive indices. These parameters are illustrated in Figure 5.1. The main properties of all shearing interferometers can be analyzed by these parameters.

Although the most popular shearing interferometer is the lateral shearing instrument, other types are equally useful. In this chapter we examine radial, rotational, and reversal shear interferometers, whose basic wavefront operations are illustrated in Figure 5.2. Many review papers (Murty, 1967; Briers, 1972; Fouéré and Malacara, 1975) and books (Bryngdahl, 1965; Steel, 1966; Baird and Hanes, 1967) have very good general descriptions of them. The radial shear interferometer produces two interfering wavefronts with identical deformations, but one of the wavefronts is contracted or expanded with respect to the other. The rotational shear interferometer

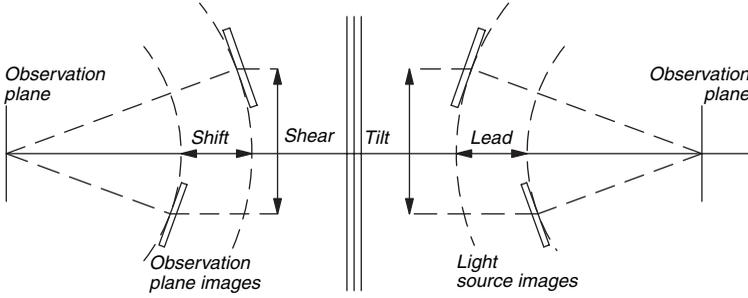


FIGURE 5.1. First order parameters for any interferometer.

produces two identical wavefronts with one of them rotated with respect to the other. The reversal shear interferometer produces two wavefronts, in which deformations on one wavefront are symmetrical with respect to those on the other wavefront, with a diameter as axis of symmetry.

For the analysis of these shearing operations, we can assume a completely general wavefront function given by

$$W(\rho, \theta) = \sum_{n=0}^k \sum_{l=0}^n \rho^n (a_{nl} \cos l\theta + b_{nl} \sin l\theta) \tag{5.1}$$

where $(n - l)$ is given (n and l have the same parity) and generally $l \leq n$.

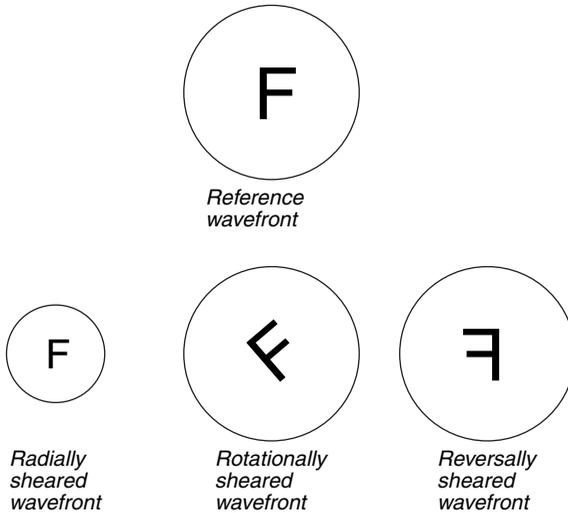


FIGURE 5.2. Three shear possibilities between two wavefronts.

5.2. RADIAL SHEAR INTERFEROMETERS

Radial shear interferometers perform the basic operations illustrated in Figure 5.3. One of the wavefronts is contracted or expanded with respect to the other. The interferometer may be thought of as an optical system producing two images of an object at A, with different magnifications, at the location A'. These two images must coincide at A' if a defocusing term is to be avoided.

As described in Chapter 2 on Twyman–Green interferometers, the interferogram to be analyzed has to be an image of the exit pupil of the system under test, especially when the wavefront deviations from the spherical shape are large. If the wavefront is almost spherical, this condition is not necessary. In the case of the radial shear interferometer, both wavefronts are deformed. Thus, both wavefronts in the interferogram must be images of the pupil. When the shear is very large, only the smaller wavefront needs to be an image of the exit pupil of the system. Fortunately, this condition may frequently be satisfied since, as described by Steel (1984), all radial shear interferometers have a second pair of conjugates, B and B', as shown in Figure 5.3, with the same shear ratio as the images A and A'. Then, ideally, the exit pupil of the system under test must be located at B and the interferogram should be analyzed at B'.

Steel (1984) also showed that by reversing the direction of the light, a radial shear interferometer with the same shear is also obtained. Thus, any system may be used in four ways to produce the same shear.

These types of interferometers directly represent the wavefront deformations when the shear is large, but even if the shear is small, they are simpler to interpret than lateral shear interferometers, and unlike them, not only in one, they provide information in all directions. Then, only one interferogram is needed. Radial and lateral shear interferometers have been compared in detail by Hariharan (1988).

Instead of the typical approach of using a long-wavelength or two-wavelength interferometry, the lower sensitivity of radial shear interferometers makes them ideal for testing wave fronts with a high degree of asphericity because the number of fringes is smaller.

When the surface under test has a central hole, the shear has to be small, otherwise the smaller of the two wavefronts would be over the hole of the larger wavefront. This

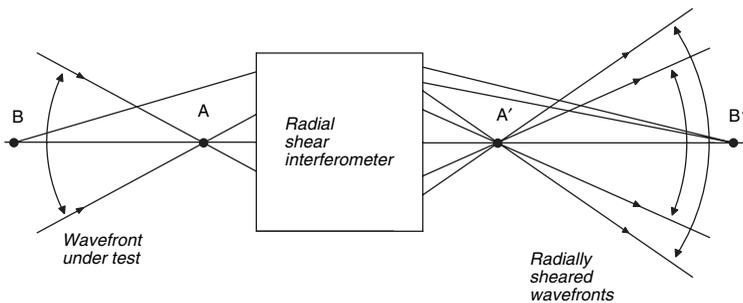


FIGURE 5.3. Schematic block for a radial shear interferometer.

small shear reduces the sensitivity, but this disadvantage may be compensated for by the use of phase shifting techniques (Hariharan et al., 1986).

We can assume that one of the wavefronts is expanded so that the ratio of the diameters of the two wavefronts is R , defined by

$$R = \frac{\rho}{\rho_e} \quad (5.2)$$

where ρ is the ratio of the radial distance of a point in the interference pattern to the maximum radius of the nonexpanded (or smaller) of the two wavefronts. Similarly, ρ_e is the ratio of the same radial distance to the maximum radius of the expanded (or larger) wavefront.

Then, the wavefront deformation for the larger (noncontracted or expanded) wavefront is represented by Eq. (5.1) and the wavefront deformation for the smaller (nonexpanded or contracted) is

$$W(\rho, \theta) = \sum_{n=0}^k \sum_{l=0}^n R^n \rho^n (a_{nl} \cos l\theta + b_{nl} \sin l\theta) \quad (5.3)$$

and hence the interference pattern will be given by the OPD,

$$\text{OPD}(\rho, \theta) = W(\rho, \theta) - W(R\rho, \theta) = \sum_{n=0}^k \sum_{l=0}^n \rho^n (1 - R^n) (a_{nl} \cos l\theta + b_{nl} \sin l\theta) \quad (5.4)$$

The sensitivity of a radial shear interferometer relative to that of a Twyman–Green interferometer is given by

$$\sigma = \frac{d\text{OPD}(\rho, \theta)/d\rho}{dW(\rho, \theta)/d\rho} = \frac{\sum_{n=0}^k \sum_{l=0}^n n\rho^{n-1} (1 - R^n) (a_{nl} \cos l\theta + b_{nl} \sin l\theta)}{\sum_{n=0}^k \sum_{l=0}^n n\rho^{n-1} (a_{nl} \cos l\theta + b_{nl} \sin l\theta)} \quad (5.5)$$

but if only one aberration (n, l) is present, the relative sensitivity can be expressed as

$$\sigma_{n,l} = 1 - R^n \quad (5.6)$$

and is plotted in Figure 5.4. for some aberrations. We can see that a moderate effective radial shear R equal to 0.5 gives a very high relative sensitivity. When testing aspherical wavefront with pure primary spherical aberration, the radial shear interferometer provides a reduced sensitivity and an increased dynamic range equivalent to the use of a larger wavelength.

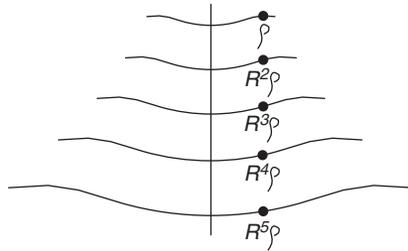


FIGURE 5.4. Radially sheared wavefronts used in iterative calculation of the wavefront.

If the expansion of one of the wavefronts becomes infinite ($\rho_e \rightarrow \infty$ and $\rho \rightarrow 0$), we have a kind of radial shear interferometer that is said to have exploded shear. This topic is treated in the chapter on common-path interferometers.

5.2.1. Wavefront Evaluation from Radial Shear Interferograms

Radial shear interferograms are basically identical to the Twyman–Green interferograms studied in Chapter 2, especially if only pure primary aberrations are present. Of special interest are wavefronts with rotational symmetry because they are obtained during the testing of aspherical rotationally symmetric surfaces, for example, an astronomical mirror. The procedure for computing the wavefront from a radial shear interferogram under these conditions has been given by Malacara (1974) and by Honda et al. (1987). However, in general, the wavefront is not rotationally symmetric and even some lateral shear may be present. Then, the wavefront may be computed by one of three possible methods.

1. The first procedure consists in calculating the a_{nl} , b_{nl} coefficients in Eq. (5.4) by a least-squares fitting of the OPD (ρ, θ) to the radial shear interferogram, by assuming that the wavefront is smooth enough, so that it can be represented by a polynomial.

2. Another method, described by Kohler and Gamiz (1986) is by successive iterations. In the first iteration, the reference wavefront (the larger wavefront) is assumed to be perfectly flat and the interferogram is sampled and fitted to a polynomial. The result of this evaluation is then used in the second iteration with a better estimation of the reference wavefront. This procedure produces a very accurate result, limited only by the sample spacing, reading errors, and the quality of the wavefront fitting.

3. The third method, described by Kohno et al. (2000) and later in a similar manner by Li et al (2002) is by an iterative process. From Eq. (5.4) we can obtain

$$\begin{aligned}
 \text{OPD}(\rho, \theta) &= W(\rho, \theta) - W(R\rho, \theta) \\
 \text{OPD}(R\rho, \theta) &= W(R\rho, \theta) - W(R^2\rho, \theta) \\
 &\dots\dots\dots \\
 \text{OPD}(R^n\rho, \theta) &= W(R^n\rho, \theta) - W(R^{n-1}\rho, \theta)
 \end{aligned}
 \tag{5.7}$$

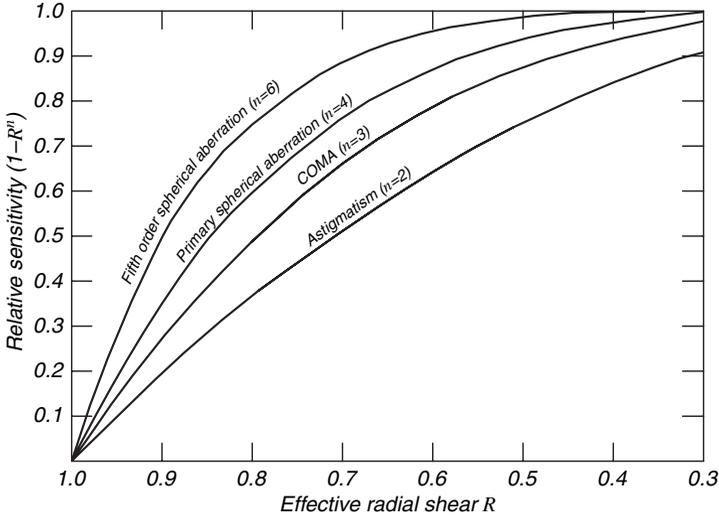


FIGURE 5.5. Relative sensitivity σ versus radial shear R .

If we now add all these terms we may write

$$W(\rho, \theta) = \sum_{i=1}^n \text{OPD}(R^i \rho, \theta) - W(R^n \rho, \theta) \quad (5.8)$$

If the number of iterations is sufficiently large, the last term becomes a small flat wavefront whose deformations can be considered zero (see Fig. 5.5). Thus, we finally obtain

$$W(\rho, \theta) = \sum_{i=1}^n \text{OPD}(R^i \rho, \theta) \quad (5.9)$$

5.2.2. Single-Pass Radial Shear Interferometers

In this kind of interferometer, the light goes through the instrument only once and two radially sheared interfering wavefronts are produced as illustrated in Figure 5.3. If the interferometer is illuminated by a small, circular extended source, the degree of coherence g_{12} between any two points on the wavefront is given (Hariharan and Sen 1961a; Murty 1964b) by

$$g_{12} = \frac{2J_1[(2\pi/\lambda)\alpha d]}{(2\pi/\lambda)\alpha d} \quad (5.10)$$

where 2α is the angular diameter of the source as seen from the wavefront under consideration and d is the distance between the two points on the same wavefront.

For a radial shear interferometer, the distance d between two points on the wavefront that interfere at a point on the interference pattern can be shown to be given by

$$d = \rho' - \rho'' = \rho'(1 - R). \quad (5.11)$$

Therefore, for those two points we obtain (Hariharan and Sen, 1961a)

$$g_{12} = \frac{2J_1[(2\pi/\lambda)(1 - R)\rho'\alpha]}{(2\pi/\lambda)(1 - R)\rho'\alpha}. \quad (5.12)$$

Since the visibility of the pattern is directly proportional to g_{12} , a fringe pattern similar to that in Figure 5.6 will be obtained. The first minimum on the visibility is obtained when the argument x of $J_1(x)$ is equal to 1.22π . Hence, for a good visibility of the whole pattern, the circular light source must have an angular semiconductor α , as seen from the wavefront (entrance pupil of the interferometer), smaller than a certain value given by

$$\alpha \leq \frac{1.22\lambda}{(1 - R)D} \quad (5.13)$$

where D is the entrance pupil diameter.

A radial shear interferometer is said to be compensated for the imperfect monochromaticity of the light (wavelength bandwidth) when the optical paths for the two

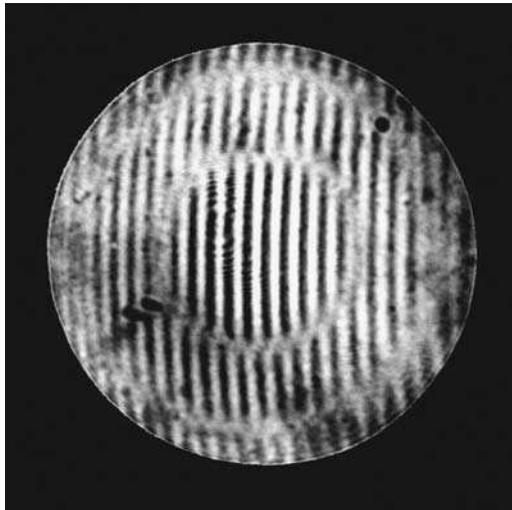


FIGURE 5.6. Fringe visibility changes in a radial shear interferometer with a large source. (From Hariharan and Sen, 1961a.)

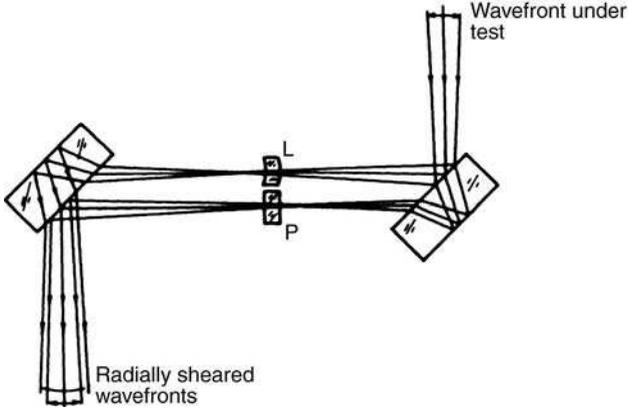


FIGURE 5.7. Brown's radial shear interferometer.

interfering beams involve the same glass and air paths. Interferometers of this kind will now be discussed in this section.

Radial shear interferometers were first considered by Brown (1959), who described the instrument shown in Figure 5.7 (Brown, 1962). Basically, this is a Jamin interferometer, but it uses convergent light and has a small meniscus lens in one of the beams. A compensating parallel plate is placed in the other beam.

Another of the first radial shear interferometers, designed by Hariharan and Sen (1961b), is shown in Figure 5.8. It consists of a plane parallel beam splitter plate P and two plane mirrors M_1 and M_2 . The radial shear is produced by two lenses L_1 and L_2 , which are placed in such a manner that their foci are at the face of the beam

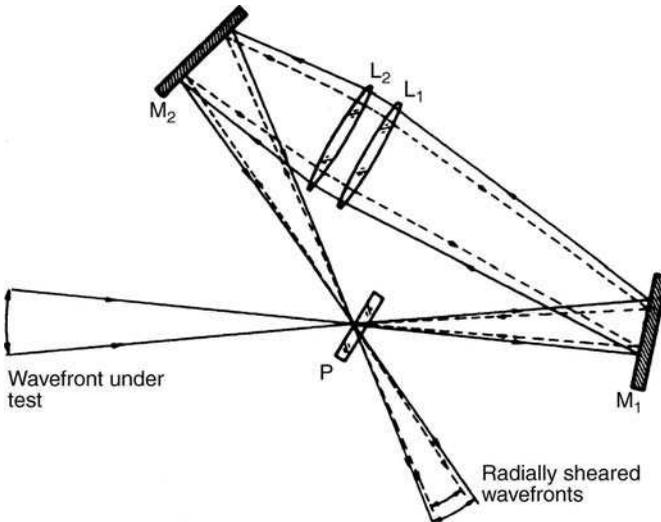


FIGURE 5.8. Hariharan and Sen's radial shear interferometer.

splitter. The radial shear is produced when the two lenses have different focal lengths f_1 and f_2 ($f_2 > f_1$) and is given by

$$R = \frac{f_1}{f_2}. \tag{5.14}$$

The two lenses could be replaced by a single lens, but two lenses serve to facilitate the elimination of aberrations produced by the lenses.

The visibility is equal to 1 only if the two beams have the same irradiance; but since the two beams are differently expanded, this is possible only if the reflectance \mathcal{R} and the transmittance \mathcal{F} of the beam splitter satisfy the equation

$$\frac{\mathcal{R}}{\mathcal{F}} = \left(\frac{f_2}{f_1}\right)^2. \tag{5.15}$$

Hariharan and Sen (1962) successfully used this interferometer to test microscope objectives.

Murty (1964a) suggested several arrangements to produce radial shear. One of them is based on the Mach-Zehnder interferometer, using telescopic systems S_1 and S_2 on each of the arms as shown in Figure 5.9. The effective radial shear F is given by $1/M^2$, where M is the magnification of each of the telescopes.

Another system is a cyclical interferometer (shown in Fig. 5.10) that resembles the interferometer designed by Hariharan and Sen. Here, however, the light entering the interferometer should be collimated.

A very practical and interesting interferometer, also described by Murty (1964a), is based on the contraction and expansion of the numerical aperture by a

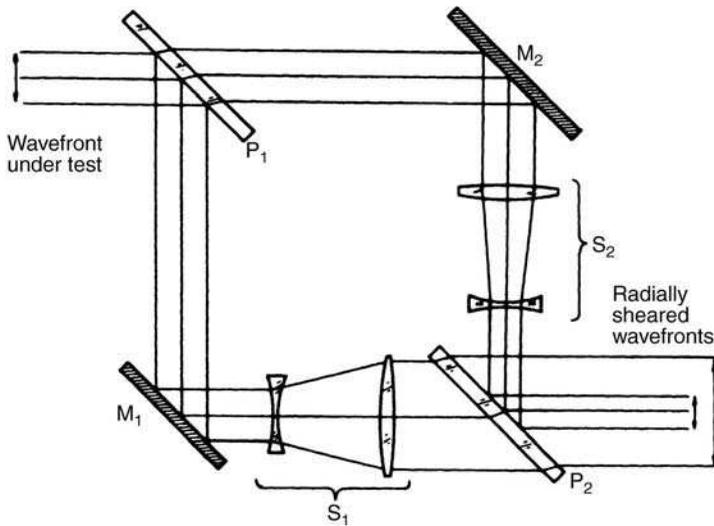


FIGURE 5.9. Telescopic systems in a Mach-Zehnder interferometer to produce radial shear.

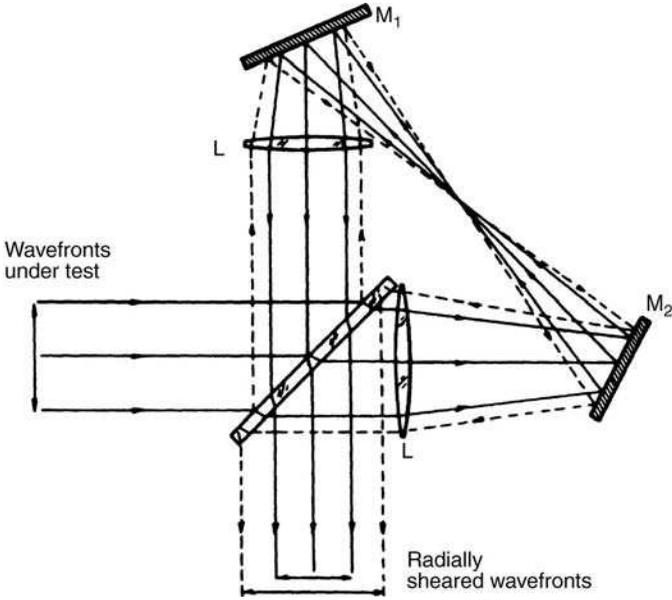


FIGURE 5.10. Cyclic radial shear interferometer for collimated beams.

hemispherical lens as illustrated in Figure 5.11. Using this principle and the basic cyclic configuration, Murty designed the interferometers shown in Figures 5.12 and 5.13. The hemispherical cavity in the second interferometer may be emptied or filled with oil, in order to obtain the desired radial shear. Some very unconventional radial shear interferometers with discontinuous stepping wave-fronts have been described by Bryngdahl (1970, 1971).

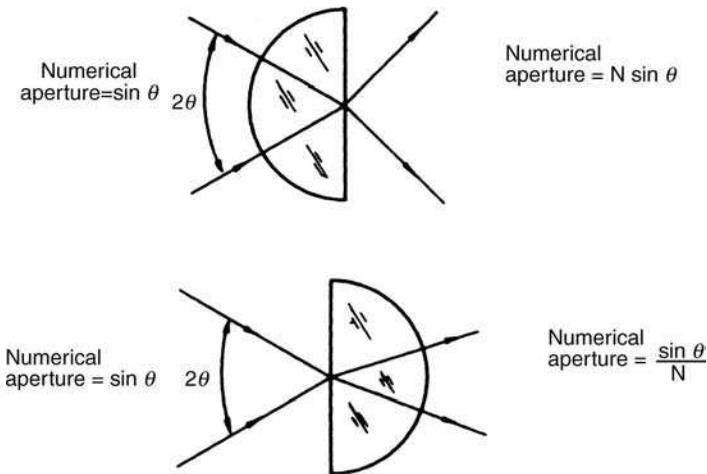


FIGURE 5.11. Contraction and expansion of the numerical aperture by a hemispherical lens.

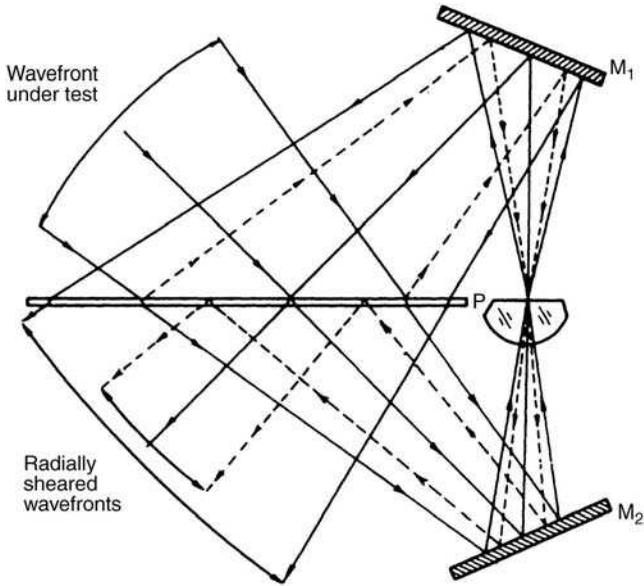


FIGURE 5.12. Murty's cyclic radial shear interferometer.

5.2.3. Double-Pass Radial Shear Interferometers

The schematic block for a double-pass radial shear interferometer is shown in Figure 5.14. It must be recalled that the single-pass interferometer in Figure 5.3 produces two interference patterns, one formed by the light passing through the interferometer and another by the light that is reflected back from the instrument in order to preserve

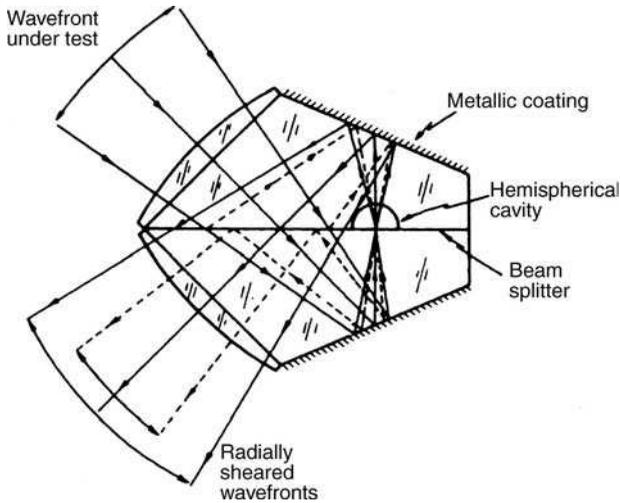


FIGURE 5.13. Murty's solid radial shear interferometer.

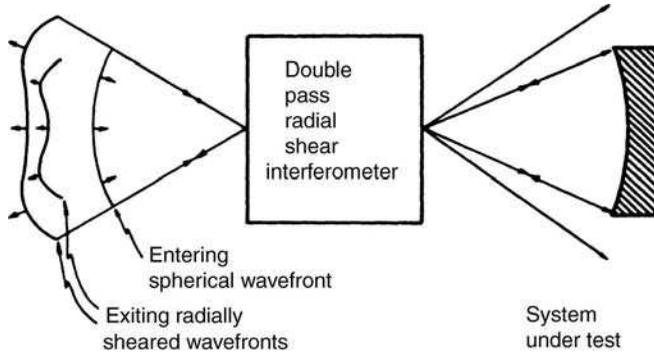


FIGURE 5.14. Double-pass radial shear interferometer.

the total amount of energy. If the reflected pattern were present in the double-pass interferometer, the observed pattern would be extremely complicated because four returning beams instead of two would be interfering. This problem can be avoided if the two sheared interfering wavefronts illuminating the system under test after the first pass do not interfere with each other. This can be achieved if the two wavefronts are produced with orthogonal polarizations, but are nevertheless coherent to each other. As pointed out by Brown (1959) and later by Steel (1965), the spatial coherence requirements are greatly relaxed on double-pass interferometers because they are essentially compensated for the size of the source. Although the surface cannot be made extremely large—only about three times (Brown, 1959) as large as that in the single-pass interferometer—this provides about ten times more light.

Two interferometers of this type have been designed by Steel to test microscope objectives. One of these is shown in Figure 5.15 (Steel, 1965). The radial shear is produced by two birefringent systems, each formed by two calcite components between two glass lenses. The glass is Schott La K11, chosen to match the ordinary index of calcite and to correct the chromatic aberration of the lens. The calcite lenses are designed to have small off-axis errors, are equiconcave, and are divided into two parts by plane surfaces. The optical axes of the two halves are at 90° with respect to each other. The plane of polarization between the halves is rotated 90° by means of a half-wave plate, so that the ordinary ray in the first half remains an ordinary ray in the second half. The two bi-refringent systems have a relative orientation such that the ordinary ray of the first system becomes the extraordinary ray of the second system and vice versa.

The two bi-refringent systems are adjusted to satisfy the following conditions: (a) the apparent point of divergence of the two radially sheared wavefronts is at the proper distance (16 cm) from the microscope objective and (b) the focal plane of the whole birefringent system coincides with the exit pupil of the microscope objective, which also coincides with the back focal plane of this objective. For the reasons explained in Chapter 2, a Dyson system is used in front of the microscope objective. Since the fringes must be observed at the exit pupil of the objective, a telescope is used to look at them.

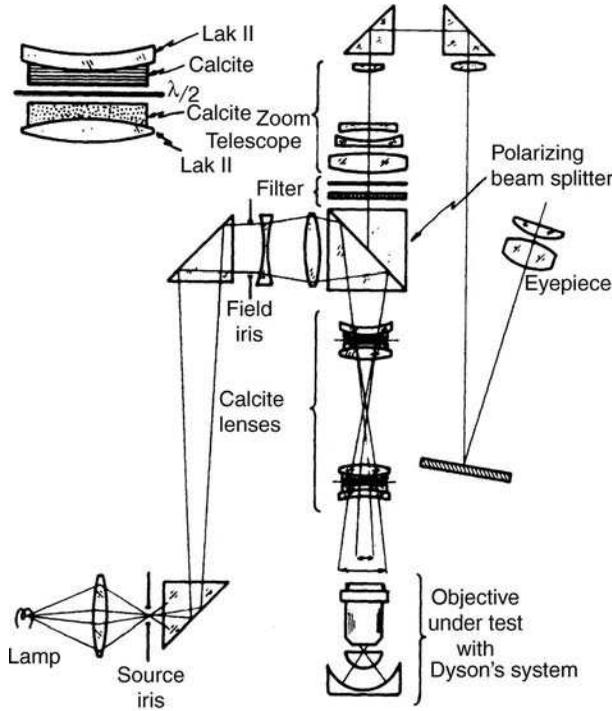


FIGURE 5.15. Steel's double-passage radial shear interferometer to test microscope objectives.

The second double-passage radial shear interferometer designed by Steel (1966) is illustrated in Figure 5.16. It is similar to the Hariharan and Sen interferometer except that the two lenses are replaced by a single lens from a low-power microscope objective.

The combination of the polarizing beam splitter 1 and the $\lambda/4$ plate forms a source of circularly polarized light. When the light returns from the instrument after being reflected on the system under test, the direction of rotation of the circularly polarized light is reversed (assuming a perfect system). Therefore, the returning light passes through the prism to go to the camera. If the system under test is imperfect, an optical path difference between the two radially sheared and orthogonally polarized beams is introduced. Then the returning light will not be circularly but elliptically polarized, giving rise to dark zones (fringes) on the camera.

5.2.4. Laser Radial Shear Interferometers

The radial shear interferometers so far described are of the equal-path type with white-light compensation. This is necessary when conventional light sources are used. When a laser is employed, the two beams do not need to have the same optical paths.

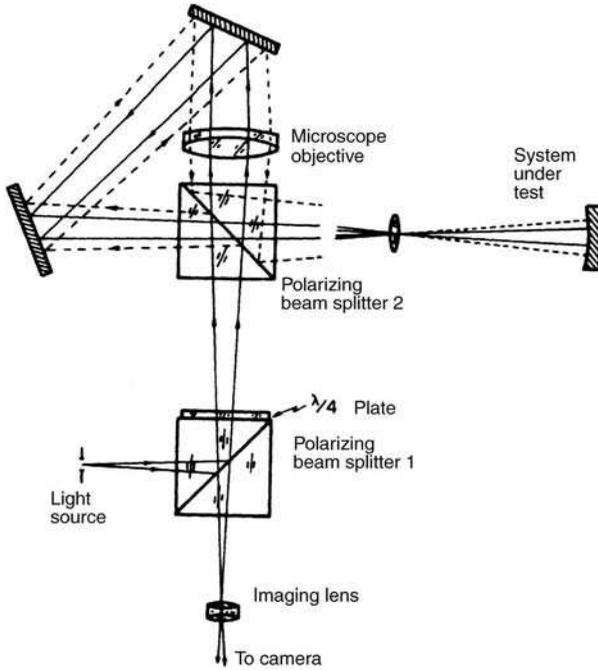


FIGURE 5.16. Steel's double-passage radial shear interferometer.

With laser light the design of the interferometer is greatly simplified, but some new problems are introduced. The main problem is the presence of many spurious fringes over the desired interference pattern due to reflections on lens surfaces and glass plates. All reflections produce interference fringes because of the long coherence length of laser light.

Probably the first laser radial shear interferometer was designed by Som (1970); it is illustrated in Figure 5.17. There is, however, a great problem with this design since the virtual points of divergence P_1 and P_2 for the two wave-fronts do not coincide, as pointed out by Murty and Shukla (1973). Hence, a perfect system produces a system of concentric circular fringes, similar to Newton rings, and it is difficult to analyze the interferogram with this as a reference. Murty points out that for easier analysis a perfect optical system under test must give either a fringe-free field or a set of straight fringes.

To eliminate this problem, Murty and Shukla (1973) modified Som's design and proposed the interferometer shown in Figure 5.18, in which one of the reflecting surfaces is spherical. If a and b are the distances from the concave and plane mirrors, respectively, to the center of the beam-dividing surface, the radius of curvature r of the reflecting surface must be

$$r = \frac{(2b - a)a}{b - a}. \tag{5.16}$$

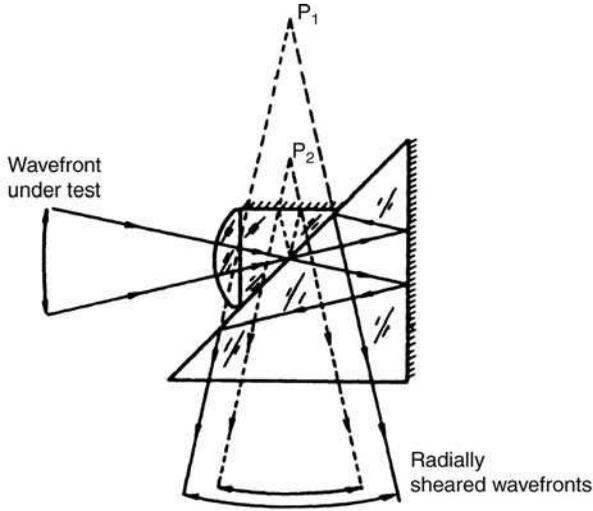


FIGURE 5.17. Som's laser radial shear interferometer.

The effective radial shear R is then

$$R = \frac{a}{2b - a}. \tag{5.17}$$

This interferometer can be fabricated very easily from a solid cube beam splitter. Ideally, the reflecting surface should be a hyperboloid of revolution and thus a

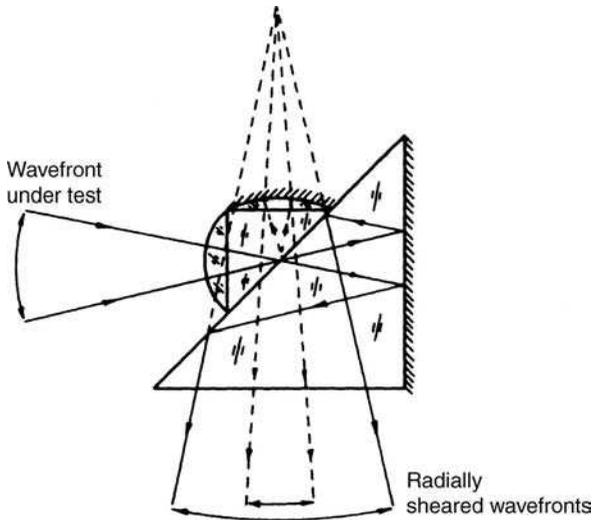


FIGURE 5.18. Murty's laser radial shear interferometer.

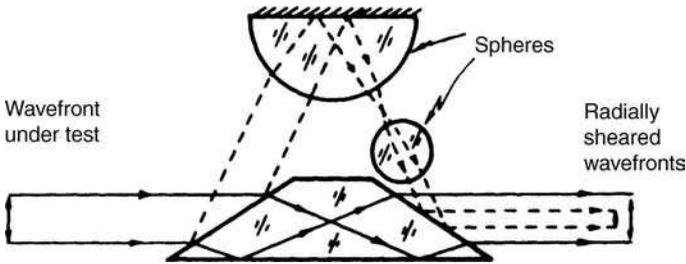


FIGURE 5.19. Steel's laser radial shear interferometer.

spherical surface introduces a small amount of spherical aberration. Murty et al. (1975) showed that the wavefront spherical aberration is given by

$$OPD = \frac{-Na^4b^2\alpha^4}{(b-a)^2r^3} \quad (5.18)$$

where N is the refractive index of the glass and α is the numerical aperture. Thus, it is necessary to reduce a as much as possible. If the numerical aperture is small, very high accuracies can be obtained.

Malacara et al. (1975) designed a lateral shearing interferometer for converging or diverging laser beams that uses a plano-concave prismatic glass plate. Similar to the interferometer just described, this instrument also produces some radial shear together with the lateral shear.

Hariharan et al. (1984a, 1984b) has used an interferometer based on the instrument designed by Murty and Shukla (1973) in order to perform phase shifting interferometry.

Another of the early laser radial shear interferometers was designed by Steel (1970); it is illustrated in Figure 5.19. The two beams are split and recombined at the two surfaces of a prism, thus eliminating the possibility of unwanted fringes on the second face of a beam splitter. The diameter of one of the beams is reduced in size by means of a telescope with 5.5 magnification. The advantage of using glass spheres is that they do not need to be squared to the beam; a disadvantage is that they introduce spherical aberration. It must be pointed out that, simultaneously with the radial shear, this interferometer also produces reversal shear, since one of the wavefronts is reversed with respect to the other.

Another laser light radial shear interferometer has been described by Shukla et al. (1992), as illustrated in Figure 5.20. It is made from a cube beam splitter either by grinding and polishing or by cementing two plano convex lenses on the opposite ends of the cube. These two convex surfaces are at different distances from the center of the beam splitter interface surface, so that their centers of curvature are located at this center. We can easily see that the radial shear is equal to the ratio of the two radii of curvature. Since the surfaces are spherical, some spherical aberration is introduced, but its magnitude can be set within a reasonable limit. It is also interesting to notice that one of the two beams is reflected twice and transmitted once on the beam

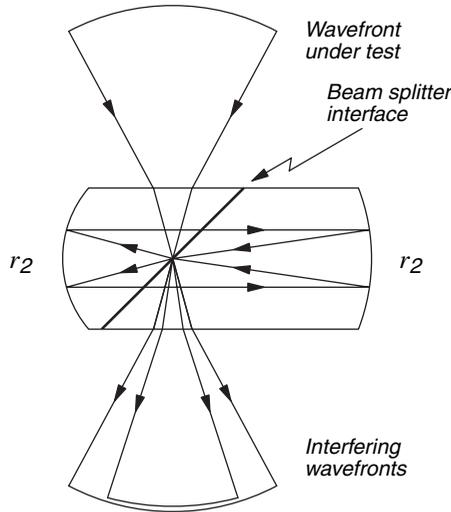


FIGURE 5.20. On-axis radial shear interferometer.

splitter, while the other beam is just transmitted once. Thus, the intensities of the two interfering beams will be quite different. However, this is partially compensated because the weaker beam is the one with a smaller diameter. Shukla et al. (1992) has also described several other radial shear interferometers based on the same working principle.

A very simple holographic radial shear interferometer, shown in Figure 5.21 was devised by Fouéré and Malacara (1974) and Fouéré (1974). The first step in making

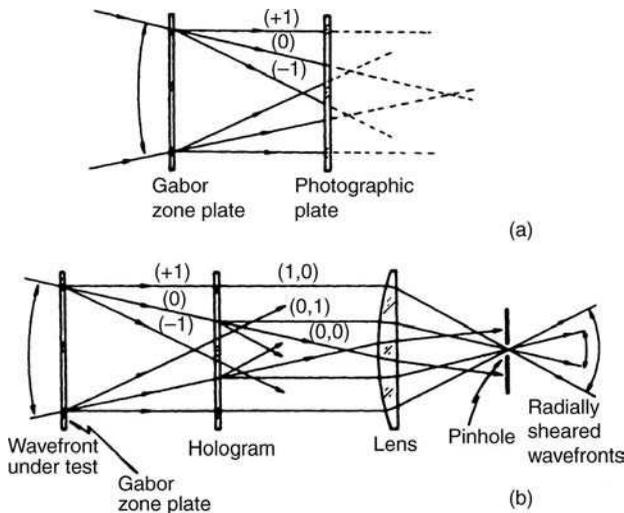


FIGURE 5.21. Holographic radial shear interferometer.

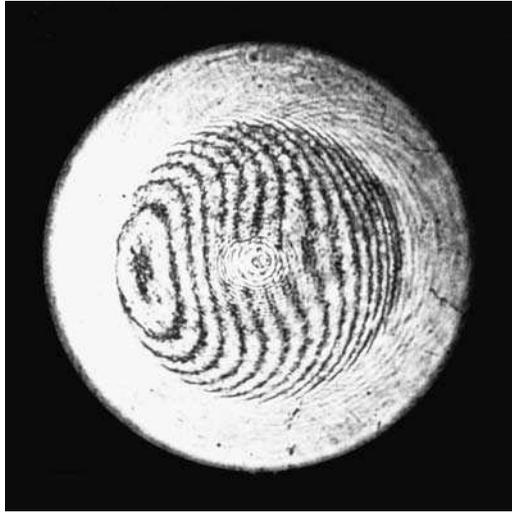


FIGURE 5.22. Interferogram obtained in holographic interferometer. (From Fouéré and Malacara, 1974.)

this interferometer is to fabricate a Gabor zone plate by photographing the interference between a convergent and a flat wavefront. The numerical aperture of this plate is the numerical aperture that the interferometer will accept. The second step is to illuminate the Gabor zone plate with a convergent wavefront. If the point of convergence corresponds exactly to the focus of the zone plate, the diffracted +1 order beam emerges as a parallel beam. Let us now place a photographic plate behind the zone plate and make an exposure. By developing the photographic plate, we obtain a hologram. The convergent beam (order 0) can be regarded as the reference beam and the diffracted parallel beam (order +1) as the object beam.

Considering Figure 5.21 now, if we place the hologram exactly in its original position, we obtain an emergent parallel beam (0, 1) by reconstruction. The incident parallel beam goes through the hologram as the parallel beam (1, 0). Thus, these two beams emerge parallel but with different magnifications, achieving a radial shear interferometer. All undesired diffracted beams are filtered out by means of a lens and a pinhole at its focus. It may be shown that all imperfections on the glass plates are automatically canceled out. Tilt and defocusing in this interferometer can very easily be obtained by small lateral and longitudinal movements, respectively, of the Gabor zone plate. Figure 5.22 shows an interferogram obtained with this interferometer.

Several variations of this basic configuration of the two-zone plates have been proposed, in order to have rotational shear also or to produce an exiting spherical, instead of flat wavefront.

5.2.5. Thick-Lens Radial Shear Interferometers

Steel (1975) and Steel and Wanzhi (1984) have described an interesting class of laser interferometers called thick-lens radial shear interferometers. Figure 5.23 shows

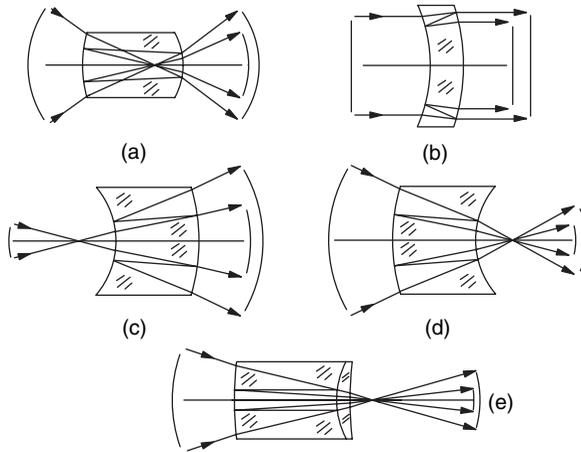


FIGURE 5.23. Thick-lens laser radial shear interferometers.

some designs. One of the beams is obtained from the direct beam, going through the thick lens. The other beam is obtained with two internal reflections, first on the second lens face and then on the first lens face. If the direct beam has an irradiance I_1 , the internally reflected beam has an irradiance I_2 given by

$$I_2 = \frac{R_1 R_2}{S^2} I_1 \tag{5.19}$$

where S is the radial shear, defined by the ratio of the diameter of the internally reflected beam to the diameter of the direct beam. The maximum constant of the fringes is obtained when $R_1 R_2 = S^2$. This is possible only if $S < 1$, in other words, when the direct beam has the smallest diameter. Steel and Wanzhi suggest keeping the contrast to a low value of the order of 0.8, to reduce the disturbing effects of higher-order ghosts due to secondary internal reflections. Unfortunately, these reflections arrive at the same focus, and thus they cannot be eliminated by a spatial filter.

Figure 5.23 shows some thick-lens interferometers. The first one, Figure 5.23(a), is a radial as well as a rotational shear interferometer; the rest are pure radial shearing. The interferometer in Figure 5.23(b) may be used to test large collimated beams. In the interferometer in Figure 5.23(c), the direct beam has a diameter smaller than the reflected beam. Thus, the contrast of the fringes cannot be controlled. The system in Figure 5.23(e) may be used in convergent beams with the exit pupil located far away.

The reflecting radial shear interferometer described by Wanzhi (1984) is a special case of the thick lens interferometer working in a reflection mode as shown in Figure 5.24(a) and 5.24(b). The fringe contrast is much better than in the transmission system. The only problem with this glass-spaced system is that the magnitude of the shear is fixed. An alternative is the air-spaced system shown in Figure 5.24(c), as described by Wanzhi (1985).

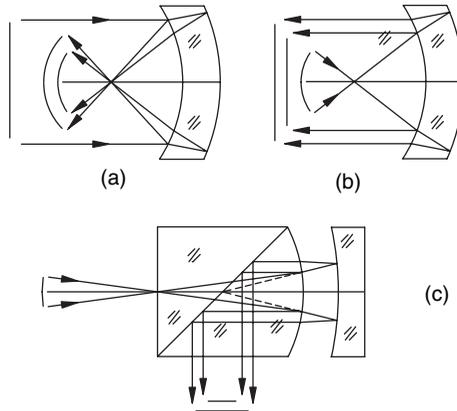


FIGURE 5.24. Reflection laser radial shear interferometers.

5.3. ROTATIONAL SHEAR INTERFEROMETERS

In rotational shear interferometers, identical wavefronts are made to interfere, but one of them is rotated with respect to the other by a certain angle about their common optical axis. They are useful for detecting and evaluating nonsymmetrical aberrations such as coma and astigmatism. Many procedures have been developed for the interpretation of rotational shearing interferograms. For example, Golikov et al. (1981) have described a method using cylindrical coordinates.

Let a wavefront be represented by $W(\rho, \theta)$. A rotational shear interferometer is one that performs the operation of rotating one wavefront with respect to the other to give an interferogram defined by

$$\text{OPD}(\rho, \theta) = W\left(\rho, \theta - \frac{\phi}{2}\right) - W\left(\rho, \theta + \frac{\phi}{2}\right) \quad (5.20)$$

where ϕ is the rotation of one wavefront with respect to the other. If the wavefront is given by general expression (5.1), we obtain

$$\text{OPD}(\rho, \theta) = \sum_{n=0}^k \sum_{l=0}^n \rho^n \left\{ a_{nl} \left[\cos l \left(\theta - \frac{\phi}{2} \right) - \cos l \left(\theta + \frac{\phi}{2} \right) \right] + b_{nl} \left[\sin l \left(\phi - \frac{\theta}{2} \right) - \sin l \left(\theta + \frac{\phi}{2} \right) \right] \right\} \quad (5.21)$$

where n and l are both even or both odd. It is interesting that all terms for $l = 0$ cancel out, rendering the rotational shear interferometers insensitive to rotationally symmetric wavefronts, as could be expected. Therefore, the sums in this expression can be started from $n = l = 1$. If we now assume that the aberrations are produced by an

axially symmetric optical system, we have the wavefront symmetric about the tangential ($y - z$) plane and therefore, all coefficients b_{nl} become zero. Thus

$$\begin{aligned} \text{OPD}(\rho, \theta) &= \sum_{n=1}^k \sum_{l=1}^n \rho^n a_{nl} \left[\cos l \left(\theta - \frac{\phi}{2} \right) - \cos l \left(\theta + \frac{\phi}{2} \right) \right] \\ &= \sum_{n=1}^k \sum_{l=1}^n 2\rho^n a_{nl} \sin l\theta \sin \frac{l\phi}{2}. \end{aligned} \quad (5.22)$$

The only two primary aberrations contained in this expression are astigmatism ($n = 2, l = 2$) and coma ($n = 3, l = 1$), in addition to a tilt of the wavefront about the x axis ($n = 1, l = 1$) that we may ignore, thus obtaining

$$\text{OPD}(\rho, \theta) = 2a_{22}\rho^2 \sin 2\theta \sin \phi + 2a_{31}\rho^3 \sin \theta \sin \frac{\phi}{2} \quad (5.23)$$

which can also be written as

$$\text{OPD}(\rho, \theta) = 2a_{22}\rho^2 \cos 2\left(\theta + \frac{\pi}{4}\right) \sin \phi + 2a_{31}\rho^3 \cos\left(\theta + \frac{\pi}{2}\right) \sin \frac{\phi}{2} \quad (5.24)$$

Observing this expression, we can see that the sensitivity σ_{ast} for astigmatism with respect to that of a Twyman-Green interferometer is given (Murty and Hagerott, 1966) by

$$\sigma_{\text{ast}} = 2 \sin \phi, \quad (5.25)$$

but it is also important to note that the interference pattern is rotated 45° with respect to the Twyman-Green pattern, as shown in Figure 5.25, and that the rotational shear interferometer is insensitive to real defocusing. The apparent focus is such that the observed pattern is similar to the Twyman-Green pattern at the intermediate focus between the tangential and sagittal foci.

The relative sensitivity σ_{coma} for coma is given (Murty and Hagerott, 1966) by

$$\sigma_{\text{coma}} = 2 \sin \frac{\phi}{2}, \quad (5.26)$$

and the interference pattern is rotated 90° with respect to the corresponding Twyman-Green interferogram, as shown in Figure 5.26. No defocusing appears, nor can it be introduced in this pattern.

The relative sensitivities σ_{ast} and σ_{coma} for astigmatism and coma, respectively, are plotted in Figure 5.27 which shows that the ability of a rotational shear interferometer to detect astigmatism and coma depends on the amount of rotational shear ϕ . We can see that the coma can be isolated at $\phi = 180^\circ$, but the astigmatism

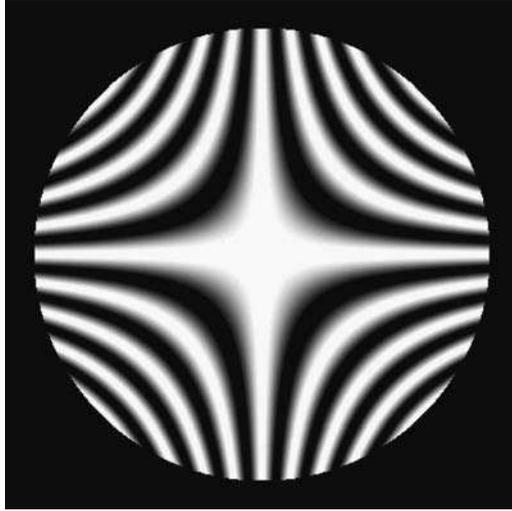


FIGURE 5.25. Astigmatism pattern in a rotational shear interferometer with maximum sensitivity ($\phi = 90^\circ$)($\alpha_{22} = 2\lambda$).

can never be isolated. However, the coma can be eliminated and the astigmatism doubled by the use of a small, flat mirror to allow the wavefront to go twice through the optical system. The wavefront goes first to the system under test (concave mirror), proceeds next to the small, flat mirror near the center of curvature of the concave mirror, and returns through the same path.

It is important to point out the difference between the two similar processes, in which either the symmetrical (power of $\cos \theta$ is even) or the antisymmetrical (power

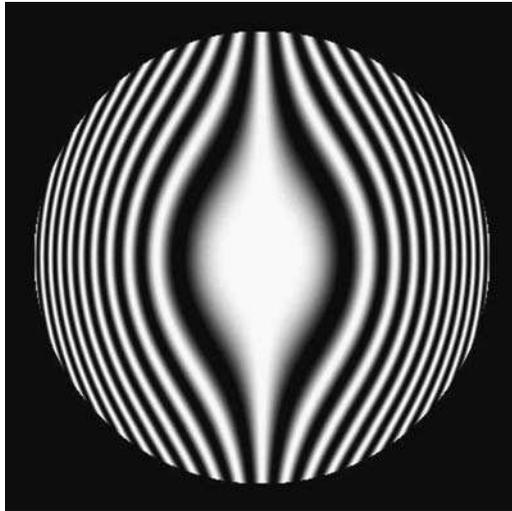


FIGURE 5.26. Coma pattern in a rotational shear interferometer with maximum sensitivity ($\phi = 180^\circ$)($\alpha_{31} = 5\lambda$).

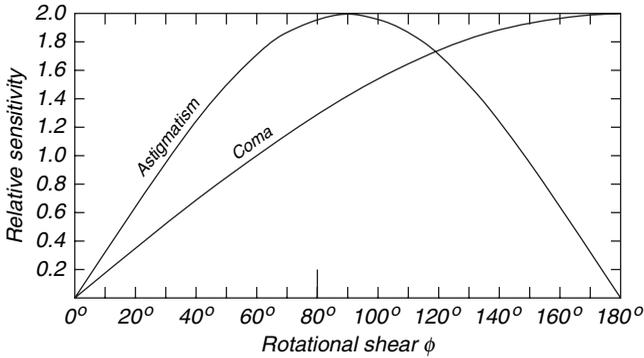


FIGURE 5.27. Relative sensitivity for astigmatism and coma in a rotational shear interferometer.

of $\cos \theta$ is odd) components of the wavefront are isolated. To isolate the asymmetrical components, the wavefront is made to interfere with an image of itself that is identical but is rotated 180° , in order to obtain the difference between them. To isolate the symmetrical components, the wavefront is rotated 180° and then passed again through the optical system under test, in order to duplicate the symmetrical errors and to eliminate the asymmetrical ones. Then the wavefront is made to interfere with an unaberrated wavefront. These two processes have been exploited by Hariharan and Sen (1961c) in a single arrangement, as described in Chapter 7 of this book.

As pointed out by Murty and Hagerott, it is interesting to consider the testing of a ribbed lightweight mirror. When such a mirror is polished, the region of the face directly on top of the rib is low, whereas the region between the ribs is high. The astigmatism of a mirror of this kind ($n = 2, l = m$) produces on a rotational shear interferometer on OPD given by

$$OPD = 2\rho^n a_{nl} \sin m\theta \sin \frac{m\phi}{2} \tag{5.27}$$

where m is the number of ribs supporting the faceplate. It should be noticed that $n = m$ are both even or both odd. The maximum relative sensitivity is 2 and occurs for $\phi = 180^\circ$. Figure 5.28 shows the interferogram for a mirror with four ribs.

5.3.1. Source Size Uncompensated Rotational Shear Interferometers

The degree of coherence g_{12} of a uniform circular source with angular diameter 2α is given in Eq. (5.10). Thus to maintain good fringe contrast, the pinhole of a rotational shear interferometer must be of the proper size. If a point of the wavefront is sheared through an angle ϕ , the distance in the wavefront between two interfering points is given by

$$d = 2\rho \sin \frac{\phi}{2}. \tag{5.28}$$

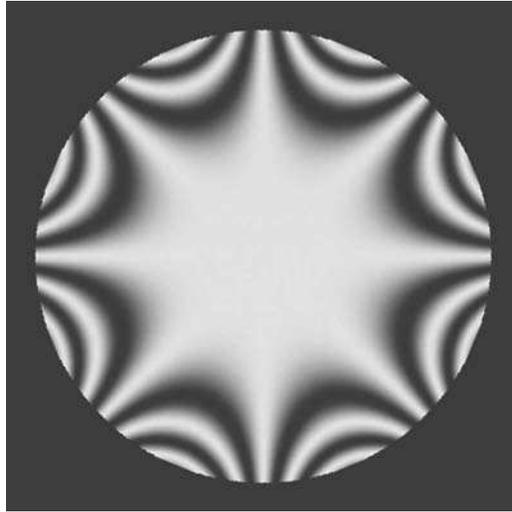


FIGURE 5.28. Pattern for a mirror with four ribs in a rotational shear interferometer with maximum sensitivity ($\phi = 45^\circ$)($\alpha_{24} = 4\lambda$).

Substituting this value in Eq. (5.10), we obtain

$$g_{12} = \frac{2J_1[(4\pi/\lambda)\alpha\rho \sin(\phi/2)]}{(4\pi/\lambda)\alpha\rho \sin(\phi/2)}. \quad (5.29)$$

For the reasons given in Section 5.2.1, to have a good contrast on the interference pattern, the pinhole must be smaller than the certain value given by Murty and Hagerott (1966) as

$$\alpha \leq \frac{1.22\lambda}{2D \sin(\phi/2)} \quad (5.30)$$

where D is the interferometer entrance pupil diameter, from which the angular diameter 2α of the source is measured.

Murty and Hagerott also designed the rotational shear interferometer shown in Figure 5.29. It is a Jamin interferometer with two identical Dove prisms between the glass plates. A microscope objective L_1 collimates the light from a point source, and the pattern is observed by looking through the microscope objective L_2 . One of the wavefronts is rotated an angle 2α by rotating one of the Dove prisms an angle α .

An ordinary Twyman–Green interferometer can be converted to a rotation shear interferometer, in order to test the illuminating wavefront, by means of several procedures. A 180° rotational shear is achieved, as suggested by Murty (1964c), by replacing the mirror in one of the arms with a cat’s-eye retroreflector or a cube corner prism. Armitage and Lohmann (1965) suggested using two roof prisms as reversing prisms instead of the flat mirrors. The magnitude of the rotational shear ϕ is

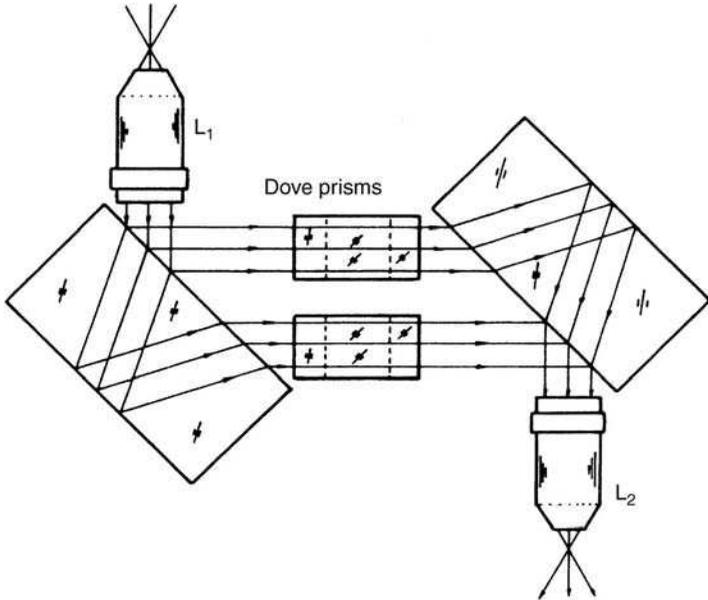


FIGURE 5.29. Murty and Hagerott's rotational shear interferometer.

changed by rotating one of the prisms about its optical axis by an amount $\phi/2$. The interferometer proposed by Armitage and Lohmann is illustrated in Figure 5.30. The wavefront under test is collimated by the lens L_1 or a pinhole is used if L_1 is the lens to be tested. The pair of lenses L_2 and L_3 forms an image of the entrance pupil at L_1 on the observing screen. To obtain good coherence, the state of polarization of both beams must be identical, but this is not so when the rotating roof prism is rotated at an angle $\phi/2$, because the roof prisms modify the state of polarization according to their positions. To solve this problem, a polarization coupling is used as follows. Two polarizers are included, one before and one after the interferometer, both at 0° . Also, two quarter-wave plates are placed between the main interferometer body and the

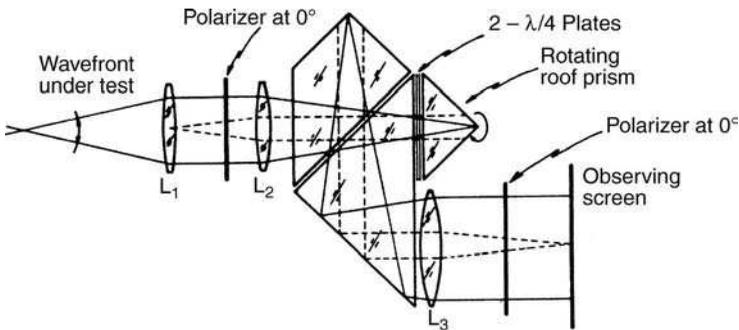


FIGURE 5.30. Armitage and Lohmann's rotational shear interferometer.

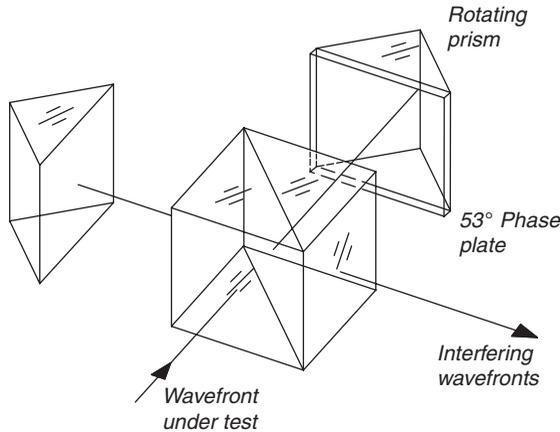


FIGURE 5.31. Rotational shear interferometer with phase compensation.

rotating roof prism. One plate is fixed at 45° to the main body of the interferometer and the other plate is fixed to the roof prism at 45° to the line of intersection of the roof planes. Between the two quarter-wave plates, the light will be circularly polarized. Within the roof prism, it will be linearly polarized and parallel to the roof ridge. Upon returning into the main body, the old direction of linear polarization is restored, independently of the rotational position of the roof prism.

The problem with the interferometer just described is that the use of the polarizers causes an appreciable loss of light. An alternative solution has been employed by Roddier et al. (1978, 1989) in the interferometer shown in Figure 5.31. There are two roof prisms, one is fixed and the other is rotatable about its optical axis, as in the interferometer by Armitage and Lohmann. Let us assume for the time being that the light entering the interferometer is linearly polarized, with the plane of polarization either parallel or perpendicular to the fixed roof ridge. If the rotatable roof prism is rotated at an angle with respect to the first prism, we may decompose the electric vector in two orthogonal components, one parallel and one perpendicular to the roof ridge of this second prism. The phase shift upon reflection for these two components will be different in general, so, they will recombine after leaving the prism producing a different polarization state from that of the entering light beam. The only way to preserve the incident polarization state is to have a phase shift difference between these two components equal to 180° . The electrical component perpendicular to the roof edge comes out of the prism in opposite direction because of the two reflections. The net effect is that the linear polarization and its direction is thus preserved.

Since the total shift upon reflection for the two electrical components is about 74° , an additional phase shift of about 106° is required. This is obtained by cementing to the entrance face (hypotenuse) of the rotatable prism a phase plate with a phase retardation of 53° , with its principal axes parallel and perpendicular to the roof ridge.

We assumed at the beginning that the light entering the interferometer was linearly polarized, with the plane of polarization either parallel or perpendicular to

the fixed roof ridge. If the light is unpolarized, the components parallel and perpendicular to the roof ridge of the fixed prism will produce different interference patterns. The same interference pattern is produced with these two components if they come out of the fixed prism in phase, as when they entered the prism. This goal is achieved by cementing another phase plate to the fixed prism. Then, unpolarized light may be used.

Another interferometer suggested by Armitage and Lohmann (1965) is based on the Sagnac or cyclic interferometer. Here the rotational shearing angle ϕ is produced by a Dove prism rotation of $\phi/4$ within the closed loop of the interferometer.

Although not really intended as a rotational shear interferometer, the inverting Fizeau interferometer designed by Sen and Puntambekar (1965, 1966) produces a rotational shear of 180° . This instrument and its adaptation to test spherical surfaces (Puntambekar and Sen, 1971) are described in Chapter 7.

5.3.2. Source Size Compensated Rotational Shear Interferometers

A large light source cannot be used in general because of the coherence problems explained in the preceding section. The only way to have good contrast with a large source is to make the two images of the source coincide in position and orientation. At the same time, the images of the object under test must be shared with respect to each other. One solution is to shear the wavefronts before they reach the object, with an equal magnitude but in an opposite sense to the main shear, which is to take place after the object. These two shearing operations cancel each other as far as coherence is concerned, but only the second shear takes place with the information about the object. These considerations, equivalent to those made before for radial shear interferometers, were also advanced by Armitage and Lohmann (1965) in suggesting several compensated interferometers. The systems they proposed consist basically of two identical interferometers placed symmetrically one after the other, with the object under test between them.

As previously explained the same compensation can be obtained by using a rotational shear interferometer in a double-pass configuration.

5.4. REVERSAL SHEAR INTERFEROMETERS

The reversal of a wavefront about a reversing axis is illustrated in Figure 5.32 where point P goes to point P' , according to the transforming equations

$$\rho' \sin \theta' = \rho \sin \theta \quad (5.31)$$

and

$$\rho' \cos \theta' = S - \rho \cos \theta \quad (5.32)$$

It is easy to see that this reversing about an arbitrary axis is equivalent to a reversion about the x axis followed by a lateral shear S in the y direction. If the

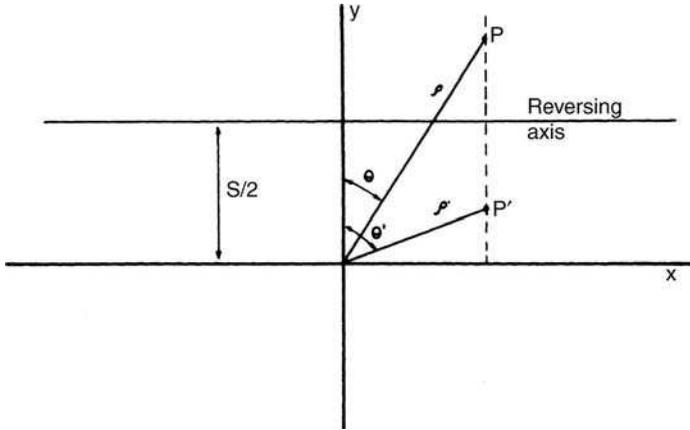


FIGURE 5.32. Reversal of a wavefront.

wavefront is defined by Eq. (5.1), the interferogram in a reversal shear interferometer is given by

$$\text{OPD} = W(\rho, \theta) - W(\rho', \theta'). \quad (5.33)$$

Let us consider a wavefront that has only primary aberrations as follows:

$$W(\rho, \theta) = a_{20}\rho^2 + a_{40}\rho^4 + a_{31}\rho^3 \cos \theta + a_{22}\rho^2 \cos 2\theta + a_{11}\rho \cos \theta \quad (5.34)$$

where the aberrations being represented are defocusing, spherical aberration, coma, astigmatism, and a tilt about the x axis, respectively.

We can see that the reversal shear interferometer has no sensitivity to symmetric aberrations, such as defocusing (a_{20}), spherical aberration (a_{40}), and astigmatism (a_{22}), if the axis of reversion coincides with the x axis ($S = 0$). The interferometric pattern and the sensitivity of this interferometer to symmetrical aberrations are, however, identical to those of a lateral shearing interferometer with shear S when the axis of reversion is shifted at a distance $S/2$, as in Figure 5.32

$$\text{OPD} = 2a_{31}\rho^3 \cos \theta + 2a_{11}\rho \cos \theta. \quad (5.35)$$

Thus the relative sensitivity of the reversal shear interferometer without any lateral shear ($S = 0$) to antisymmetric aberrations, such as coma and tilt about the x axis, is equal to 2.

If the interferometer is uncompensated for the size of the light source, that is, if the reversion affects not only the object under test but also the light source, the contrast in the fringe pattern is given by Eq. (5.10) (Murty, 1964c). The distance for any two interfering points is ($2\rho \cos \theta - S$). Therefore, its maximum value d is

$$d = D + S \quad (5.36)$$

where D is the waveform diameter. Then the maximum angular diameter α of the pinhole will be

$$\alpha \leq \frac{1.22\lambda}{D + S}. \tag{5.37}$$

An uncompensated reversal shear interferometer of this type is represented by a Twyman–Green interferometer with a right-angle prism used as a roof prism in one of the arms.

5.4.1. Some Reversal Shear Interferometers

In this section, three different versions of a prism reversal shear interferometer are described. They are variations of a basic prism system invented by Kösters (1934), who pointed out that any combination of two exactly similar prisms can be used as an interferometer in the way he described, as long as one of the angles of the prism adjoining the common face is exactly half the angle that is opposite this common face.

One of the two systems suggested by Gates (1955b), based on Kösters’s prism, is illustrated in Figure 5.33; it uses two 30° - 60° - 90° prisms. The intersection of the plane defined by the beam-splitting surface with the optical system defines the axis of reversion. Lateral shear S is produced if the axis of reversion does not coincide with a diameter of the optical system under test. A tilting of the mirror under test provides a control of the separation of the interference fringes. A lens can be tested if an autocollimating flat mirror is used behind the lens.

The interferometer is compensated for use with white light because the optical paths are exactly equal. It is also compensated for the size of the light source, but in practice this size is limited to about 0.5 mm in diameter by small errors in construction, as pointed out by Gates.

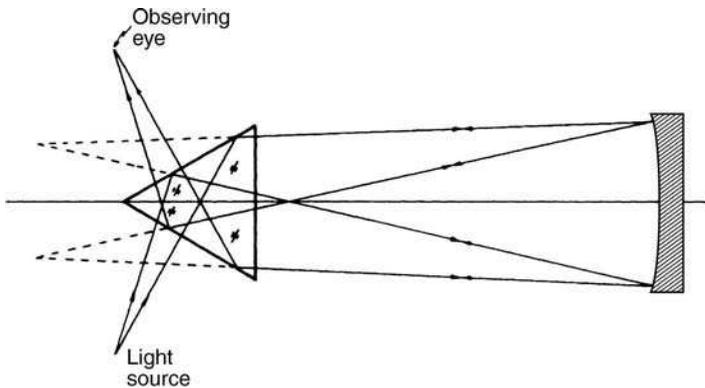


FIGURE 5.33. Köster’s reversal shear interferometer.

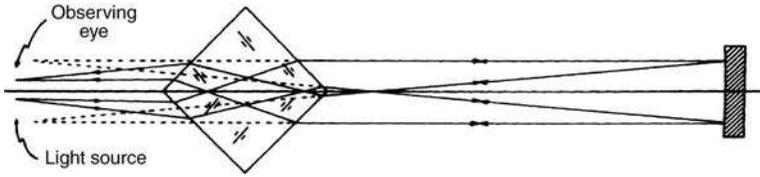


FIGURE 5.34. Gates' reversal shear interferometer.

Using the same principle, Gates (1955a) also suggested using the popular beam splitter cube as shown in Figure 5.34. This, however, has the disadvantage that only optical systems with small numerical apertures can be tested.

Two common disadvantages of both instruments described by Gates are that (a) the virtual light source and the image do not coincide but have some small separation and (b) some aberrations are introduced on the plane exit face of the prism. These problems are not present in a prism system invented by Saunders (1955) shown in Figure 5.35. The use of this instrument for the testing of optical systems, as described by Saunders (1962), is illustrated in Figure 5.36. Many other applications to metrology have also been found (Strong, 1958; Saunders, 1960). The construction method has been very well described by Saunders (1957).

Another interesting reversal shear interferometer to test lenses was described by Waetzman (1912) and later by Murty (1964b). It is illustrated in Figure 5.37. This instrument, which has also been used to test right-angle and cube corner prisms by Murty (1964b) and Sen and Puntambekar (1966), produces reversal and lateral shears. It is interesting to observe that the asymmetric aberrations are canceled out because of the double passage through the lens and that the reversal interferometer is

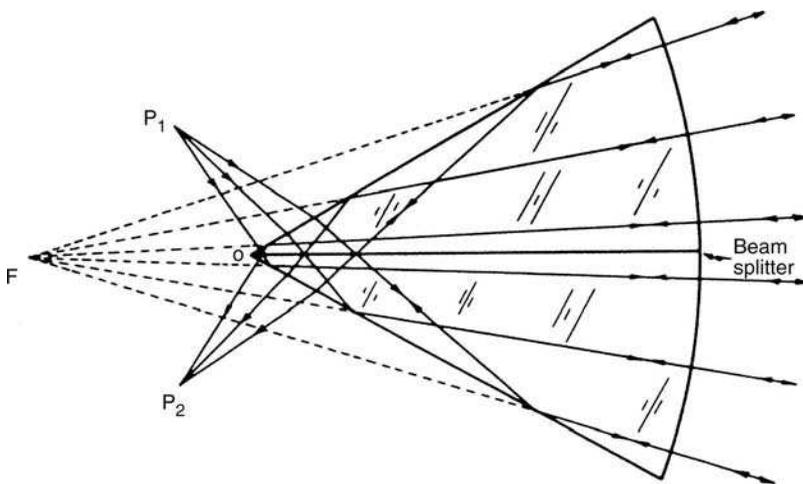


FIGURE 5.35. Saunders's prism system.

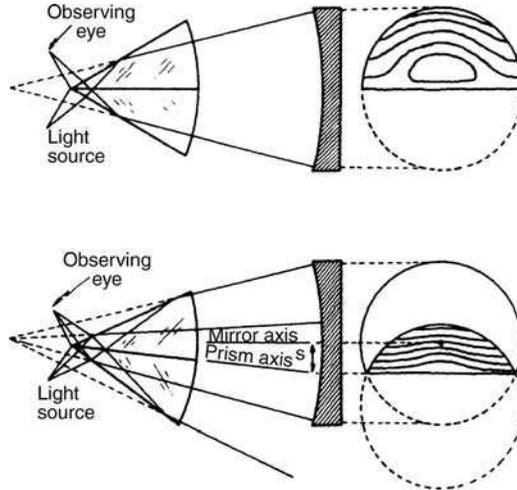


FIGURE 5.36. Saunders' reversal shear interferometer.

not sensitive to symmetric aberrations. Since the symmetric aberrations are detected by means of lateral shear only, this instrument can more properly be considered as a lateral shearing interferometer.

Holographic techniques may also be used to obtain a reversal shear interferometer as shown by Partiban et al. (1987, 1988).

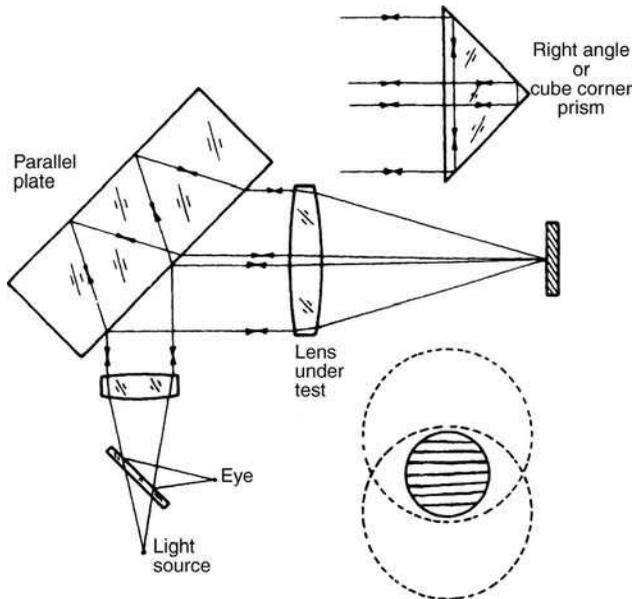


FIGURE 5.37. Modified Jamin interferometer to test lenses and prisms.

REFERENCES

- Armitage J. D. and A. Lohmann, "Rotary Shearing Interferometry," *Opt. Acta*, **12**, 185 (1965).
- Baird K. M. and G. R. Hanes, in: *Applied Optics and Optical Engineering*, Vol. 4, R. Kingslake, Ed., Academic Press, New York, 1967, p. 336.
- Briers J. D., "Interferometric Testing of Optical Systems and Components: A Review" *Opt. Laser Technol.*, **4**, 28(1972).
- Brown D. S., "Radial Shear Interferograms," *Interferometry N.P.L. Symposium* No. 11, Her Majesty's Stationary Office, London, 1959, p. 253.
- Brown D. S., "Radial Shear Interferometry," *J. Sci. Instrum.*, **39**, 71 (1962).
- Bryngdahl O., in: *Progress in Optics*, Vol. IV, E. Wolf, Ed., North-Holland, Amsterdam, 1965, Chap. II.
- Bryngdahl O., "Reversed-Radial Shearing Interferometry," *J. Opt. Soc. Am.*, **60**, 915 (1970).
- Bryngdahl O., "Shearing Interferometry with Constant Radial Displacement," *J. Opt. Soc. Am.*, **61**, 169 (1971).
- Fouéré J. C., "Holographic Interferometers for Optical Testing," *Opt. Laser Technol.*, **6**, 181 (1974).
- Fouéré J. C. and D. Malacara, "Holographic Radial Shear Interferometer," *Appl. Opt.*, **13**, 2035 (1974).
- Fouéré J. C. and D. Malacara, "Generalized Shearing Interferometry," *Bol. Inst. Tonantzintla*, **1**, 227 (1975).
- Gates J. W., "Reverse Shearing Interferometry," *Nature*, **176**, 359 (1955a).
- Gates J. W., "The Measurement of Comatic Aberrations by Interferometry," *Proc. Phys. Soc.*, **B68**, 1065 (1955b).
- Golikov A. P. M., L. Gurari, and S. I. Prytkov, "Interpretation of Rotational-Shearing Interferograms," *Sov. J. Opt. Technol.*, **48**, 676 (1981).
- Hariharan P. and D. Sen, "Effects of Partial Coherence in Two Beam Interference," *J. Opt. Soc. Am.*, **51**, 1307 (1961a).
- Hariharan P. and D. Sen, "Radial Shearing Interferometer," *J. Sci. Instrum.*, **38**, 428 (1961b).
- Hariharan P. and D. Sen, "The Separation of Symmetrical and Asymmetrical Wavefront Aberrations in the Twyman Interferometer," *Proc. Phys. Soc.*, **77**, 328 (1961c).
- Hariharan P. and D. Sen, "Interferometric Measurements of the Aberrations of Microscope Objectives," *Opt. Acta*, **9**, 159 (1962).
- Hariharan P., B. F. Oreb, and Z. Wanzhi, "A Digital Radial Shear Interferometer for Testing Aspheric Surfaces," *Proceedings of 13th Congr. ICO.*, Sapporo, Japan, 1984a, P464.
- Hariharan P., B. F. Oreb, and Z. Wanzhi, "Measurement of Aspheric Surfaces Using a Microcomputer-Controlled Digital Radial-Shear Interferometer," *Opt. Acta*, **31**, 989 (1984b).
- Hariharan P., B. F. Oreb, and Z. Wanzhi, "Digital Radial Shearing Interferometry: Testing Mirrors with a Central Hole," *Opt. Acta*, **33**, 251 (1986).
- Hariharan P., "Lateral and Radial Shearing Interferometers: A Comparison," *Appl. Opt.*, **27**, 3594 (1988).
- Honda T., J. Huang, J. Tsujiuchi, and J. C. Wyant, "Shape Measurement of Deep Aspheric Optical Surfaces by Radial Shear Interferometry," *Proc. SPIE*, **813**, 351 (1987).

- Horton R. F., "Design of a White Light Radial Shear Interferometer for Segmented Mirror Control," *Opt. Eng.*, **27**, 1063 (1988).
- Kohler D. R. and V. L. Gamiz, "Interferogram Reduction for Radial-Shear and Local-Reference-Holographic Interferograms," *Appl. Opt.*, **25**, 10, (1986).
- Kohn T., D. Matsumoto, T. Yazawa, and Y. Uda, "Radial Shearing Interferometer for In-Process Measurement of Diamond Turning," *Opt. Eng.*, **39**, 2696–2699 (2000).
- Kösters W., "Interferenzdoppelprisma für Messwecke," German Patent 595211, 1934.
- Li D., C. Huaixin, and Z. Chen, "Simple Algorithms of Wavefront Reconstruction for Cyclic Radial Shearing Interferometer," *Opt. Eng.*, **41**, 1893–1898 (2002).
- Malacara D., "Mathematical Interpretation of Radial Shearing Interferometers," *Appl. Opt.*, **13**, 1781 (1974).
- Malacara D., A. Cornejo, and M. V. R. K. Murty, "A Shearing Interferometer for Convergent or Divergent Beams," *Bol. Inst. Tonantzintla*, **1**, 233 (1975).
- Murty M. V. R. K., "A Compact Radial Shearing Interferometer Based on the Law of Refraction," *Appl. Opt.*, **3**, 853, (1964a).
- Murty M. V. R. K., "Some Modifications of the Jamin Interferometer Useful in Optical Testing," *Appl. Opt.*, **4**, 535 (1964b).
- Murty M. V. R. K., "Interference between Wavefronts Rotated or Reversed with Respect to Each Other and Its Relation to Spatial Coherence," *J. Opt. Soc. Am.*, **54**, 1187 (1964c).
- Murty M. V. R. K., "Interferometry Applied to Testing to Optics," *Bull. Opt. Soc. India*, **1**, 29 (1967).
- Murty M. V. R. K. and E. C. Hagerott, "Rotational Shearing Interferometry," *Appl. Opt.*, **5**, 615 (1966).
- Murty M. V. R. K. and R. P. Shukla, "Radial Shearing Interferometers Using a Laser Source," *Appl. Opt.*, **12**, 2765 (1973).
- Murty M. V. R. K., R. P. Shukla, and A. Cornejo, "Aberration in a Radial Shearing Interferometer Using a Laser Source," *Indian J. Pure Appl. Phys.*, **13**, 384 (1975).
- Parthiban V., C. Joenathan, and R. S. Sirohi, "Inversion and Folding Shear Interferometers Using Holographic Optical Elements," *Proc. SPIE*, **813**, 211 (1987).
- Parthiban V., C. Joenathan, and R. S. Sirohi, "Simple Inverting Interferometer With Holoelements," *Appl. Opt.*, **27**, 1913 (1988).
- Puntambekar P. N. and D. Sen, "A Simple Inverting Interferometer," *Opt. Acta*, **18**, 719 (1971).
- Roddier C., F. Roddier, and J. Demarcq, "Compact Rotational Shearing Interferometer for Astronomical Application," *Opt. Eng.*, **28**, 66 (1989).
- Roddier F., C. Roddier, and J. Demarcq, "A Rotation Shearing Interferometer with Phase-Compensated Roof Prisms," *J. Optics (Paris)*, **9**, 145 (1978).
- Saunders J. B., "Inverting Interferometer," *J. Opt. Soc. Am.*, **45**, 133 (1955).
- Saunders, J. B., "Construction of a Kösters Double-Image Prism," *J. Res. Nat. Bur. Stand.*, **58**, 21 (1957).
- Saunders J. B., in: *Optics in Metrology*, P. Mollet, Ed., Pergamon, Oxford, 1960, p. 227.
- Saunders J. B., "Measurement of Wavefronts without a Reference Standard, 2: The Wavefront Reversing Interferometer," *J. Res. Nat. Bur. Stand.*, **66B**, 29 (1962).
- Sen D. and P. N. Puntambekar, "An Inverting Fizeau Interferometer," *Opt. Acta*, **12**, 137 (1965).

- Sen D. and P. N. Puntambekar, "Shearing Interferometers for Testing Corner Cubes and Right Angle Prisms," *Appl. Opt.*, **5**, 1009 (1966).
- Shukla R. P., M. Moghbel, and P. Venkateswarlu, "Compact In-Line Laser Radial Shear Interferometer," *Appl. Opt.*, **31**, 4125–4131 (1992).
- Smart R. N. and W. H. Steel, "Zone-Plate Radial-Shear Interferometers. A Study of Possible Configurations," *Opt. Acta*, **32**, 1475 (1985).
- Som S. C., "Theory of a Compact Radial Shearing Laser Interferometer," *Opt. Acta*, **17**, 107 (1970).
- Steel W. H., "A Radial Shear Interferometer for Testing Microscope Objectives," *J. Sci. Instrum.*, **42**, 102 (1965).
- Steel W. H., *Interferometry*, 2nd Ed., Cambridge University Press, New York, 1983, p. 90.
- Steel W. H., in: *Progress in Optics*, Vol. V, E. Wolf, Ed., North-Holland, Amsterdam, 1966, Chap. III.
- Steel W. H., "A Radial-Shear Interferometer for Use with a Laser Source," *Opt. Acta*, **17**, 721 (1970).
- Steel W. H., "A Simple Radial Shear Interferometer," *Opt. Commun.*, **14**, 108 (1975).
- Steel W. H. and Z. Wanzhi, "A Survey of Thick-Lens Radial-Shear Interferometers," *Opt. Acta*, **31**, 379, (1984).
- Strong J., *Concepts of Classical Optics*. W. H. Freeman, San Francisco, 1958, Appendix C.
- Waetzman E., "Interferenzmethode sur Untersuchung der Abbildungsfehler optischer Systeme," (Interference Method for Determination of Aberrations of Optical Systems). *Ann. Phys.*, **39**, 1042 (1912).
- Walmsley I. A. and D. Malacara, "First-Order Parameters for a General Two-Beam Interferometer," *Appl. Opt.*, **34**, 3571–3575 (1995).
- Wanzhi Z., "Reflecting Radial Shear Interferometer," *Opt. Commun.*, **49**, 83 (1984).
- Wanzhi Z., "Reflecting Radial-Shear Interferometers with an Air-Spaced System," *Opt. Commun.*, **53**, 74 (1985).

6

Multiple-Beam Interferometers

C. Roychoudhuri

6.1. BRIEF HISTORICAL INTRODUCTION

The historical origin of multiple-beam interference was found as early as 1836, when Airy derived an expression for the multiple-beam interference pattern that would be produced by a plane parallel plate. However, the idea remained unexploited since high-reflectance coatings were not available and uncoated glass plate has a reflectance of only 0.04. It can produce good visibility fringes but of a two-beam type in reflection; no recognizable fringes can be seen in transmission.¹ Then Fizeau (1862a, 1862b) devised his celebrated interferometer, which now bears his name (see Chapter 1). His invention led to the idea of studying surface topography by optical interferometer (Laurent, 1883). The interferometer, being formed by two uncoated glass plates, gives interference fringes of a two-beam (cosine) type in reflection that contour the surface topography. Then Boulouch (1893) revived Airy's (1836) derivation of multiple-beam interference and demonstrated that with increasing reflectance of the Fizeau surfaces the multiple-beam fringes, both in reflection and in transmission, become increasingly sharper.

A few years later Fabry and Perot (1897) recognized the potential of the interferometer that consists of two plane parallel surfaces of high-reflection coating with variable separation, now known as the Fabry–Perot interferometer. Today it is still one of the most compact and highest resolving power spectrometers (Cook, 1971). To exploit this capability for high resolving power, the Fabry–Perot plates were usually used with a large separation that obscured the great potential of this instrument for mapping surface microtopography with ultrahigh local precision. Then in 1913, Benoit, Fabry, and Perot used a Fizeau interferometric arrangement with coated surfaces for their determination of the standard meter. They also missed the “optimum conditions” under which this multiple-beam Fizeau interferometer could be used for precision surface testing. These were provided and demonstrated

¹The relative strengths of the first three reflected beams are 0.04, 0.037, and 0.000059, and those of the transmitted beams are 0.92, 0.0015, and 0.0000024.

by Tolansky (1944), although he himself noted (Tolansky 1948a, pp. 7, 184) from a private communication sent to him from Williams that Adams Hilger Ltd. of England had been using multiple-beam Fizeau fringes for optical flat testing for some years. A lucid and detailed analysis and application of a multiple-beam interferometer of the Fizeau type can be found in Tolansky's books (1948a, 1948b, 1948c, 1960, 1966, 1968). Section 6.3 summarizes the essentials of the multiple-beam Fizeau interferometer.

To obtain sharp fringes as well as localized precision with the multiple-beam Fizeau, the proper conditions of Tolansky (the following section) require a very small separation and wedge angle between the mirrors. This mandates a high-quality test flat to give only a few narrow fringes of $\lambda/2$ interval² in the entire field of view, leaving most of the surface without any topographic information. The next year, Tolansky (1945a, 1945b) developed the so-called fringes of equal chromatic order (FECO) interferometry, where the illumination is by white light, rather than by quasimonochromatic light, and one observes small sections of the test surface successively through a prism spectrometer that displays fringes of equal chromatic order. The advantages of FECO over multiple-beam Fizeau are the greater precision of the former and its capability for distinguishing "hills" from "valleys" in a straightforward manner from the direction of bending of the colored fringes. Again, a good expression of the field is given in Tolansky's book (1948a). Contributions by Koehler (Koehler, 1953, 1955a, 1955b; Koehler and Eberstein, 1953; Koehler and White, 1955) are also worth mentioning. We describe FECO interferometry briefly in Section 6.4, where we also mention a somewhat different but very useful technique of illuminating the Fizeau interferometer by two different wavelengths that can be chosen suitably from two different calibrated monochromators (Shaan and Little 1975).

Later, several persons introduced various techniques to reduce the fringe interval of $\lambda/2$ so that microtopographic information can be obtained from otherwise fringe-free regions, using monochromatic or quasimonochromatic illumination. The first of these techniques, by Saunders (1951), introduces optical path differences into the interference film (formed by the two mirrors) by changing the air pressure by controlled amounts and taking a multiple photographic exposure of the fringe systems. Another technique to reduce the contour interval is to use a suitably chosen discrete set of wavelengths for illumination (Herroitt, 1961; Schwider, 1968; Pilston and Steinberg, 1969). A somewhat simpler technique by Murty (1962) employs regular quasimonochromatic illumination but a spatially separated set of pinholes. These techniques for reducing the fringe interval are described briefly in Section 6.5.

Some parallel development has also taken place in using the conventional Fabry–Perot interferometer with the plates parallel (rather than wedged) for deriving surface microtopography (Benedetti-Michelangeli, 1968; Hodgkinson, 1969) and for precision measurement of thin-film thickness (Schulz 1950a, 1950b). See Section 6.6 for methods of using a Fabry–Perot interferometer for testing.

²In some multiple-pass interferometers, the interval between consecutive fringes is less than $\lambda/2$; see Chapter 7.

A somewhat different set of multiple-beam fringes, produced by a Fabry–Perot illuminated by a point source but lacking the conventional fringe-focusing lens after the Fabry–Perot plates (see Figs. 6.2 and 6.17), was first observed and used by Tolansky (1943, 1946). The utility of these “Tolansky fringes” is described in Section 6.7.

The development of the spherical Fabry–Perot interferometer (Connes, 1958) of general curved mirror laser cavities (Boyd and Gordon, 1961; Fox and Li, 1961) and of the scanning spherical Fabry–Perot (Herroitt, 1963) paved the way for an elegant technique for testing curved surfaces with multiple-beam fringes preserving the localized precision (Herroitt, 1966).

A similar spherical wave interferometer was also developed independently by Perkin-Elmer of the United States (Heintze et al., 1967) and SIRA (1967) of the United Kingdom (Biddles, 1969). A strong desire to test many spherical surfaces by multiple-beam fringes against a single surface or a few master surfaces, coupled with the existing knowledge of Williams’ (1950) interferometer, paved the way for this development. We describe such multiple-beam interferometers with curved surfaces in Section 6.8.

Other developments worthy of mention in the field of multiple-beam interferometry are the so-called dual interferometry and holographic multiple-beam interferometry. In the former, one combines the multiple-beam fringes with another wavefront in a dual interferometric setup to utilize a live Moiré technique or to control the background contrast to advantage, especially while testing opaque surfaces in reflection by multiple-beam Fizeau (Langenbeck, 1968; Pastor and Lee, 1968). This is described in Section 6.9. The latter development can be exploited when one has already chosen to use holographic interferometry by modifying the system from two-beam to multiple-beam interferometry, as suggested by Matsumoto (1969) and Bryngdahl (1969). This is mentioned briefly in Section 6.10. For a detailed exposition of Moiré and holographic interferometry see Chapter 12.

Readers with broader interest on interferometry should consult the following review articles: Kuhn (1951), Baird (1967), Baird and Hanes (1967), Koppelman (1969), Briers (1972), Malacara et al. (1975), and Vrubel and Brown (1975). For a good review on precision interferometric testing, see Schulz and Schwider (1976) and Hariharan (1991). Readers interested in Fabry–Perot interferometer as a high-resolution spectrometer should consult the books by Hernandez (1988), Vaughan (1989) and Hillebrands (1999).

6.2. PRECISION IN MULTIPLE-BEAM INTERFEROMETRY

All the preceding chapters of this book dealing with various interferometric tests of optical components have one thing in common: The final fringe pattern, contouring the surface or the wavefront under test is formed by interference between two wavefronts. Then the recorded intensity variation follows the $\cos^2 \phi$ or $(1 + \cos 2\phi)$ type curve shown in Figure 6.1(a). Such fringes are said to have the so-called fringe quality, the finesse number, equal to 2. This is understood from the definition of finesse, which is the ratio of the fringe interval to the width of the fringe

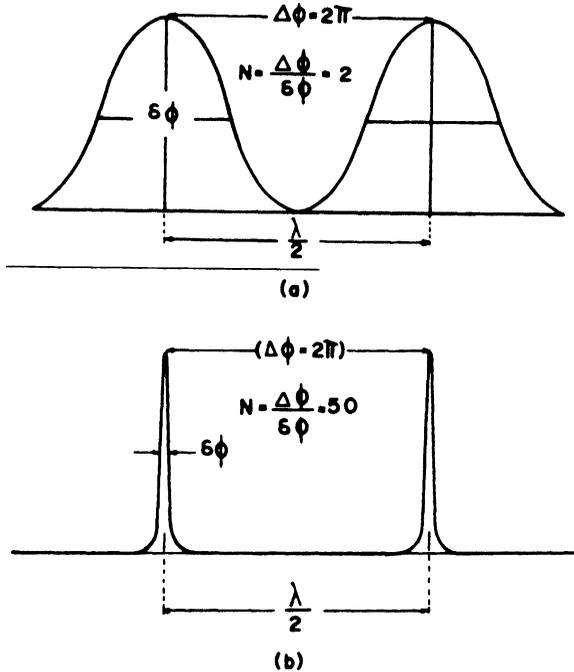


FIGURE 6.1. The finesse, the ratio of the fringe interval to the full width of the fringe at half its height, (a) for two-beam cosine fringes and (b) for multiple-beam Fabry–Perot fringes. For example, the finesse of (b) is 50.

at half its height (Fig. 6.1). In general, the fringe interval being $\lambda/2$, visual observation of two-beam cosine fringes can rarely achieve an error estimation of precision better than $\lambda/20$.

In contrast, multiple beam fringes are extremely sharp and can have finesse as high as 200 to 300 for commercial Fabry–Perot interferometers (FP) and many tens of thousands for specialized FP’s with mirrors having super flatness and super high reflectance. Very large super finesse FP’s are now used for trace gas detection [O’Keefe and Deacon, 1988] and gravitational wave detection [LIGO, 2005]. Micro super finesse cavities are used in studying cavity quantum electrodynamics [Kimble, 1994; Blais et al., 2004]. And the mirrors for such super finesse FP can be tested only as apart of a super finesse FP in the production facility [Herbelin et al., 1980; Itoh et al., 2001].

It must be realized, however, that to obtain multiple-beam fringes the interferometer plates must be coated with high-reflectance coating. This unquestionably interrupts the polishing process for a period of time, in contrast to two-beam interferometry, in which the surfaces under test can be simply cleaned and tested directly as frequently as the polisher wishes. For this reason multiple-beam interferometric tests are applicable chiefly to the field of thin-film technology (Bennett and Bennett, 1967; Eastman, 1975). The technique is also applied with

relative ease (say, instead of electron microscopy) for special precision testing of the surface roughness of high-quality optical surfaces (Koehler, 1955b; Koehler and White, 1955; Hodgkinson, 1970), but here one must be careful to choose a surface coating material (usually silver is selected; see Tolansky, 1960) that contours the substrate as closely as possible and also to analyze the phase change on reflection from the coating substance (Tolansky, 1948a; Koehler, 1953). These problems usually do not arise in two-beam interferometry. Multiple-beam interferometry is not advisable for routine shop testing unless the surface to be tested is definitely better than $\lambda/20$.

Of course, multiple-beam interferometry is not the only method for obtaining high spatial precision. If fringe sharpening is the main criterion, one can simply image the two-beam Fizeau fringes on a high-resolution vidicon camera and observe the fringes on closed circuit television. The fringes can be artificially sharpened through electronic control of the contrast. The precision by simple visual observation can certainly exceed $\lambda/50$, and an on-line computer analysis may give ultrahigh precision. Fringe sharpening can also be achieved through nonlinear photography. Precision densitometric traces of a regular photographic record of two-beam fringes or a direct photoelectric scanning of the fringe field, when suitably analyzed, can also give very high precision (Rosier, 1962; Roesler and Traub, 1966; Dew 1964, 1966). A more recent technique, oscillating mirror interferometry (see Chapter 13) with a simple two-beam instrument but with an electronic analyzing system, can also achieve a precision of $\lambda/1000$ (Raymond, 1970; Bruce and Sharpless, 1975). In fact, Moss et al. (1971) and Logan (1973) stated that their two-beam heterodyne (Michelson) interferometry, used in a gravitational wave detection system, can detect a displacement of a mirror with a precision of 10^{-6} Å. A somewhat different technique of detecting an optical path variation down to 10^{-5} Å through the measurement of the beat signal from a three-longitudinal-mode three-mirror laser cavity has also been reported in the literature (Boersch et al., 1974).

Here one should remember that the overall thermal and mechanical stability of the interferometer assembly must be higher by an order of magnitude than the precision expected from the system, unless the very purpose of the testing is to measure the relative "instability" (Dyson, 1968). The ultimate precision in interferometry is limited by the noise inherent in photoelectric detection (Hanes, 1959, 1963; Hill and Bruce, 1962; Raymond, 1970) and also by the diffraction phenomena that lead to the "optical uncertainty principle," analogous to Heisenberg's uncertainty principle (Heisenberg, 1949), discussed by Tolansky and Emara (1955), Thornton (1957), Koppelman (1966), and Lang and Scott (1968). The very high resolution, on the order of an angstrom or a fraction thereof, that is obtained through multiple-beam interferometry is in the longitudinal, not in the lateral direction. The resolution in the lateral direction is determined by the wavelength of the radiation due to diffraction.

The choice of multiple-beam interferometry should be guided by the following considerations: (a) the time available for the test, (b) the maximum precision of the test that is really necessary for the particular job the test surface is designed for, and (c) the equipment available in the laboratory.

6.3. MULTIPLE-BEAM FIZEAU INTERFEROMETER

6.3.1. Conditions for Fringe Formation

A multiple-beam Fizeau interferometer constitutes a very thin wedge-shaped film formed by two highly reflecting flat surfaces. Any incident beam falling on this film will, in general, produce a series of increasingly diverging beams (Fig. 6.3). These multiply reflected beams cannot be superposed in any plane with exactness. However, within certain approximations multiple-beam fringes are formed. Because the necessary conditions are more easily appreciated after following the process of formation of the ideal multiple-beam fringe by a Fabry–Perot interferometer, we shall briefly digress to describe fringe formation by this instrument.

The Fabry-Perot interferometer consists of two perfectly plane and parallel mirrors (Fig. 6.2). Naturally, a single incident wave will produce a series of waves by multiple reflection, and all of them can be superposed to form ideal Airy fringes (Tolansky, 1948a) at the focal plane of a lens.

We are assuming that the incident wave is continuous with a single carrier frequency ν . The intensity reflectance is R and the transmittance is T . The mirror spacing is d and hence the round trip phase delay, $\varphi = (2\pi/\lambda)2d = 2\pi\nu\tau$, τ being the round trip time delay. Then the normalized amplitude transmittance $i_{cw}(\nu, \tau)$ is given by

$$i_{cw}(\nu, \tau) = \sum_{n=0}^{\infty} TR^n e^{in\varphi} = \sum_{n=0}^{\infty} TR^n e^{in2\pi\nu\tau} = T/(1 - Re^{i2\pi\nu\tau}) \quad (6.1)$$

Then the normalized irradiance is given by the square modulus of Eq. (6.1),

$$I_{cw}(\nu, \tau) = T^2 / [(1 - R)^2 + 4R \sin^2 \pi\nu\tau] \quad (6.2)$$

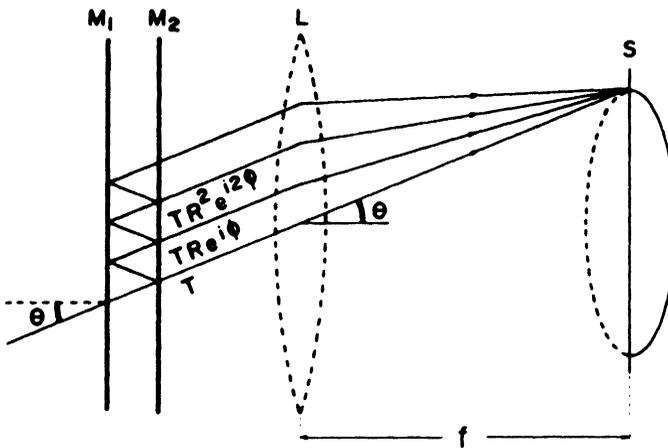


FIGURE 6.2. Fringe formation by a Fabry–Perot interferometer consisting of a pair of plane parallel mirrors M_1, M_2 followed by a fringe focusing lens L , S is the fringe plane in the focal plane of L .

The ideal reflective finesse N for these fringes (neglecting mirror aberrations and misalignments) is given by

$$N = \pi\sqrt{R}/(1 - R) \quad (6.3)$$

where T and R are the intensity transmittance and reflectance, respectively, of both mirrors and ϕ is the total effective phase or optical path delay between any two consecutive wavefronts.

Although the series in Eq. (6.1) is infinite, in practice, the total effective number of interfering beams is finite because of the fact that the energy carried by the n th transmitted beam, T^2R^{2n} , becomes negligible for sufficiently large n . [See the solid curves of Figure 6.4, which indicates the number of superposed beams required to form a multiple-beam fringe, for a particular reflectance, within 1% of the ideal (infinite sum) of the Airy curve.] This is a useful practical point to remember for any type of multiple-beam interferometry, as we shall soon see. If it were really necessary in practice to superpose the infinite sequence of multiply reflected beams to form the ideal Airy fringe, the interferometrist would have to wait throughout eternity to observe it (Roychoudhuri, 1975)! For recent developments in time domain use of FP for testing and characterizing super finesse Fabry–Perot mirror see Section 6.11.

The next point of practical interest in multiple-beam Fabry–Perot fringe formation is that the phase delay ϕ between any two consecutive wavefronts is constant. Then, once ϕ has been determined, the entire series of multiply reflected beams combines to form the appropriate part of the ideal Airy fringe. Thus, when ϕ is either $2n\pi$ or $(2n + 1)\pi$ (n any integer), the entire series of beams, when superposed, either combines constructively to form a bright fringe or adds destructively to form a dark fringe.

This is where the major point of departure appears between a plane parallel Fabry–Perot and a wedge-shaped Fizeau interferometer. The phase delay between the consecutive wavefronts produced by the wedged Fizeau mirrors is a progressively increasing quantity, rather than a constant one as in a parallel Fabry–Perot. The other difference is the spatial walk-off of the beams with multiple reflection (in Fizeau interferometry) that cannot be compensated for perfectly by any focusing or imaging device. Both these effects are displayed in Figure 6.3 (see also Rogers, 1982).

The progressive phase delay is easily derived by using Brossel's (1947) very general but elegant method developed to compute the intensity distribution and localization of Fizeau fringes (see also Born and Wolf, 1975, p. 286). We shall consider the particular case of perpendicular illumination of one of the mirrors by a collimated beam that is used in many optical shops (Fig. 6.3 and 6.6). Two partially transmitting mirrors M_1 and M_2 form a very small wedge angle ϵ , whose apex is at O . An incident plane wavefront, parallel to the mirror M_2 , produces a series of beams due to multiple reflection. The arrow heads in Figure 6.3(a) correspond to the "center" of the incident wavefront and hence demonstrate the beam walk-off defect of the Fizeau interferometer. Figure 6.3(b) shows the position of the multiply reflected wavefronts relative to the incident one at the plane of the mirror M_2 (Y plane). The choice of such a diagram to compute the phase delay (Brossel, 1947;

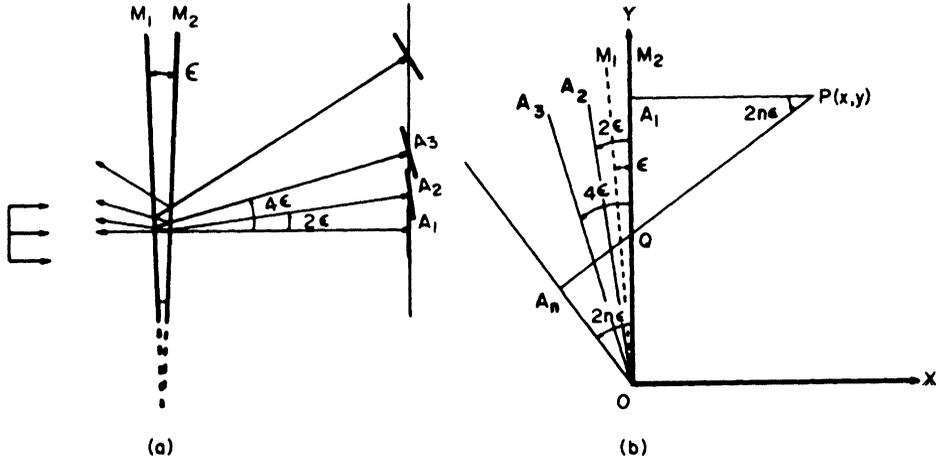


FIGURE 6.3. (a) Multiply reflected beams produced by a Fizeau interferometer consisting of a pair of “plane” mirrors M_1, M_2 with a wedge angle ϵ between them when illuminated by a collimated wavefront parallel to M_2 . (b) Geometrical construction to aid in the computation of the relative path difference between multiply reflected beams produced by a Fizeau interferometer.

Tolansky, 1948a) was based on the fact that none of the multiply reflected wavefronts suffers any phase delay relative to the incident one (OA_1) at the apex of the wedge O , formed by the two Fizeau mirrors. The tilts of the multiply reflected wavefronts are designed by lines OA_2 ($\langle 2\epsilon$), OA_3 ($\langle 4\epsilon$), \dots , OA_n ($\langle 2n\epsilon$). If the effect of superposition is observed in transmission at a general point $P(x, y)$, the path difference between the first and the n th wavefronts is

$$d_n(x, y) = PA_n - PA_1 \tag{6.4}$$

But

$$\begin{aligned} PA_n &= PQ + QA_n = PQ + OQ \sin 2n\epsilon \\ &= PQ + (y - QA_1) \sin 2n\epsilon = x \cos 2n\epsilon + y \sin 2n\epsilon \end{aligned}$$

Then

$$d_n(x, y) = x(\cos 2n\epsilon - 1) + y \sin 2n\epsilon \tag{6.5}$$

If the fringes are observed on mirror surface M_2 ($x = 0$), as is customarily done in optical shops through an imaging device, the expression for the phase difference simplifies to

$$d_n(0, y) = 2tn - \frac{4}{3}n^3\epsilon^2t \tag{6.6}$$

where we have used $t = y\epsilon$, the separation between the plates at y and the first two terms of the series expansion of $\sin 2n\epsilon$, considering $2n\epsilon$ to be still a small angle. For a parallel plate Fabry–Perot with orthogonal collimated illumination, the phase difference between the first and n th wavefronts is $2tn$. Then the n th beam of a Fizeau interferometer lags behind in phase, compared to $2tn$, by

$$\delta d_n = \frac{4}{3}n^3\epsilon^2t. \quad (6.7)$$

This lesser delay in the Fizeau than in the Fabry–Perot arithmetic series ($2tn$) is the single most important restriction against forming high-quality symmetric and sharp fringes, as was first realized by Tolansky. (This problem is avoided in white-light FECO interferometry, described in the following section, where the mirrors are set as nearly parallel as possible.) In particular, if

$$\frac{4}{3}n^3\epsilon^2t = \frac{\lambda}{2} \quad (6.8)$$

the n th beam will add destructively rather than constructively to the first one. Hence to obtain symmetric and sharp Fizeau fringes, one should restrict one-self to the Tolansky inequality,

$$\frac{4}{3}n^3\epsilon^2t < \frac{\lambda}{2}. \quad (6.9)$$

Here the most rapidly varying quantity is n , the number of interfering beams. But we cannot set a low value for it because the very purpose of a multiple-beam Fizeau is to produce high-quality, high-finesse fringes that require superposition of a large number of regularly delayed beams; this number, in turn, depends on the reflectance of the surfaces. The broken curve in Figure 6.4 shows how the finesse N_R (the ratio of the fringe interval to the full width of a fringe at half the height from its peak) of Fabry–Perot fringes increases with increasing reflectance. The solid curve in the same figure shows the effective number of beams M added by a computer to obtain a peak transmittance within 0.1 % of the ideal Airy value [infinite sum of Eq. (6.1)] for different reflectances. It can be seen that for a reflectance below 90%, the effective number of beams M that are superposed to obtain a finesse N_R is roughly equal to $2N_R$. For higher reflectivity, the required value of M steadily increases. See also Eq. (6.5) and the rationale given there.

These curves can be exploited for Fizeau interferometry in the following manner. If the surfaces to be tested have a given reflectance R (or equivalently, $R = \sqrt{R_1R_2}$ when the reflectances are different for the surfaces), or one requires a desired finesse N_R such that the reflectance is R , then the curves of Figure 6.4 indicate the value of M , the effective number of beams that are superposed. Then to maintain the Tolansky inequality, one substitutes M for n in Eq. (6.9) and reduces the values of ϵ and t accordingly. Tolansky (1948a) gave some typical values of n , ϵ , and t for obtaining symmetric Fizeau fringes. The lowest limit of t , the plate separation, is usually

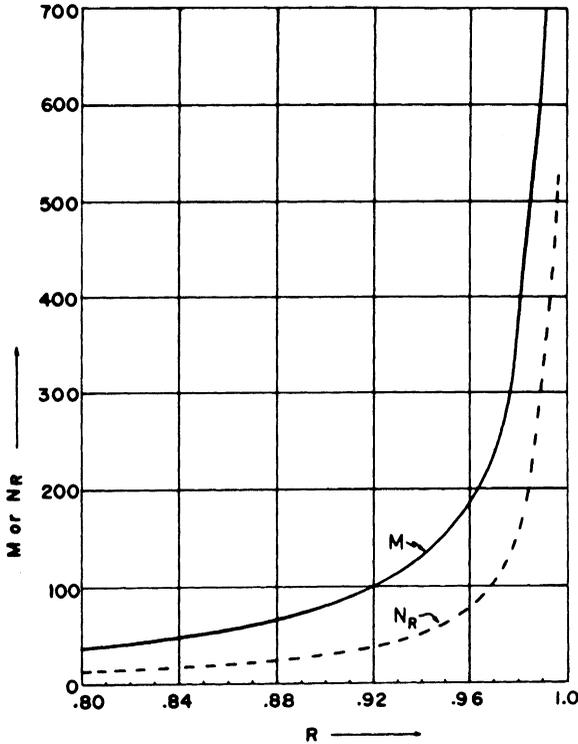


FIGURE 6.4. The number of interfering beams M that gives a value for the ideal Airy curve within 0.1% plotted against the reflectance R (continuous curve.) The discontinuous curve shows the reflective finesse N_R plotted against the reflectance R .

determined by omnipresent dust particles whose average size is generally of the order of $1\ \mu\text{m}$. The lowest limit for the wedge angle ϵ is determined by the minimum number of fringes one wants within the field of view. The smaller the angle the larger is the spatial separation between two adjacent fringes.

When the Tolansky inequality [Eq. (6.9)] is not satisfied (in other words, under the most general condition), the intensity profile of each Fizeau fringe at high reflectance is an asymmetric and composite one and constitutes a primary maximum followed by a series of secondary maxima on the side of the thicker wedge. The intensity of these secondary maxima rapidly decreases in the same direction as the increasing wedge separation. The primary maximum is broadened and is no longer symmetric like the Fabry–Perot fringes; also, its peak shifts toward the opening of the wedge with reduced height (Kinosita, 1953). Figure 6.5 shows a qualitative representation of a general Fizeau fringe profile plotted against one due to Fabry–Perot. These general Fizeau fringes show a remarkable similarity to Tolansky fringes (Fig. 6.17) produced through a plane parallel Fabry–Perot but illuminated by various tilted rays from a point source. Such composite fringes will not destroy precision testing if the smooth edge of the fringes is used (Polster, 1969). Characteristic Fizeau fringes are formed

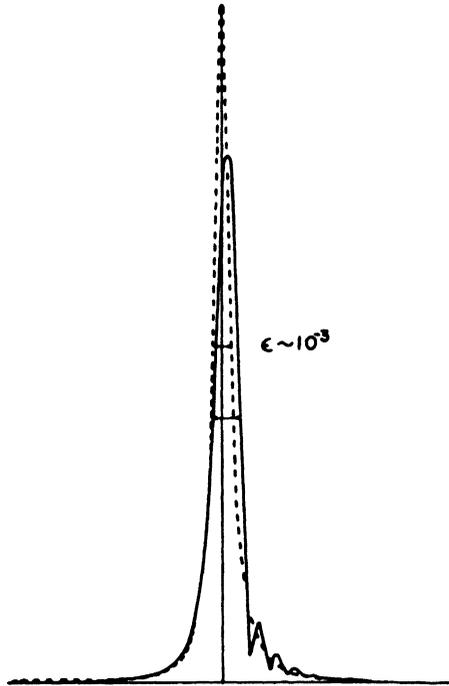


FIGURE 6.5. A composite Fizeau fringe (continuous curve) produced by a pair of flat wedged ($\sim 10^{-3}$ radian) mirrors compared with a symmetric Fabry–Perot fringe (discontinuous curve) produced by the same mirrors when they are perfectly parallel.

not only on the plane of the thin wedged film but also at various well-defined distances on planes named after Feussner (Barakat et al., 1965).

There is another important constraining parameter of a Fizeau interferometer: the lateral displacement Δ_n of the n th beam due to the wedge angle ϵ (Tolansky, 1948a),

$$\Delta_n = 2n^2t\epsilon. \quad (6.10)$$

This implies that each multiple-beam fringe formed by a Fizeau interferometer always gives the contour map of the average surface topography over a region spanned by Δ_n , rather than a precise geometrical point-by-point mapping.

6.3.2. Fizeau Interferometry

The essential components of a Fizeau interferometer arrangement are shown in Figure 6.6. A monochromatic or quasimonochromatic collimated beam illuminates the coated flats (M_1 , M_2) forming a wedge. The fringes can be observed both in transmission (O_T) and in reflection (O_R). The fringes in transmission are sharp and bright, in very high contrast against an almost dark background, and those in

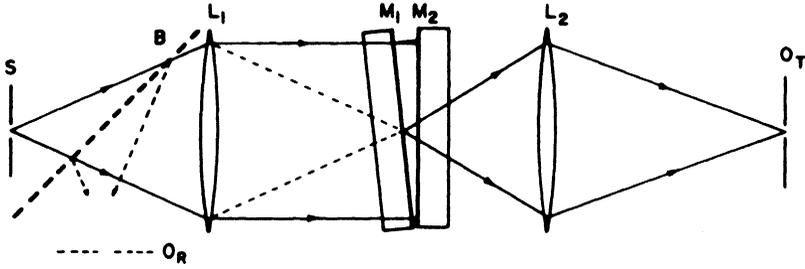


FIGURE 6.6. The essential components of Fizeau interferometric arrangement: S, point source. B, beam splitter to observe fringes in reflection (O_R); L_1 , collimator; M_1 , M_2 , Fizeau mirrors, the wedge and separation very much exaggerated; L_2 , observing lens, which images the localized fringes between the mirrors at the observation plane O_T .

reflection are exactly complementary (by simple energy conservation), that is, they are dark fringes against a bright background (Fig. 6.7). The fringes in reflection, however, may have very poor contrast when the reflection coating (such as a silver layer) has a very high absorption coefficient (Tolansky, 1948a). All Fizeau fringes have the same fringe interval of $\lambda/2$, just as do the multiple-beam Fabry–Perot or regular two-beam fringes. This is the reason why sharp Fizeau fringes leave most of the surface area under test without much available information. Alternatives to this arrangement are discussed in the following sections.

Since the surface contour fringes show the total optical path variation of both flats, one of them should be a very high quality reference flat so that the fringes can be interpreted as contours only because of the flat being tested (Clapham and Dew, 1967). One way of obtaining an extremely high quality flat is to use a liquid surface such as liquid mercury. A detailed description of such a Fizeau interferometer has been given by Bünnagel et al. (1968). However, deviation of a surface from absolute planeness can also be obtained without the use of an absolute reference flat. The method (Schulz, 1967; Schulz and Schwider, 1967; Schulz et al., 1971) requires the use of three flats, none of which is a high-quality reference flat.

Because of the Tolansky inequality [Eq. (6.9)], good-quality Fizeau fringes are obtained more easily if the separation between the plates, t , is set to a very small value. However, this is not an absolute necessity, as has been demonstrated by Moos et al. (1963), who used a highly collimated laser beam to obtain the surface contours with a plate separation as large as 20 cm. Of course, the tilt must be very small; in their case, it was less than 10^{-4} rad. This point is worth remembering for shopwork because a large separation ensures against developing scratches or spoiling the reflection coating by physical contact. A large separation requires a high-quality mounting to maintain a stable relative alignment.

In an interesting example of off-axis illumination for improving fringe sharpening, described by Langenbeck (1970), the incident angle is chosen in such a way that the incident beam is first reflected toward the apex of the Fizeau plates and then, after a controllable finite number of reflections, it is reflected away from the apex. This can

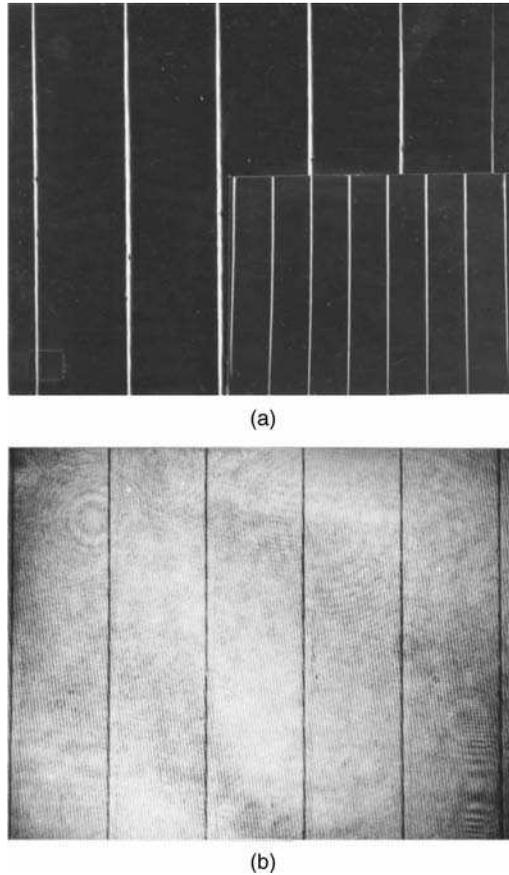


FIGURE 6.7. Fizeau fringes in transmission and in reflection, (a) Fringes in transmission constitute sharp, bright lines against a dark background. Insert, the same transmission fringes but without the collimating lens L_1 of Figure 6.6. (b) Fringes in reflection constitute complementary fringes of dark lines against a bright background. (Fine fringes are due to back reflection from the beam splitter.)

be appreciated from Figure 6.8, where if the illumination follows the direction of the ray SA_1 , after a few reflections it becomes perpendicular to mirror M_2 at A_n and starts retracing its path. In this manner, sharper Fizeau fringes can be formed in special situations where a relatively large separation between Fizeau mirrors is required by a particular test object, such as a corner cube (Langenbeck, 1970).

Eastman and Baumeister (1974) designed a regular Fizeau interferometer with one of the mirrors mounted on a piezoelectric scanning device. Instead of photographic recording, the fringes can be detected and analyzed electronically. The reported accuracy is about 20 \AA , but techniques to improve the measurement precision are probably available. Fizeau interferometry with illumination of two wavelengths (Shaalan and Little, 1975) is discussed briefly in the following section.

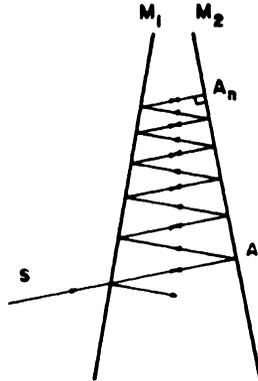


FIGURE 6.8. Fringe sharpening in a Fizeau interferometer. For a given wedge between the Fizeau mirrors M_1 , M_2 , a suitable choice of the direction of illumination SA_1 can reduce the range of beam walk-off and increase the effective number of interfering beams to sharpen the fringes.

6.4. FRINGES OF EQUAL CHROMATIC ORDER

In 1945, Tolansky developed a new technique of surface microtopography using white-light illumination. The interferometer has evolved from the multiple-beam Fizeau interferometer, where the fringes are formed by a thin, wedge-shaped film bounded by two highly reflecting surfaces. This wedge reduces the fringe sharpness (finesse) because of beam walk-off (Fig. 6.3). Therefore, Tolansky (1945a, 1945b) had set the surfaces as parallel as possible, increasing the inherent finesse. With this arrangement a collimated and orthogonal illumination with white light will produce a channeled spectrum (Jenkins and White, 1957) of all the wavelengths, λ_i , that satisfy the relation

$$2t = n_i \lambda_i. \quad (6.11)$$

However to see the spectrum, one needs to use a spectrograph for dispersion. When a narrow section of an ideal interference film (t constant) is imaged through a spectrograph, a series of laterally separated and parallel straight fringes (the channeled spectrum) that satisfy Eq. (6.11) is displayed. If there is a variation in t within the imaged section of the film, each point is passing a group of waves corresponding to the local value of t . The spectrographic image then consists of nonstraight “fringes,” each one of which shifts along the wavelength scale, keeping t/λ constant. As stated earlier, the name fringes of equal chromatic order (FECO) is used since along each fringe the order number n is constant.

The basic interferometric setup is illustrated in Figure 6.9. Figure 6.10(a) shows some FECO fringes (black and white reproduction) of a cleaved mica surface, the topographic variation of which is shown in Figure 6.10(b) (Tolansky, 1945a). Determination of the “hills” and “valleys” of a surface microtopography becomes very simple since fringes are convex to the violet on the “hills” and are concave to

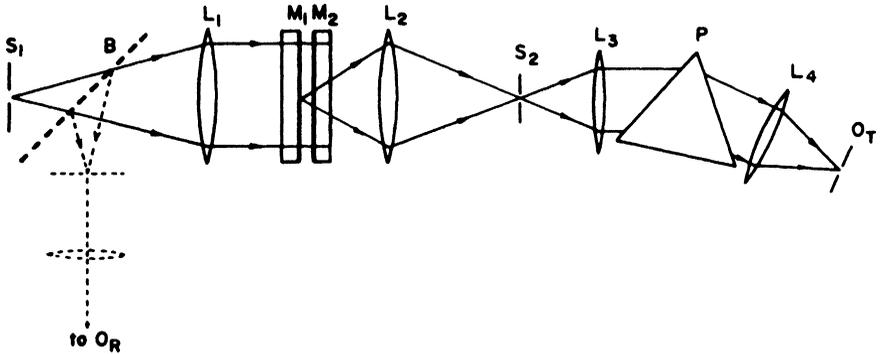


FIGURE 6.9. The essentials of the interferometric arrangements for observing fringes of equal chromatic order (FECO) either in transmission (O_T) or in reflection (O_R): S_1 , white-light point source; B , beam splitter to aid observation in reflection; L_1 , collimating lens; M_1, M_2 , parallel mirrors forming the interferometer, one of which is being compared to the other as a reference; L_2 , lens that images the channelled spectrum from a small section of M_1 – M_2 on the spectrographic slit S_2 (slit S_2 , lenses L_3 and L_4 , and prism P form the spectrograph); O_T , observation plane for fringes in transmission.

the violet on the “valleys.” The steepness of the incline at a local region of a surface is given by the number of fringes intersected per unit length of the vertical section of the fringes (Fig. 6.10). For a detailed exposition of FECO interferometry, the reader should refer to Tolansky (1948a).

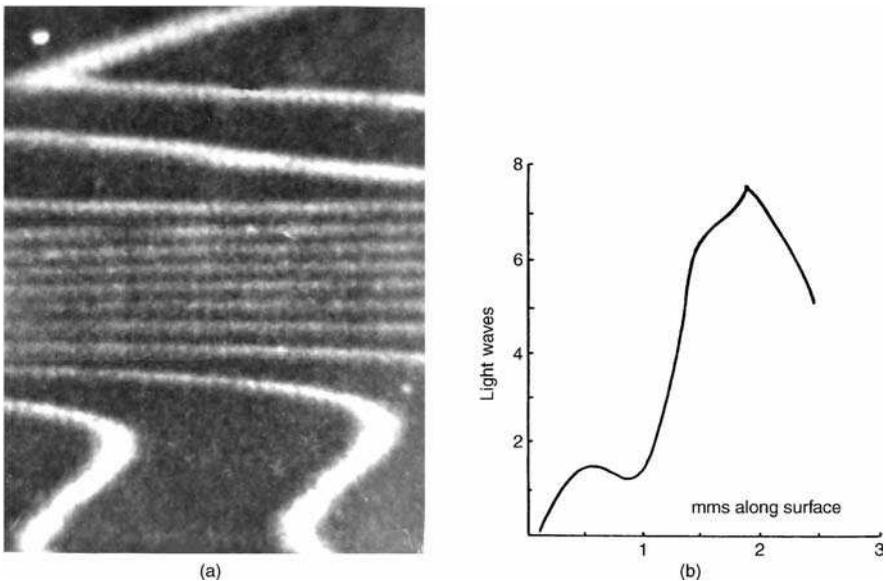


FIGURE 6.10. (a) Fringes of equal chromatic order (in black and white) from a freshly cleaved and silvered mica surface tested against a reference flat mirror. The dense fringes at the central region along the vertical direction indicate a sharp ridge, (b) A quantitative plot of the surface height variation along the vertical direction depicted by the fringes of (a). (From Tolansky, 1948a.)

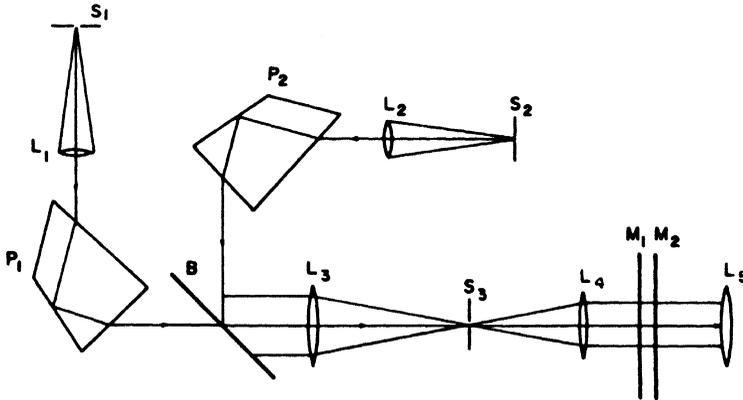


FIGURE 6.11. Fizeau interferometry with white light, using two calibrated monochromators: S_1 , L_1 , P_1 and S_2 , L_2 , P_2 , the two monochromators; B , beam splitter; L_3 , common spectrum-forming lens for both monochromators; S_3 , entrance slit for the Fizeau interferometer (M_1 – M_2); L_4 , collimator for the interferometer; L_5 , observing lens.

Fringes of equal chromatic order interferometry find a very useful application in the accurate determination of thin-film thickness and in surface roughness measurement (Bennett and Bennett, 1967; Eastman, 1975; Bennett, 1976). The steadily increasing demand for very high quality optical surfaces in modern optical technology has motivated the development of FECO interferometry for studying surface roughness and has led to a better understanding of the mechanism of polishing (Koehler and White, 1955; Koehler, 1955b; Vinokurov et al., 1962; Hodgkinson, 1970). One should be careful, however, to consider the problem of the phase change on reflection due to the material coated on the substrate (Tolansky, 1948a; Koehler, 1953; Schulz, 1951a) and also the dependence of the phase change on the thickness of the thin-film material (Schulz and Scheibner, 1950). It has been observed that the original surface microtopography is more faithfully contoured by a metallic silver film (Tolansky, 1948a) than by a dielectric film.

Following the lead of Tolansky, Shaalan and Little (1975) reported a different technique of exploiting white-light fringes for the study of surface micro-topography. To exploit the full advantage, one should use an illumination for the Fizeau interferometer that has two wavelengths simultaneously present from two calibrated monochromators (like constant deviation spectrometers). Figure 6.11 shows the experimental arrangement. With this technique, the direction and the height of crystal cleavage steps or of thin films can be measured with great facility. Any region of the surface under test can be scrutinized by first choosing one of the wavelengths to form a fringe in the region of interest, and then adjusting the second monochromator to produce a different colored but matched fringe on the other side of the step (Fig. 6.12). Since the two wavelengths are known, the direction and the height of the step are determined easily using Eq. (6.11); but one should take into account the phase change on reflection, which is not explicit in this equation (Shaalan and Little, 1975).

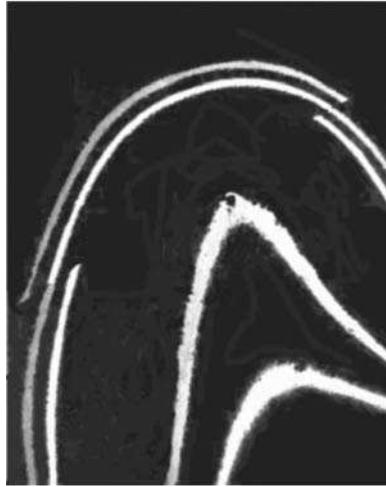


FIGURE 6.12. Multiple fringes of variable chromaticity meeting across a step. The lighter fringe corresponds to the red and the darker one to the green radiation as they appear in the original, which was in color. (From Shaalam and Little, 1975.)

6.5. REDUCTION OF FRINGE INTERVAL IN MULTIPLE-BEAM INTERFEROMETRY³

In most regular interferometry with monochromatic or quasimonochromatic illumination, the fringe interval is $\lambda/2$, the fringes being contours of equal optical thickness. This is the reason why in multiple-beam interferometry, the fringes being very sharp and of high contrast, most of the surface area under test produces almost no information about the surface topography. Even in the case of interferometry with white light, discussed in Section 6.4, only a small surface area is imaged on the spectrometer slit. One can scan the surface of the test object region by region. However, this procedure is somewhat tedious.

A simpler solution (Saunders, 1951), using “pressure scanning” with normal quasimonochromatic illumination, gives sharp Fizeau fringes but obtains a fringe interval less than $\lambda/2$. The basic technique requires placing the entire interferometer in an airtight chamber and taking a series of exposures of the fringe system, the fringes being shifted between every exposure by the desired amount. The shift is achieved by a controlled amount of change in the optical path through a change in air pressure. The same objective can be achieved a bit more conveniently by mounting one of the mirrors on piezoelectric scanning devices and applying a suitable staircase voltage to the piezoelectric (Roychoudhuri, 1974). This eliminates the necessity of putting the entire interferometer within an airtight system.

In a different solution demonstrated by Herriott (1961), one illuminates the interferometer simultaneously with a discrete set of wavelengths obtained through

³The fringe interval can also be decreased by multipass interferometry, which is discussed in Chapter 7.

a monochromator whose entrance illumination consists of a set of spatially separated slits, instead of a single slit. Later Murty (1962) suggested a simpler solution, whereby one can use a regular quasimonochromatic (or monochromatic) source with a set of spatially separated pinholes to illuminate the interferometer. A multiple set of fringes with a reduced fringe interval is provided. A simple computation to achieve the desired fringe interval has been given by Murty (1962). Schwider (1968) also used Herriott's idea of multiple-wavelength illumination but employed the channeled spectra from a suitable Fabry–Perot. The flexibility of the test is greatly increased if one uses a tunable laser source, as described by Pilston and Steinberg (1969). With the continuously tunable dye lasers currently available, such tests can be carried out very conveniently and rapidly.

6.6. PLANE PARALLEL FABRY–PEROT INTERFEROMETER

It is somewhat surprising that even though the capabilities of the plane parallel Fabry–Perot (1897) interferometer as a high-resolution spectroscopic instrument were well understood, the interferometer was not used for surface measurements until fairly recently (Schulz, 1950a, 1950b). The standard techniques that evolved for its utility as a spectroscopic instrument require (a) illumination by an extended source to obtain fringes of sufficient brightness and (b) a reasonable separation between the plates to provide sufficient resolving power. These two general conditions probably inhibited the use of this interferometer for surface topography because the fringes formed under these conditions cannot describe the local defects region by region; rather, every part of the fringes is characteristic of the overall defects of the entire surface (Chabbal, 1953, 1958). In the following discussion we describe briefly, in chronological order, a series of uses of the Fabry–Perot interferometer for surface testing.

6.6.1. Measurement of Thin-Film Thickness

The thickness of thin evaporated films can be measured to an accuracy of $\pm 15 \text{ \AA}$ (or better) by using conventional Fabry–Perot fringes (Schulz, 1950a, 1950b). The method exploits the fact that a simple sodium vapor lamp emits a close doublet (5890 and 5896 \AA) that forms, in general, a pair of closely spaced fringes and, in particular, consonance and dissonance of the pair of fringes, depending on the separation between the plates. Then a knowledge of the wavelengths and plate separations for suitable positions of the pair of fringes on the two different layers of the plate gives one the thickness of the thin film. The precise method of quantitative analysis has been described by Schulz (1950a, 1950b). The advantage of the method over the multiple-beam Fizeau is that the interferometer-forming surfaces do not have to be pressed together so closely as to endanger the high-quality surfaces and/or reflection coatings. [But here it should be remembered that Fizeau fringes can also be formed with large separation of the plates but with a very small wedge angle (Moos et al., 1963).]

Schulz (1951a, 1951b) extended this technique to measure the phase change on reflection from a thin film, its dependence on the wavelength and the thickness of the material, and also the absorption coefficient of the film.

6.6.2. Surface Deviation from Planeness

Benedetti-Michelangeli (1968) developed a method to measure the local defects of plane surfaces by using a narrow collimated laser beam to illuminate the local region of the Fabry-Perot interferometer formed by the test and the reference plate. The computation of the defect is carried out by using the standard relation for the diameter of the Fabry-Perot fringes with the interferometer parameters as explained in this paper. The method of illumination consists of a narrow collimated beam incident on the parallel plates at an angle that can be varied in a smooth manner. The reported accuracy of the flatness variation is $\lambda/400$, but, as has been rightly claimed, the potential limit is much higher.

Again, one advantage of this method is that the plates under test do not have to be pressed together so closely as to endanger the surfaces. [An important virtue of this short paper of Benedetti-Michelangeli (1968) is that it describes briefly but succinctly all the important factors that influence the fringe qualities of the Fabry-Perot]. A more accurate method is to exploit the steep slope of an Airy transmission curve and use a narrow collimated illumination that is orthogonal to the parallel plates. The regional defects are obtained by changing the position of the narrow collimated beam. An accuracy of $\lambda/1000$ can be obtained by this method (Koppelman and Krebs, 1961). A major defect of either of these methods is that the minimum area for which the average defect can be determined is limited by the spatial size of the collimated scanning beam, which can be scarcely smaller than a millimeter. Methods for determining point-by-point defects of surfaces are described in the following paragraphs.

It appears that Hodgkinson (1969) was the first one to exploit the full potentiality of the Fabry-Perot interferometer for the study of surface defects. Most multiple-beam interferometric tests for surface topography suffer from the general disadvantage that they do not record the complete topography of the entire surface at a time; rather, they sample it only along the narrow fringes, leaving most of the surface without any available information, as mentioned before. Techniques to reduce the fringe interval (Section 6.5) alleviate the problem only partially. Hodgkinson (1969) developed a method of recording an integrated interferogram with the transmitted wavefront by slowly moving one of the mirrors of the Fabry-Perot parallel to itself. The illuminating beam being an orthogonal collimated monochromatic radiation, the transmission of the interferometer at every point is proportional to the local plate separation. All the topographic information as well as the defect distribution function (Chabbal, 1958) can be obtained from the transmission characteristics of this integrated interferogram when it is properly developed (Hodgkinson, 1969). The exposure and the development are such as to produce a transmission characteristic that is almost linearly proportional to the surface defects. Because of the limits on the precision obtainable from photographic

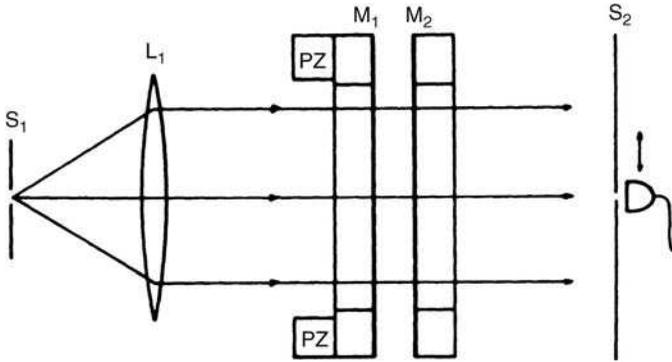


FIGURE 6.13. A Fabry–Perot interferometric arrangement for evaluating the surface defects of all the points of a pair of mirrors. The entire interferometer is illuminated by a collimated beam. Defects are determined from the change in transmission. One of the mirrors is mounted on piezoelectric material (PZ) to choose the desired value of transmission.

work, the measurement of surface defects by this method has a precision of around $\lambda/500$.

It is possible to dispense with the intermediate record of an integrated interferogram if the transmitted wavefront is sensed by a spatially scanning photo detector (or a high-resolution vidicon camera) and the information is stored for detailed analysis (Fig. 6.13). The reference mirror can be mounted with piezoelectrical scanning devices, and the separation between the mirrors adjusted precisely so that the detector records half the peak transmission, say, from a small central area. Then the transmission from this central spot can be used as the reference signal to measure the deviation from planeness of the other points. For a direct graphical computation, the transmission curve along any diameter of the test surface is compared against the “ideal” Airy curve, recorded through the reference central spot as an oscilloscopic trace while scanning one of the mirrors. This is illustrated in Figure 6.14. The half-width of the “ideal” Airy curve is given by

$$pp' = \frac{2\pi}{N_R} \equiv \frac{\lambda}{2N_R} \quad (6.12)$$

Also, the abscissa [Fig. 6.14(a)] is linear in phase, and the finesse N is known either by measurement from the actual Airy curve obtained from a small central reference region or from the value of the reflection coating (Fig. 6.15). Then any fluctuation in the transmission, AQ or BR , that corresponds to an optical path change of pq or pr , respectively, can be directly transformed into fractions of a wavelength using Eq. (6.12). The precision of the method depends on the minimum detectable signal like pq that is controlled by the slope of the Airy curve (Polster, 1969). This precision is limited by the noise inherent in the photodetection. Nevertheless, the measurement precision good to $\lambda/2300$ for surface flatness and roughness has been achieved by Itoh et al. (2001). This paper demonstrates the extreme precision achievable by using the principle of multiple-beam Fabry–Perot interferometer.

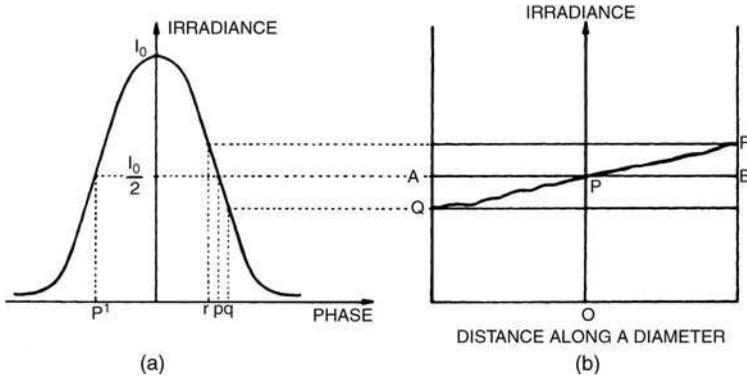


FIGURE 6.14. Determination of the surface defects every point of a pair of mirrors by Fabry-Perot interferometry. (a) Reference (or ideal) Airy transmission curve for a Fabry-Perot with calibrated axes, (b) Transmission curve along a particular diameter of the Fabry-Perot when the transmission from the small central reference region was adjusted (using the PZ) to exactly half the peak transmission.

We conclude this section by mentioning another advantage of Fabry-Perot interferometry. So far, we have discussed the problem of determining the positions and sizes of local defects of surfaces. However in most optical testing, the approximate overall (average) quality, λ/m , of the entire surface is also of interest. This is

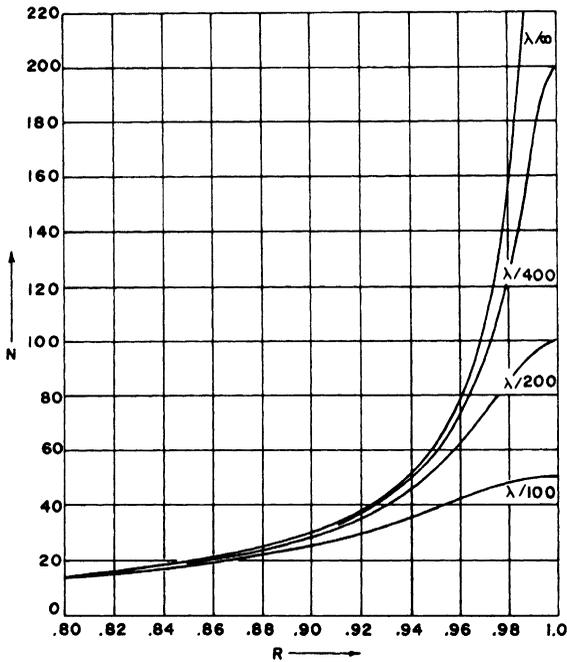


FIGURE 6.15. Effective surface finess N plotted against reflectance R for a Fabry-Perot with mirrors having a surface defect of residual spherical curvature. Curves for four different cases of surface flatness are shown.

most readily obtained by assembling a master plate and the plate under test to form a Fabry–Perot of known ideal reflective finesse and comparing this quantity with the experimental finesse given by the interferometer. Presented in Figures 6.15 and 6.16 are two sets of curves for effective finesse against surface reflectivity; the first one is for pure residual spherical curvature and the second one is for Gaussian random deviation from perfect flatness. These curves were computed (Roychoudhuri, 1973b) using Chabbal’s analysis (1958). The usefulness of the curves is illustrated by a simple example. Suppose that the reference flat and the flat under test have 98% reflectivity, giving an ideal reflective finesse close to 155, but the experimental finesse is, say, 80 then with a careful look at the curves of Figures 6.15 and 6.16, one can conclude that if the surface under test has a pure residual spherical curvature, the magnitude of the mean deviation from flatness is somewhat below $\lambda/200$; or if the deviation is of Gaussian nature or is a combination of both regular and Gaussian, the mean deviation must certainly be less than $\lambda/400$. Furthermore, one can use the value of the peak transmission (Chabbal, 1958; Jacquinot, 1960; Hodgkinson, 1969) or the nature of the broadened Fabry–Perot fringes (Hill 1963; Bhatnagar et al., 1974) to discern and characterize the surface deviation more precisely.

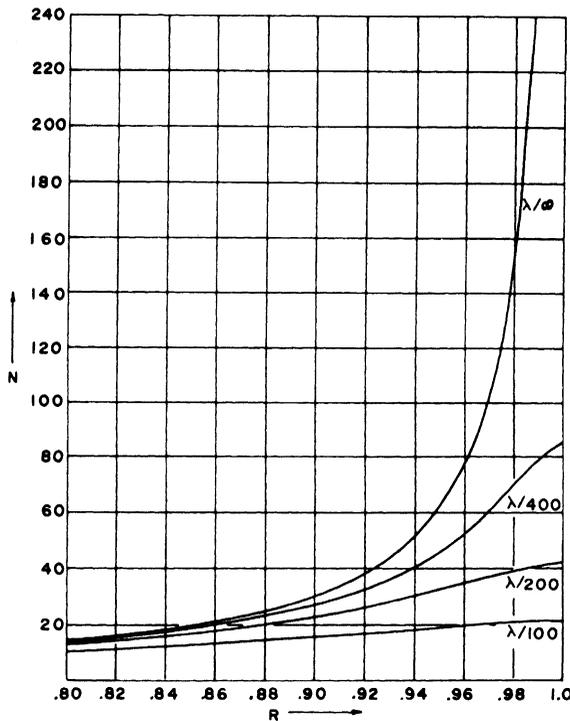


FIGURE 6.16. Effective surface finesse N plotted against reflectance R for a Fabry–Perot with mirrors having a surface defect of the Gaussian type. Curves for four different cases of surface flatness are shown [Roychoudhuri and Hercher, 1978].

6.7. TOLANSKY FRINGES WITH FABRY-PEROT INTERFEROMETER

When a pair of highly reflecting plane parallel surfaces is illuminated by a point source, a characteristic series of composite circular fringes (with decaying secondary maxima) can be observed at any distance from the surfaces. Such nonlocalized fringes appearing on diverging conical surfaces were first noted and used for the study of thin crystal plates by Tolansky (1943, 1946). Therefore, we shall call these nonlocalized multiple-beam fringes Tolansky fringes.

Tolansky fringes can be used for localized precision testing by imaging a point source between the mirror surfaces. For this method to be effective, the mirrors must be fairly close. The observation can be carried out without the use of any microscope or similar device by simply intercepting the fringes with a distant wall (Fig. 6.17). Tolansky (1948a) gave a simplified analysis. The angular diameters of the fringes are approximately given by $2\theta_n$, where θ_n follows the Fabry-Perot formula,

$$2t \cos \theta_n = n\lambda. \tag{6.13}$$

A detailed computation of the characteristics of these fringes was made by Aebischer(1971).

Tolansky fringes also find useful application in the quick alignment of a laser cavity (Bergman and Thompson, 1968) and the Fabry-Perot interferometer (Ford and Shaw, 1969; Roychoudhuri, 1973b). For this purpose, if transmission fringes are used, the secondary maxima should be aligned to be perfectly concentric with the primary maxima; a clear residual tilt in the alignment is indicated when the secondary maxima cross the primary maxima [Figure 6.18(a,b)]. However, when

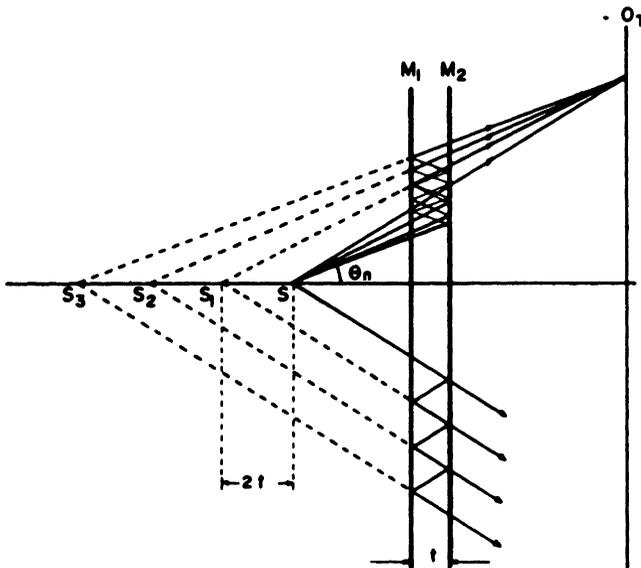


FIGURE 6.17. Formation of nonlocalized Tolansky fringes at plane (O_T) by a Fabry-Perot (M_1 - M_2) when it is illuminated by a point source S.

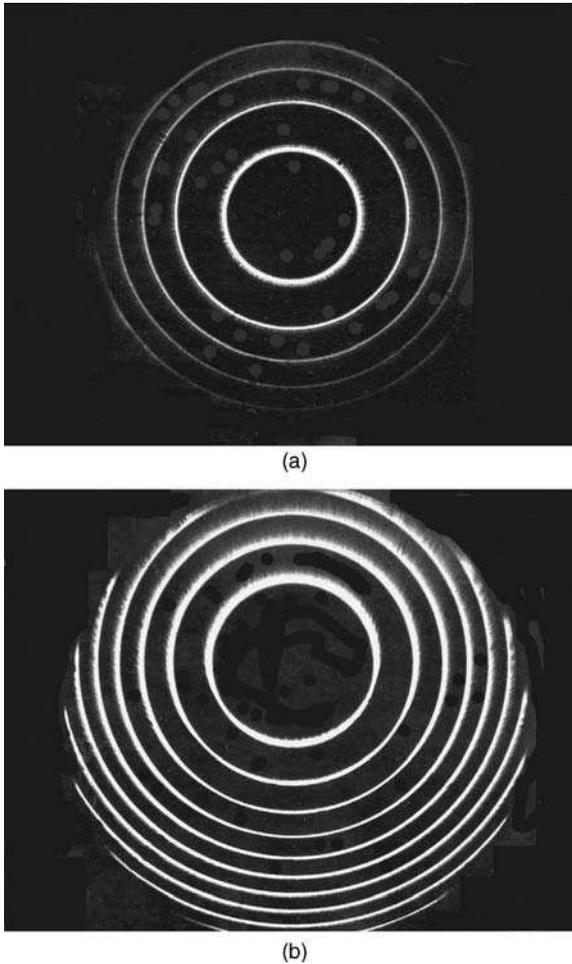


FIGURE 6.18. Some nonlocalized Tolansky fringes, (a) Fabry–Perot plates were “perfectly” parallel and (b) the plates were tilted.

the fringes in reflection are used for alignment (the weak secondary, dark fringes are almost invisible), the fringe system should be made concentric with the axis of the cone of fringes (or the illuminating point source).

Another application of Tolansky fringes involves measurements of the long-term stability of high-quality Fabry–Perot interferometers (or a laser cavity). As mentioned before, the fringes formed by the Fabry–Perot under test can be intercepted with a distant wall, and the contraction (expansion) of the fringes measured in millimeters to obtain the longitudinal expansion (contraction) of the Fabry–Perot separation in fractions of a wavelength. The tilt of the plates can also be measured by observing the crossing of the secondary fringes over the primary ones. We have used this technique to measure the long-term stability of a commercial

Fabry–Perot (all Invar structure) and one of our own constructions (all Cervit structure with thermally compensated mounts) (Roychoudhuri, 1973a, 1973b). The latter Fabry–Perot shows better stability.

6.8. MULTIPLE-BEAM INTERFEROMETER FOR CURVED SURFACES

As mentioned in the introduction, multiple-beam interferometers to test curved surfaces were developed independently by Herriott (1966), by Heintze et al. (1967), and by Biddles (1969). Such interferometers are very useful for precision-testing of various surfaces of different curvatures against a suitable master surface when they are arranged as a concentric system. A similar interferometric setup can also be adapted to other possible precision measurements, such as thermal, pressure, or composition gradients in wind tunnels, shock tubes, and so on (Herriott, 1966). The generic natures of interference fringes formed by multiple spherical wave fronts have been developed by Zerbino et al. (1984).

The essential elements of the interferometer are portrayed in Figure 6.19, where a master aplanatic M_1 and a surface under test M_2 form a concentric system. The point source S is imaged at C , the common center, by the lens L_1 . With the precise concentric arrangement, one finds at the observation plane (O_T or O_R) a uniform wavefront with perfectly spherical surfaces. To produce contour fringes, one displaces one of the mirrors by a short distance either laterally (to obtain straight fringes) or longitudinally (to obtain circular fringes). [The latter technique is also used with a confocal spherical Fabry–Perot to produce high-dispersion spectral fringes (Persin and Vukicevic, 1973).] Since such displacement introduces walk-off of the beams, Herriott (1966; see Fig. 5) introduced a compensating lens at the common center to image one mirror to the other in such a way that the reflected rays are deviated back to the same points on the mirrors, thus preserving the localized surface testing capability. For various modifications of this basic interferometer see the original references (Herriott, 1966, Heintze et al., 1967; Biddies, 1969). See Rafalowski (1988, 1990) for testing coma of decentration and asymmetric wavefronts with confocal Fabry–Perot interferometers.

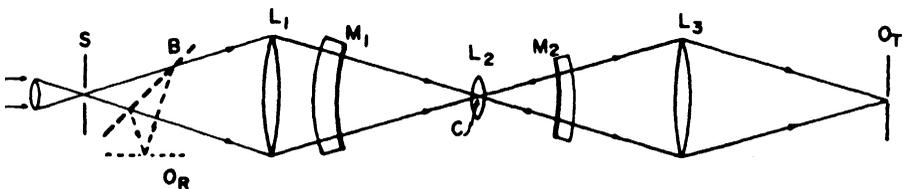


FIGURE 6.19. Multiple-beam interferometric arrangement for testing curved surfaces. Mirrors M_1 and M_2 are set almost concentric with point C . Lens L_1 images point source S at common center C . Observation can be carried out both in transmission (O_T) and reflection (O_R). Lens L_2 images surface M_1 on M_2 and vice versa to correct the walk-off defect.

Such spherical interferometers can also be used with multiple (say, n) wavelength illumination (Herriott, 1966) to reduce the fringe interval from $\lambda/2$ to some $\lambda/2n$ and thus increase the precision of testing, discussed in Section 6.5. Since such interferometers normally have large separation, one can use regular multilongitudinal mode lasers to advantage by appropriately matching the lengths of the laser and the interferometer cavities. In matching the cavities, one should carefully take into account the separation of the laser modes (laser cavity length), the free spectral range of the interferometer (interferometric cavity length), and the finesse of the interferometer. Here matching does not imply equality of the cavity lengths; they can be integral or fractional multiples of each other as the situation demands (Herriott, 1966). In using a regular multilongitudinal laser for long-path interferometers, one must be aware of the difficulties in obtaining stable high-contrast fringes (Batishko and Shannon, 1972).

6.9. COUPLED AND SERIES INTERFEROMETERS

In this section, we describe two interesting and useful modifications of multiple-beam Fizeau interferometers. The first one consists in coupling a Fizeau interferometer into a Twyman-Green interferometer as the dual interferometric arrangement shown in Figure 6.20 (Pastor and Lee, 1968; Langenbeck, 1968; Aebischer, 1970). For more details on such interferometers, see Cagnet (1954) and Candler (1951). The second modification uses three plates in series instead of two (see Fig. 6.22), as in conventional Fizeau interferometers (Post, 1954; Saunders, 1954; Roberts and Langenbeck, 1969).

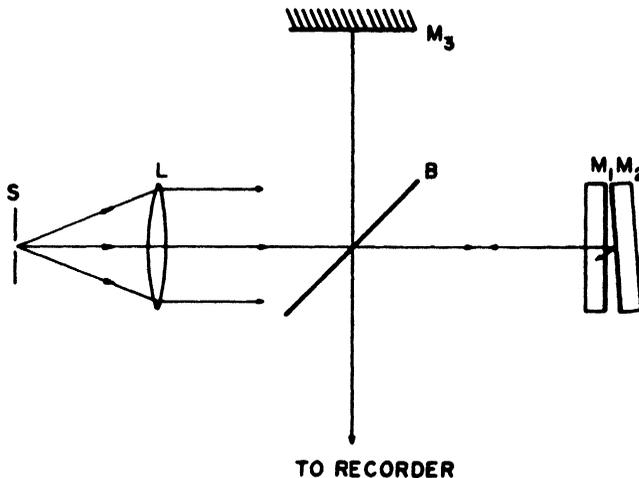


FIGURE 6.20. A coupled interferometric arrangement between a Twyman-Green and a Fizeau to control the contrast and exploit the dynamic moire technique in Fizeau interferometry.

6.9.1. Coupled Interferometer

The coupled interferometer, which consists of a Twyman-Green interferometer with one of its mirrors replaced by a Fizeau interferometer, combines several advantages in a single setup. Studying opaque optical flats in a Fizeau setup poses a general drawback because the fringes in reflection are sharp black lines against a bright background. This in itself does not create the main problem, although the contrast is poor with absorbing reflection coatings like silver film (Tolansky, 1948a; Schulz, 1951b). The Tolansky condition [Eq. (6.9)] for good-quality Fizeau fringes forces one to have low-frequency fringes in the field of view, and since the fringes are very sharp, most of the area does not produce any information, as has already been mentioned. (Several solutions to this problem were discussed in Sections 6.4 and 6.5.) The use of white-light (Section 6.4) or multiple-wavelength illumination (Section 6.5) is ruled out for reflection fringes since the large bright background due to one wavelength will wash out the sharp dark fringes due to another wavelength. This problem, which is inherent in studying opaque surfaces by reflection Fizeau, can be eliminated by the use of this dual interferometer (Fig. 6.20) to reverse the contrast and obtain transmission-like fringes in reflection Fizeau interferometry (Pastor and Lee, 1968). However, in such a coupled interferometer the reference surface (of the Fizeau arm), the Twyman mirror surface, the beam splitter, and the light collimation must be of very good optical quality.

The other advantages of this coupled interferometer have been described in detail by Langenbeck (1968), and we shall mention only the major points. It is well known that the Moiré interferometry increases the sensitivity of the test under proper conditions (see Chapter 12 for detailed background), where one superposes the test interferogram against a suitable master reference pattern. A Moiré pattern displaying only the absolute error will be presented, provided that the master reference pattern has been made to include the inherent aberrations of the interferometer itself (e.g., the errors introduced by the beam splitter in a Twyman-Green). Such a master reference can be obtained “live” from the Fizeau wedge in one arm of the dual interferometer. Such “live” fringes also have another very useful advantage: The absolute direction of a surface deviation (“hills” and “valleys”) can be read directly from the resultant interferogram, especially if the Fizeau fringes are used as the reference Moiré grid (because of the direct knowledge of the order of the reference fringes). This is illustrated by the photograph in Figure 6.21, taken from Langenbeck (1968).

Recently Gillen and Guha (2005) have demonstrated the use of the principle of the coupled interferometers, a Michelson and an FP, to independently measure the thickness and the refractive index of a substrate when available as a parallel plate. Schwider (1968) has demonstrated a superposition fringe shear interferometer (SFSI), which is a combination of an FP etalon and a shear plate to measure all the complex aberrations of a microscope objective in a single and very stable set up. Later Schwider (1997) developed this concept into a white light Fizeau interferometer where the illumination passes through an FP etalon. When the air gap (spacing) between the FP etalon plates and the Fizeau plates are equalized, one

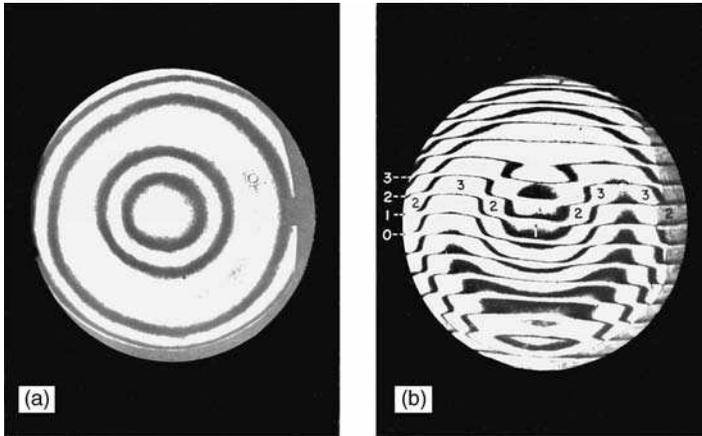


FIGURE 6.21. (a) Twyman-Green interferogram of a deformed mirror. Concentric fringes may be high or low areas. Dynamic observation is needed to determine the direction of order or interference. (b) Same, but with superimposed wedge field, permitting determination of the direction of the order of interference. (From Langenbeck 1968.)

can take advantage of the order number of the fringes with reference to the $OPD=0$ fringe.

6.9.2. Series Interferometer

In this second modification of a Fizeau interferometer, introduced to measure the homogeneity of optical plates, the plate under test is inserted between two reference Fizeau mirrors (Figure 6.22). Such a device is necessary for testing high-quality beam splitters whose function is not only to reflect but also to transmit wavefronts without distortion. Hence, a simple surface flatness test of the reflecting surface of the beam splitter is not sufficient. The Fizeau fringes alone in reflection, with the plate under test on the side of the observer, cannot directly map the index variation when there is a simultaneous surface variation. Since the interfering light beams in such a three-plate interferometer (Fig. 6.22) pass through the plate under test many times, the sensitivity of the measurement of the index variation or homogeneity increases by a large factor.

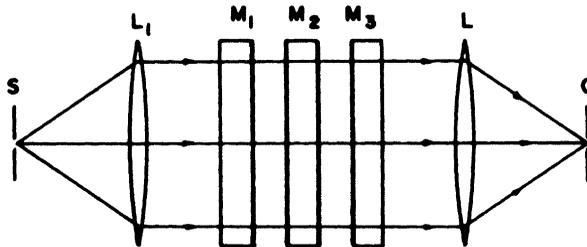


FIGURE 6.22. A series interferometer with three mirrors for precision testing of the homogeneity of a flat (M_2).

Such interferometers have been named by their independent discoverers as the “in-line interferometer” (Saunders, 1954) and the “series interferometer” (Post, 1954).

The various different conditions and reflectances for the plates of such interferometers have been described in the references cited. Of the more recent papers along similar lines by Ashton and Marchant (1967) and by Roberts and Lan-genbeck (1969), the latter describes how to evaluate and obtain a contour map of fractional index variation better than 10^{-6} .

6.10. HOLOGRAPHIC MULTIPLE-BEAM INTERFEROMETERS

The advantages of holographic interferometric tests of optical components are described in Chapter 12. When the choice has been made in favor of holographic interferometry, one can introduce the precision associated with multiple-beam interferometry by employing the hologram to reconstruct many wave-fronts that can be combined to produce sharp multiple-beam fringes. This is achieved by the hologram, nonlinearly recorded, that produces several higher order reconstructed beams. Matsumoto (1969) and Bryngdahl (1969) described such techniques, by which they recorded single-exposure holograms with single-wavelength illumination. An obvious advantage of this method is that one can obtain multiple-beam Fizeau type fringes from diffuse surfaces under test.

Burch et al. (1966) described a different method that uses linear holographic recording, but has multiple exposures made while bending the object. Multiple exposures can also be obtained by changing the direction of illumination or by using multiple-wavelength illumination, instead of bending the object. But before undertaking multiple-beam holographic testing (Shalan and Jonathan, 1978) with its attendant difficulties, one should make reasonably certain that simple two-beam interferometry with a good-quality fringe detecting device is insufficient for the precision necessary.

6.11. TEMPORAL EVOLUTION OF FP FRINGES AND ITS MODERN APPLICATIONS

With the advancement of sophisticated mirror coatings, polishing, and alignment technologies, people have been able to achieve reflectance exceeding 0.9999 and effective finesses exceeding 10,000. Such super finesse FPs are becoming critically important tools for various specialized applications such as gravitational wave detection [LIGO], trace amount of gas detection by methods called CRDS (cavity ring down spectroscopy) (Casaes et al., 2002), and study of micro cavity quantum electrodynamics (Kimble, 1994; Blais et al., 2004). Optical testing of the surface quality and/or alignment of such mirrors is a daunting task, but the principle of operation of FP itself comes to rescue (Herbelin, 1980; Itoh, 2001). In fact, Herbelin et al. exploited the long photon life time of super finesse FP to measure the high reflectance, low absorption, and scattering losses using a modulated laser beam. Then

this basic concept has been reinvented (O'Keefe and Deacon, 1988) for ultra sensitive measurement of very low concentration of gases inside the FP cavity.

In Section 6.3.1, we have underscored that the multiply reflected infinite FP series can be effectively terminated at $n = N$, the finesse number [Eq. (6.3)], because the energy carried by the beams after the N th reflection become very small and hence negligible. Here we justify it numerically. The relative reduction in intensity, b , of the n th transmitted beam compared to the first one can be given by

$$b = I_n/I_1 = T^2 R^{2n}/T^2 \text{ or } \ln b = 2n \ln R \quad (6.14)$$

For a range of moderate to high finesse FP with $R = 0.9000$ to 0.9999 , the intensity of the N th beam will be reduced approximately by the same factor of $b = 1.87 \times 10^{-3}$. So, one can safely terminate the infinite FP series by a finite sum of N terms, and the normalized intensity can be reformulated as a series by complex conjugate multiplication of the transmitted amplitude from Eq. (6.1),

$$\begin{aligned} I_{cw}(v, \tau) &\cong \left(\sum_{n=0}^{N-1} TR^n e^{i2\pi n v \tau} \right) \left(\sum_{m=0}^{N-1} TR^m e^{-i2\pi m v \tau} \right) \\ &= \sum_{n=0}^{N-1} T^2 R^{2n} + 2 \sum_{n \neq m}^{N-1} T^2 R^{n+m} \cos[2\pi(n-m)v\tau] \end{aligned} \quad (6.15)$$

We have numerically verified that this Eq. (6.5) quite accurately represents the traditional FP Eq. (6.1) as the superposition of a finite number of cosine fringes of varying contrast (fringe visibility). We can now appreciate that for a high finesses FP, the quantitative representation of FP fringes would require one to wait for a time τ_0 till the FP has been able to generate N transmitted beams. Thus, we define a very important parameter for an FP; its characteristic time constant as [Kastler, 1974; Roychoudhuri, 1975],

$$\tau_0 = N\tau \quad (6.16)$$

This time duration τ_0 is required for the FP to establish its steady state fringe pattern if a light beam is suddenly turned on. With the routine availability of light sources of very short duration and very high speed detectors, the use of super finesses FP (large τ_0) requires a simple classical understanding of the evolution of the FP fringes. If the incident light beam, $a(t) \exp[i2\pi v t]$, is a pulse of finite duration, δt , then following the model of Eq. (6.15), the time varying, transmitted intensity can be given by (Roychoudhuri et al., 2003; Roychoudhuri, 2004),

$$\begin{aligned} I'_{pls}(t, v, \tau) &= \left| \sum_{n=0}^{N-1} a(t - n\tau) TR^n e^{i2\pi v(t-n\tau)} \right|^2 \\ &= \sum_{n=m=0}^{N-1} T^2 R^{2n} a^2(t - n\tau) + 2 \sum_{n>m}^{N-1} T^2 R^{n+m} a(t - n\tau) a(t - m\tau) \cos[2\pi(n-m)v\tau] \end{aligned} \quad (6.17)$$

After integration over the entire train of the N_R replicated pulses,

$$I_{pls}(v, \tau) = \sum_{n=0}^{N-1} T^2 R^{2n} + 2 \sum_{n \neq m}^{N-1} T^2 R^{n+m} \gamma(|n - m|\tau) \cos[2\pi(n - m)v\tau] \quad (6.18)$$

Here, $\gamma(\tau)$ is the normalized autocorrelation function between the m th and the n th replicated pulses that reduce to unity when the width of the pulse $a(t)$ is theoretically infinite. We have found that numerically it only needs to exceed $\tau_0 = N\tau$.

$$\text{Lt}_{\delta t \rightarrow \tau_0 = N\tau} [I_{pls}(v, \tau)] \cong I_{cw}(v, \tau) \quad (6.19)$$

This is a very important result because it shows by very simple mathematics that the generalized time integrated fringe pattern for a pulse converges to the standard CW fringe pattern (from Eq. (6.15) to Eq. (6.18)); this recognizes that an FP is a pair of beam splitters that replicates an incident pulse into a train of N_R effective pulses with the periodic delay of τ . One does not need to complicate this simple physical picture by introducing time-frequency Fourier transformations that tend to bring confusion between the actual carrier frequency of the incident radiation and the mathematical Fourier frequencies of a transformed pulse. Thus, if the width of the incident pulse exceeds the characteristic time constant of the FP, it produces CW-like fringes; otherwise the fringe becomes broader than the CW situation due to the factor $\gamma(\tau)$. This $\gamma(\tau)$ represents the reduction in the fringe contrast due to physical superposition of unequal amplitudes (translated pulses). In the extreme case where $\delta t < \tau$, a fast detector can only detect an exponentially dying train of pulses for lack of superposition, irrespective of whether the carrier frequency of the pulse is in resonance or antiresonance with FP spacing. This last case of nonoverlapping pulses (Fig. 6.23c) can be represented by after using Eq. (6.7) and neglecting the cross terms due to absence of overlapping,

$$I_{pls}(t, v) = T^2 \sum_n R^{2n} a^2(t) \quad (6.20)$$

If one considers only the peak intensity, then a^2 at the peak is a fixed number and the time envelope of the intensity fall can be represented by

$$\ln |I_{pls}(t = n\tau)|_{peak} = \ln [a^2 T^2 R^{2n}] = (2 \ln R)n + 2 \ln (aT) \quad (6.21)$$

If one plots the log of the transmitted peak intensities against the pulse number, n , this will give a straight line with a slope of $(2 \ln R)$ and an intercept of $(2 \ln (aT))$. Since

$$T + R + A = 1 \quad (6.22)$$

where A represents the absorption and scattering losses, one can use a single short pulse to characterize the FP mirror properties or the absorption loss of trace gases if its absorption frequency matches with the carrier frequency of the incident pulse. Conceptually, this is the foundation behind characterizing FP mirrors with a modulated laser (Herbelin et al., 1980) and the success of CRDS (O'Keefe and Deacon, 1988). Readers should note that our $\tau_0 = N_R\tau$, although closely similar, is

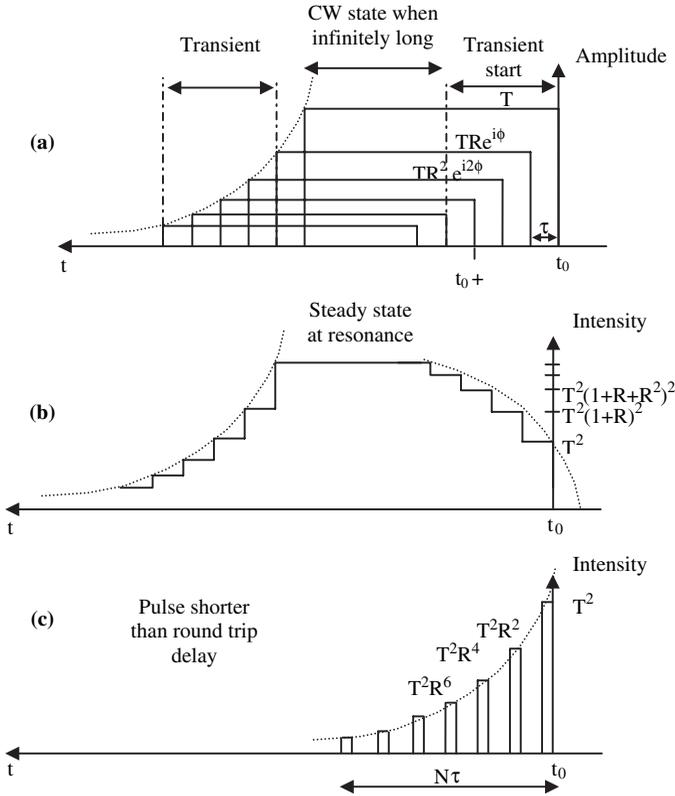


FIGURE 6.23. A pair of beam splitters create a train of delayed, identical pulses of reduced amplitudes from a single incident pulse. The width of the pulse, δt , relative to the FP transit time, $\tau = 2d/c$, determines the degree of real physical superposition of the folded pulse inside the FP. For $\delta t \sim \tau$, the partial superposition of the amplitudes is depicted in (a), and the corresponding, resultant intensity transmittance is depicted in (b) at the FP resonance ($2d = m\lambda$). When the pulse is very much longer than τ , we have a steady state condition, or CW interference. When $\delta t \ll \tau$, there is no physical overlap between the train of pulses. This is depicted in (c). In the absence of superposition, the transmitted intensity peaks are given by the square modulus of the individual replicated pulses; the roundtrip phase shift is no longer operative and the resonance enhanced higher transmittance is also absent (Lee, 2005).

not the photon life time for an FP, as is popularly defined in the literature using FP with short pulses. We have defined it as the FP time constant that represents the actual physical time period during which useful N_R pulses are generated by the FP mirror pair.

6.12. FINAL COMMENTS

Multiple-beam interferometry provides very high precision surface measurements compared to two-beam interferometry under comparable data processing environments. However, the requirements on source coherence and the basic

experimental setup are more expensive and time consuming, so multiple-beam interferometry is less useful in routine shop floor testing, where $\lambda/10$ precision is adequate. Measuring surface quality to much better than $\lambda/50$ is more efficient using various forms of multiple-beam interferometry exploiting principles of Fabry-Perot interferometry and multiple-beam Fizeau interferometers including fringes of equal chromatic order (FECO). There are many unique niches where multiple-beam interferometers are ideally suited. The first historic example is the characterization of surface microtopography using multiple-beam Fizeau interferometer (Bennett and Bennett, 1967). However, readers interested in measuring surface roughness to a root mean square (rms) accuracy of 20 Å using two-beam Fizeau interferometer may see Eastman (1980), where he uses a piezoelectrically scanned interferometer. A natural application of Fabry-Perot interferometry is to test the component Fabry-Perot plates or solid etalons requiring a precision of $\lambda/200$ or better (Hariharan et al., 1984; Hernandez, 1988). Measurements of parallelism of optically contacted etalons have been described by Killeen et al. (1981). Modern electronic industry requires bonding optically flat, large silicon crystal wafers with precision micrometric spacing. Rhee et al. (1990) have successfully exploited the principle of Fabry-Perot interferometry.

Multiple-beam interferometry using grating as one of the two conventional mirrors has been developed for special applications. Useful examples can be found from Bates and Li (1986) and also from Rodriguez-Zurita et al. (1998) and from references there. A combination of a Fabry-Perot interferometer and a two-beam interferometer has been cleverly used to very accurately measure the thickness of air layers by Schwider (1979). Precision associated with multiple-beam fringes can also be generated using holography that is suitable for realtime testing of deformation of microobjects (Shaalan and Jonathan, 1978).

Sometimes, four-beam and three-beam interferometry are also useful. Ma kosch and Jaerisch (1978) have described four-beam Fizeau interferometry for contactless surface testing using a grating as the reference flat. Lin and Cowley (1986) have described the advantages of using three-beam Ronchigram in measuring primary aberrations. Strains of curved surfaces that are coated with photosensitive material have been measured by Chiang and Kim (1984) by three-beam holographic interferometric techniques.

It is worth mentioning that there have been rapid technological advancements in the diverse use of the principle of multiple-beam Fabry-Perot interferometry that are beyond the domain of this book. Miniaturization of all instruments has become the general technological drive. With the advancements in fiber Bragg gratings, micro-opto-electromechanics (MOEMs) and nano photonics, Fabry-Perot interferometers are also getting integrated into miniature and micro-optical systems. Testing such micro Fabry-Perots will become the next challenge.

REFERENCES

- Acbischer N., "Etudes d'Interférences en Ondes Multiples par Diagramme Complexe Visualiser les Franges en Réflexion." *Nouv. Rev. Opte. Appl.*, **1**, 233 (1970).

- Aebischer N., "Calculs de Profils Dissymetriques Observables sur des Figures d'Interférences en Ondes Multiples Sphériques," *Nouv. Rev. Opte. Appl.*, **2**, 351 (1971).
- Airy G., *Mathematical Transactions*, 2nd ed., 1836, p. 301. (obtainable from Library of British Museum, London).
- Ashton A. and A. C. Marchant, "Note on the Testing of Large Glass Panels," *Opt. Acta*, **14**, 203 (1967).
- Baird K. M., "Interferometry: Some Modern Techniques," in *Advances in Optical Techniques*, A. C. S. Van Heel, Ed., North-Holland, Amsterdam, 1967, Chap. 4.
- Baird K. M. and G. R. Hanes, "Interferometers," in: *Applied Optics and Optical Engineering*, vol. IV, R. Kingslake, Ed., Academic Press, New York, 1967, Chap. 9.
- Barakat N., A. S. Farghaly, and A. Abd-El-Aziz, "Studies on Multiple-Beam Interference Fringes Formed on High Order Planes of Localization: Intensity Distribution and Fringe Shift between Successive Planes of Localization," *Opt. Acta*, **12**, 205 (1965).
- Bates B. and Y. Li, "Multiple-Beam Grating Interferometry: A Comparison of Techniques Applied to the Measurement of Coating Thickness," *Appl. Opt.*, **25**, 835–836 (1986).
- Batishko C. R. and R. R. Shannon, "Problem in Large-Path-Difference Laser Interferometry," *Appl. Opt.*, **11**, 195 (1972).
- Benedetti-Michelangeli G., "New Technique for the Evaluation of the Quality of Plane Surfaces," *Appl. Opt.*, **7**, 712 (1968).
- Bennett H. E. and J. M. Bennett, "Precision Measurements in Thin Film Optics," in: *Physics of Thin Films*, vol. IV, G. Hass and R. F. Thun, Eds., Academic Press, New York, 1967, Chap. 1.
- Bennett J. M., "Measurement of the RMS Roughness, Autocovariance Function and Other Statistical Properties of Optical Surfaces Using a FECO Scanning Interferometer," *Appl. Opt.*, **15**, 2705 (1976).
- Bergman T. G. and J. L. Thompson, "An Interference Method for Determining the Degree of Parallelism of Laser Surfaces," *Appl. Opt.*, **7**, 923 (1968).
- Bhatnagar G. S., K. Singh, and B. N. Gupta, "Transmission Profile of a Fabry-Perot Interferometer Suffering from Asymmetric Surface Defects," *Nouv. Rev. Opt.*, **5**, 237 (1974).
- Biddies B. J., "A Non-contacting Interferometer for Testing Steeply Curved Surfaces," *Opt. Acta*, **16**, 137 (1969).
- Blais A., R-S. Huang, A. Wallraff, S. M. Girrin, and R.J. Schoelkopf, "Cavity quantum electrodynamics for superconducting Electrical Circuits: An Architecture for Quantum Computation," *Phys. Rev. A*, **69**, 062320 (2004).
- Boersch H., H. J. Eichler, M. Pfundstein, and W. Wisemann, "Measurement of Length Shifts Down to 10^{-5} Å with a Three-Mode Laser," *IEEE J. Quantum Electron.*, **QE-10**, 501 (1974).
- Bom M. and E. Wolf, *Principles of Optics*, 5th ed., Pergamon, Oxford, 1975.
- Boulouch M. R., "Dedoublement des Franges d'Interférence en Lumière Naturelle," *J. Phys.*, 3rd series, **2**, 316 (1893).
- Boyd G. D. and J. P. Gordon, "Confocal Multimode Resonator for Millimeter through Optical Wavelength Masers," *Belt Syst. Tech. J.*, **40**, 489 (1961).
- Briers J. D. "Interferometric Testing of Optical Systems and Components: a Review," *Opt. Laser Technol.*, **4**, 28 (1972).

- Brossel J., "Multiple-Beam Localized Fringes. I: Intensity Distribution and Localization II: Conditions of Observation and Formation of Ghosts," *Proc. Phys. Soc.*, **59**, 224, 234 (1947).
- Bruce C. F. and F. P. Sharpless, "Relative Flatness Measurement of Uncoated Optical Flats," *Appl. Opt.*, **14**, 3082 (1975).
- Bryngdahl O., "Multiple-Beam Interferometry by Wavefront Reconstruction," *J. Opt. Soc. Am.*, **59**, 1171 (1969).
- Bünnagel R., H. A. Oehring, and K. Steiner, "Fizeau Interferometer for Measuring the Flatness of Optical Surfaces," *Appl. Opt.*, **7**, 331 (1968).
- Burch J. M., A. E. Ennos, and R. J. Wilton, "Dual- and Multiple-Beam Interferometry by Wavefront Reconstruction," *Nature*, **209**, 1015 (1966).
- Cagnet M., "Méthodes Interférométriques Utilisant les Franges de Superposition," *Rev. Opt.*, **33**, 1, 113, 229 (1954).
- Casaes R., J. P. Provencal, and R. J. Saykally, "High Resolution Pulsed Infrared Cavity Ring Down Spectroscopy: Application to Laser Ablated Carbon Clusters," *J. Chem. Phys.*, **116**, 6640–6647 (2002).
- Candler C, *Modern Interferometers*, Hilger and Watts, London, 1951.
- Chabbal R., "Recherche des Meilleures Conditions d'Utilisation d'un Spectromètre Photoélectrique Fabry-Perot," *J. Rech. Cent. Nat. Rech. Sci.*, **24**, 138 (1953).
- Chabbal R., "Finise Limite d'un Fabry-Perot Forme de Lames Imparfaites," *J. Phys. Radium.*, **19**, 295 (1958).
- Chiang F. P. and C. C. Kin, "Three-Beam Interferometric Technique for Determination of Strain of Curved Surfaces," *Opt. Eng.*, **23**, 766–768 (1984).
- Clapham P. B. and G. D. Dew, "Surface Coated Reference Flats for Testing Fully Aluminized Surfaces by Means of a Fizeau Interferometer," *J. Sci. Instrum.*, **44**, 899 (1967).
- Connes P., "L'Étalon de Fabry-Perot Sphérique," *J. Phys. Radium*, **19**, 262 (1958).
- Cook A. H., *Interference of Electromagnetic Waves*, Clarendon Press, Oxford, 1971.
- Dew G. D., "A Method for Precise Evaluation of Interferograms," *J. Sci. Instrum.*, **41**, 160 (1964).
- Dew G. D., "The Measurement of Optical Flatness," *J. Sci. Instrum.*, **43**, 409 (1966).
- Dyson J., "Optics in a Hostile Environment," *Appl. Opt.*, **7**, 569, (1968).
- Eastman J. M., "Effects and Measurement of Scattering and Absorption of Thin Films," *Proc. Soc. Photo-Opt. Instrum. Eng.*, **50**, 43 (1975).
- Eastman J. M., "The Scanning Fizeau Interferometer: An Automated Instrument for Characterizing Optical Surfaces," *Opt. Eng.*, **19**, 810–814 (1980).
- Eastman J. M. and P. W. Baumeister, "Measurement of the Microtopography of Optical Surfaces Using a Scanning Fizeau Interferometer," *J. Opt. Soc. Am.*, **64**, 1369A (1974).
- Fabry C. and A. Perot, "Sur les Franges des Lames Minces Argentées et Leur Application a la Mesure de Petites Epaisseurs d'Air," *Ann. Chim. Phys.*, **12**, 459 (1897).
- Fizeau M. H. "Recherches sur les Modifications que Subit la Vitesse de la Lumière dans le Verre et Plusieurs," *Ann. Chim. Phys.*, **66**, 429 (1862a).
- Fizeau M. H., "Recherches sur les Modifications que Subit la Vitesse de la Lumière dans le Verre et Plusieurs Autres Corps Solides sous L'Influence de la Chaleur," *C. R. Acad. Sci. Paris*, **54**, 1237 (1862b).
- Ford D. L. and J. H. Shaw, "Rapid Method of Aligning Fabry-Perot Etalons," *Appl. Opt.*, **8**, 2555 (1969).

- Fox A. G. and T. Li, "Resonant Modes in a Master Interferometer," *Bell Syst. Tech. J.*, **40**, 453 (1961).
- Gillen G. D. and S. Guha, "Use of Michelson and Fabry–Perot interferometry for independent determination of the refractive index and physical thickness of wafers," *Appl. Opt.*, **44**, 344–347 (2005).
- Hanes G. R., "Limiting Precision in Optical Interferometry," *Can. J. Phys.*, **37**, 1283 (1959).
- Hanes G. R., "Quantum Limit to Precision of Wavelength Determination," *Appl. Opt.*, **2**, 465 (1963).
- Hariharan P., "Optical Interferometry," *Rep. Prog. Phys.*, **54**, 339–390 (1991). doi: 10.1088/0034-4885/54/3/001.
- Hariharan P. B., F. Oreb, and A. J. Leistner, "High Precision Digital Interferometry: Its Application to the Production of An Ultrathin Solid Fabry-Perot Etalon," *Opt. Eng.*, **23**(3), 294–297, (1984).
- Heintze L. R., H. D. Polster, and J. Vrabel, "A Multiple-Beam Interferometer for Use with Spherical Wavefront," *Appl. Opt.*, **6**, 1924 (1967).
- Heisenberg W., *The Physical Principles of the Quantum Theory*, Dover, New York, 1949.
- Hillebrands B., "Progress in Multipass Tandem Fabry-Perot Interferometry: 1. A Fully Automated, Easy to Use, Self-Aligning Spectrometer with Increased Stability and Flexibility," *Rev. Sci. Instrum.*, **70**, 1589–1598 (1999).
- Herbelin J. M., J. A. McKay, M. A. Kwok, R. H. Ueunten, D. S. Urevig, D. J. Spencer, and Benard DJ., "Sensitive Measurement of Photon Life Time and True Reflectances in an Optical Cavity by a Phase-shift Method," *Appl. Opt.*, **19**, 144 (1980).
- Hernandez G., *Fabry–Perot Interferometers*, Cambridge University Press, Cambridge, 1988.
- Herriott D. R., "Multiple-Wavelength Multiple-Beam Interferometric Observations of Flat Surfaces," *J. Opt. Soc. Am.*, **51**, 1142 (1961).
- Herriot D. R., "Spherical-Mirror Oscillating Interferometer," *Appl. Opt.*, **2**, 865 (1963).
- Herriot D. R., "Long-Path Multiple-Wavelength Multiple-Beam Interference Fringes," *J. Opt. Soc. Am.*, **56**, 719 (1966).
- Hill R. M., "Some Fringe-Broadening Defects in a Fabry–Perot Etalon," *Opt. Ada*, **10**, 141 (1963).
- Hill R. M. and C. F. Bruce, "Limiting Precision in a Scanning Optical Interferometer," *Aust. J. Phys.*, **15**, 194 (1962).
- Hodgkinson I. J., "A Method for Mapping and Determining the Surface Defects Function of Pairs of Coated Optical Flats," *Appl. Opt.*, **8**, 1373 (1969).
- Hodgkinson I. J., "The Application of Fringes of Equal Chromatic Order to the Assessment of the Surface Roughness of Polished Fused Silica," *J. Phys.*, **E3**, 300 (1970).
- Itoh S-L, Y. Yamazaki, K. Katoh, and J. Chen, "High Precision Testing Method for Fabry–Perot Etalon," *Optical Review* **8**, 179–183 (2001).
- Jacquinot M. P., "New Developments in Interference Spectroscopy," *Rep. Prog. Phys.*, **23**, 267 (1960).
- Jenkins F. A. and H. E. White, *Fundamentals of Optics*, 3rd ed., McGraw-Hill, New York, 1957.
- Kastler A. "Transmission d'une Impulsion Lumineuse par un interféromètre Fabry–Perot", *Nouv. Rev. Opt.*, **5**, 133–139 (1974).
- Killeen T. L., P. B. Hays, and J. DeVos, "Parallelism Maps for Optically Contacted Etalons," *Appl. Opt.*, **20**, 26116–26119 (1981).

- Kimble H., *Cavity Quantum Electrodynamics*, Academic Press, San Diego, 1994.
- Kinosita K., "Numerical Evaluation of the Intensity Curve of a Multiple-Beam Fizeau Fringe," *J. Phys. Soc. Jap.*, **8**, 219 (1953).
- Koehler W. F., "Multiple-Beam Fringes of Equal Chromatic Order. I: Phase Change Considerations. II: Mechanism of Polishing Glasses," *J. Opt. Soc. Am.*, **43**, 738, 743 (1953).
- Koehler W. F., "Multiple-Beam Fringes of Equal Chromatic Order. IV: Use of Multilayer Film," *J. Opt. Soc. Am.*, **45**, 934 (1955a).
- Koehler W. F., "Multiple-Beam Fringes of Equal Chromatic Order. VII: Mechanism of Polishing Glass," *J. Opt. Soc. Am.*, **45**, 1015 (1955b).
- Koehler W. F. and A. Eberstein, "Multiple-Beam Fringes of Equal Chromatic Order. III: The Cleavage of Topaz," *J. Opt. Soc. Am.*, **43**, 747 (1953).
- Koehler W. F. and W. C. White, "Multiple-Beam Fringes of Equal Chromatic Order. V: Fringe Fine Structure; VI: Method of Measuring Roughness," *J. Opt. Soc. Am.*, **45**, 940, 1011 (1955).
- Koppelman G., "Eine beugungsbedingte Auflösungs Grenze in der Mehrstrahl-Interferometrie," *Opt. Acta*, **13**, 211 (1966).
- Koppelman G., "Multiple-Beam Interference and Natural Modes in Open Resonators," in: *Progress in Optics*, vol. VII, E. Wolf, Ed., Wiley-Interscience, New York, 1969, Chap. 1.
- Koppelman G. and K. Krebs, "Eine Registriermethode zur Vermessung der Reliefs höchsterbener Oberflächen," *Optik*, **18**, 349 (1961).
- Kuhn H., "New Technique in Optical Interferometry," *Rep. Phys. Soc. Prog. Phys.*, **14**, 80 (1951).
- Lang J. and G. Scott, "Resolution Limits in Multiple-Beam Interferometry," *J. Opt. Soc. Am.*, **58**, 81 (1968).
- Langenbeck P., "Optical Wavefront Mapping by Dual Interferometry," *J. Opt. Soc. Am.*, **58**, 499 (1968).
- Langenbeck P., "Fizeau Interferometer-Fringe Sharpening," *Appl. Opt.*, **9**, 2053 (1970).
- Laurent L., "Sur Phisieurs Appareils d'Optique, Destinés à Controler les Surfaces Planes: Parallèles, Perpendiculaires et Obliques," *C. R. Acad. Sci. Paris*, **96**, 1035 (1883).
- Lee D., "A Comprehensive View of Temporal Domain Interference and Spectral Interpretation of Short Pulses," Ph. D. Thesis, 2005, University of Connecticut.
- LIGO, "Laser Interferometer Gravitational Wave Observatory," www.ligo.caltech.edu
- Lin J.-A. and J. M. Cowley, "Aberration Analysis by Three-Beam Interferograms," *Appl. Opt.*, **25**, 2245–2246 (1986).
- Logan J. L., "Gravitational Waves—A Progress Report," *Phys. Today*, March, 44 (1973).
- Makosch, G. and W. Jaerich, "Mapping of Optical Surfaces with Quarter Wavelength Fringes," **17**, 744–747, **17**, 1990–1991 (1978).
- Malacara D., A. Cornejo, and M. V. R. K. Murty, "Bibliography of Various Optical Testing Methods," *Appl. Opt.*, **14**, 1065 (1975).
- Matsumoto K., "Holographic Multiple-Beam Interferometry," *J. Opt. Soc. Am.*, **59**, 777 (1969).
- Moos H. W., G. F. Imbusch, L. F. Mollenauer, and A. L. Schawlow, "Tilted-Plate Interferometry with Large Plate Separations," *Appl. Opt.*, **2**, 817 (1963).
- Moss G. E., L. R. Miller, and R. L. Forward, "Photon-Noise-Limited Laser Transducer for Gravitational Antenna," *Appl. Opt.*, **10**, 2495 (1971).

- Murty M. V. R. K., "Multiple-Pinhole Multiple-Beam Interferometric Observation of Flat Surfaces," *Appl. Opt.*, **1**, 364 (1962).
- O'Keefe A. and D. A. G. Deacon, "Cavity Ring-down Optical Spectrometer for Absorption Measurements using Pulsed Laser Sources," *Rev. Sci. Instrum.* **59**, 2544–2551 (1988).
- Pastor J. and P. H. Lee, "Transmission Fringes in Reflection Multiple-Beam Interferometry," *J. Opt. Soc. Am.*, **58**, 149 (1968).
- Persin A. and D. Vukicevic, "Block Defocused Spherical Fabry-Perot Interferometer," *Appl. Opt.*, **12**, 275 (1973).
- Pilston R. G. and G. N. Steinberg, "Multiple-Wavelength Interferometry with Tunable Source," *Appl. Opt.*, **8**, 553 (1969).
- Polster H. D., "Multiple-Beam Interferometry," *Appl. Opt.*, **8**, 522 (1969).
- Post D., "Characteristics of the Series Interferometer," *J. Opt. Soc. Am.*, **44**, 243 (1954).
- Rafalowski M., "Testing the Coma of Decentration With a Confocal Fabry-Perot Interferometer," *Appl. Opt.*, **27**, 3046–3050 (1988).
- Rafalowski M., "Precise Testing of Asymmetric Wavefronts With an Inverting Interferometer," *Appl. Opt.*, **29**, 1477–1481 (1990).
- Raymond O. J., "Limiting Fringe Pointing Precision in a Scanning Two-Beam Interferometry," *Appl. Opt.*, **9**, 1140 (1970).
- Rhee I., F. M. Gasparini, A. Petrou, and D. J. Bishop: "Si Wafers Uniformly Spaced; Bonding and Diagnostics," *Rev. Sci. Instrum.*, **61**, 1528–1536 (1990).
- Roberts F. E. and P. Langenbeck, "Homogeneity Evaluation of Very Large Disks," *Appl. Opt.*, **8**, 2311 (1969).
- Rodriguez-Zurita, G. J., Pedraza-Contreras, R. Pastrana-Sánchez, A. Cornejo-Rodriguez and F. Vázquez-Castillo, "Multiple-Beam Interference with Wear-Grazing Waves in Dielectric Wedges: Monochromatic and Polychromatic Cases," *Rev. Mex. Phys.*, **44**, 147–154 (1998).
- Rosier F. L., "Mapping of High Quality Optical Flats without Reflection Coating," *J. Opt. Soc. Am.*, **52**, 471 (1962).
- Roesler F. L. and W. Traub, "Precision Mapping of Pairs of Uncoated Optical Flats," *Appl. Opt.*, **5**, 463 (1966).
- Rogers J. R.: "Fringe Shifts in Multiple Beam Fizeau Interferometry," *J. Opt. Soc. Am.*, **72**, 638–643 (1982).
- Roychoudhuri C., "Brillouin Spectra of CaF₂ Microcrystals Using a Stable 3-Pass Fabry-Perot Interferometer," *Appl. Phys. Lett.*, **23**, 543 (1973a).
- Roychoudhuri C., *Multi-pass Fabry-Perot Interferometer for Brillouin Scatter Measurements*, Ph.D. Thesis, University of Rochester, New York, 1973b; University Microfilms No. 74-14413.
- Roychoudhuri C., "Dynamic and Multiplex Holography with Scanning Fabry-Perot Fringes," *Opt. Commun.*, **10**, 160 (1974).
- Roychoudhuri C., "Response of Fabry-Perot Interferometers to Light Pulses of Very Short Duration," *J. Opt. Soc. Am.*, **65**, 1418 (1975).
- Roychoudhuri C. and M. Hercher, "Stable Multi-pass Fabry-Perot Interferometer: Design and Analysis," *Appl. Opt.*, **16**, 2514 (1978).
- Roychoudhuri C., D. Lee, Y. Jiang, S. Kittaka, N. Nara, V. Serikov, and M. Oikawa, "Limits of DWDM With Gratings and Fabry-Perots and Alternate Solutions," Invited Paper at the ITCOM Conference, *Proc. SPIE Proc.*, **5246**, 333–344, 2003.

- Roychoudhuri C., "Propagating Fourier Frequencies vs. Carrier Frequency of a Pulse Through Spectrometers and Other Media," in: *Interferometry XII: Techniques and Analysis*, SPIE Proc. vol. 5531, (2004), pp. 450–461.
- Saunders J. B., "Precise Topography of Optical Surfaces," *J. Res. Nat. Bur. Stand.*, **47**, 148 (1951).
- Saunders J. B., "In-line Interferometer," *J. Opt. Soc. Am.*, **44**, 241 (1954).
- Schulz G., "Ein Interferenzverfahren zur absoluten Ebenheitsprüfung langs beliebiger Zentralschnitte," *Opt. Ada*, **14**, 375 (1967).
- Schulz G. and J. Schwider, "Precise Measurement of Planeness," *Appl. Opt.*, **6**, 1077 (1967).
- Schulz G. and J. Schwider, "Interferometric Testing of Smooth Surfaces," in: *Progress in Optics*, vol. XIII, pp.93, E. Wolf, Ed., Elsevier North-Holland, New York, 1976.
- Schulz G., J. Schwider, C. Hiller, and B. Kicker, "Establishing an Optical Flatness Standard," *Appl. Opt.*, **10**, 929 (1971).
- Schulz L. G., "Accurate Thickness Measurements with a Fabry-Perot Interferometer," *J. Opt. Soc. Am.*, **40**, 177 (1950a).
- Schulz L. G., "An Interferometric Method for Accurate Thickness Measurements of Thin Evaporated Films," *J. Opt. Soc. Am.*, **40**, 690 (1950b).
- Schulz, L. G., "The Effect of Phase Changes in White Light Interferometry," *J. Opt. Soc. Am.*, **41**, 261 (1951a).
- Schulz L. G., "An Interferometric Method for the Determination of the Absorption Coefficients of Metals, With Results for Silver and Aluminum," *J. Opt. Soc. Am.*, **41**, 1047 (1951b).
- Schulz L. G. and E. J. Scheibner, "An Experimental Study of the Change in Phase Accompanying Reflection of Light from Thin Evaporated Films," *J. Opt. Soc. Am.*, **40**, 761 (1950).
- Schwider J., "Informationssteigerung in der Vielstrahlinterferometrie," *Opt. Acta*, **15**, 351 (1968).
- Schwider J., "Superposition Fringers As A Measuring Tool in Optical Testing," *Appl. Opt.*, **18**, 2364–2367 (1979).
- Shaan M. S. and J. M. Jonathan. "Multiple Beam Holographic Interferometry," *Opt. Acta*. **25**, 1025–1034 (1978).
- Shaan M. S. and V. I. Little, "The Application of Multiple-Beam White-Light Fringes to the Study of Surfaces," *J. Phys.*, **D8**, 1003 (1975).
- SIRA, "New Aid for the Optical Industry: an Interferometer for Testing Deeply Curved Lens Surfaces," *SIRA News-Sheet* ("Impact"), No. 5, 1967.
- Thorton B. S., "An Uncertainty Relation in Interferometry," *Opt. Acta*, **4**, 41 (1957).
- Tolansky S., "New Non-localized Interference Fringes," *Philos. Mag.*, **34**, 555 (1943).
- Tolansky S., "New Contributions to Interferometry. II: New Interference Phenomena with Newton's Rings; III: The Differential Polarization Phase Change on Reflection at a Thin Silver Film," *Philos. Mag.*, Ser. 7, **35**, 120, 179 (1944).
- Tolansky S., "New Contributions to Interferometry. V: New Multiple-Beam White Light Interference Fringes and Their Applications," *Philos. Mag.*, Ser. 7, **36**, 225 (1945a).
- Tolansky S., "Topography of Crystal Faces. I: The Topography of a (100) Face of a Left-Handed Quartz Crystal; II: The Topography of Cleavage Faces of Mica and Selenite," *Proc. Roy. Soc.*, **A184**, 41, 51 (1945b).
- Tolansky S., "Further Interferometric Studies with Mica: New Multiple-Beam Fringes and Their Applications," *Proc. Roy. Soc.*, **A186**, 261 (1946).

- Tolansky S., *Multiple-Beam Interferometry of Surfaces and Films*, Oxford University Press, Oxford, 1948a, Dover, New York, 1970.
- Tolansky S., *Multiple-Beam Interferometry*, Oxford University Press, Oxford, 1948b.
- Tolansky S., *Multiple-Beam Fringes*, Clarendon Press, Oxford, 1948c.
- Tolansky S., *Surface Microtopography*, Interscience, New York, 1960.
- Tolansky S., *Introduction to Interferometry*, Longmans, London, 1966.
- Tolansky S., *Microstructure of Surfaces Using Interferometry*, Edward Arnold, London, 1968.
- Tolansky S. and S. H. Emara, "Precision Multiple-Beam Interference Fringes with High Lateral Microscopic Resolution," *J. Opt. Soc. Am.*, **45**, 792 (1955).
- Vinokurov V. M., A. L. Ardamatski, and L. V. Popov, "The Surface of the Disrupted Layer," in: *Generation of Optical Surfaces*, K. G. Kumanin, Ed., Focal Library, London, 1962, Chap. I.
- Vrabel J. and E. B. Brown, "The Practice of Interferometry," *Opt. Eng.*, **14**, 124 (1975).
- Williams W. E., *Applications of Interferometry*, Methuen, London, 1950.
- Zerbino K. M., R. Torroba, N. Rodriguez, and M. Garavaglia, "Interference Profiles with Multiple Spherical Waves: General Case," *J. Opt. Soc. Am.*, **A1**(5), 495–501 (1984).

7

Multiple-Pass Interferometers

P. Hariharan

This chapter discusses some variants of the conventional interferometers used for optical testing in which one (or more) of the wavefronts is sent back and makes two or more traverses of either the whole system or a part of it. Such double- or multiple-pass interferometers offer definite advantages for some testing applications.

7.1. DOUBLE-PASS INTERFEROMETERS

7.1.1. Separation of Aberrations

The interference pattern obtained with a lens in a Twyman–Green interferometer gives a contour map of the wavefront leaving the lens aperture; however, when many aberrations are present, estimation of the individual aberrations becomes difficult. This problem can be simplified if the Twyman–Green interferometer is used in a double-pass configuration (Hariharan and Sen, 1961d) so that the symmetrical and the antisymmetrical parts of the wave aberration (see Chapter 13) are displayed in separate interferograms.

As shown in Figure 7.1, the beams emerging from the interferometer through the lens L_2 are reflected back through it by the plane mirror M_3 placed at its focus, and the double-pass beams emerging from L_1 are brought to a focus at the eye stop by the auxiliary beam divider S_2 . If the source is shifted very slightly sideways, the two images formed at the eye stop move off the axis in opposite directions, and it is possible to view either the fringes produced by the double-pass beams or the normal interference pattern.

The four double-pass rays derived from a ray incident on the beam divider S_1 at O can be identified as the AA' ray ($SOAOM_3O'A'O'S$), the AB' ray ($SOAOM_3O'B'O'S$), the BA' ray ($SOBOM_3O'A'O'S$), and the BB' ray ($SOBOM_3O'B'O'S$), corresponding to the paths they follow on the outward and return journeys. Since the wavefronts emerging from the interferometer are inverted before they are sent back, it is easily seen that if on the first pass the BB' ray traverses the pupil of the lens under test at a

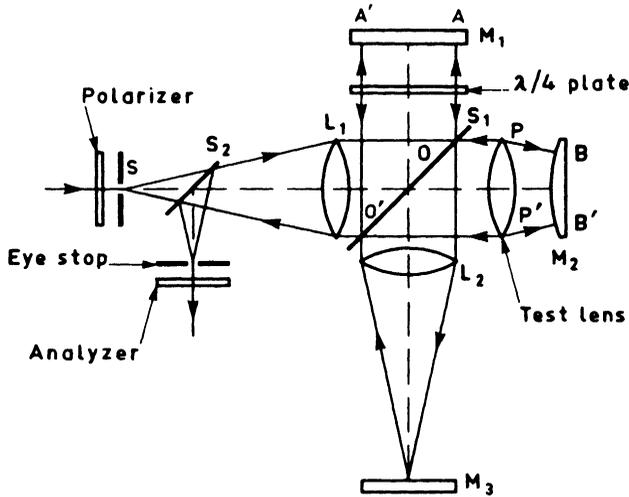


FIGURE 7.1. Double-pass Twyman-Green interferometer for separation of symmetrical and antisymmetrical wavefront aberrations. (From Hariharan and Sen, 1961d.)

point $P(r, \phi)$, it traverses the pupil on the second pass at the diametrically opposite point $P'(r, \pi + \phi)$. In the same manner, while the BA' ray traverses the pupil at P , the AB' ray traverses it at P' .

At these two points, the terms in the expressions for the total wavefront aberration (Hopkins, 1950) involving even powers of $\cos \phi$ (i.e., the defocusing, spherical aberration, and astigmatism terms) have the same value, whereas the terms involving odd powers of $\cos \phi$ (i.e., the distortion and coma terms) are of equal magnitude but have opposite signs. The total paths of the four double-pass rays can therefore be written as follows:

$$\begin{aligned}
 D_{AA'} &= D \text{ (say)} \\
 D_{BA'} &= D + 2W_{\text{even}} + 2W_{\text{odd}} + \Delta D \\
 D_{AB'} &= D + 2W_{\text{even}} - 2W_{\text{odd}} + \Delta D \\
 D_{BB'} &= D + 2W_{\text{even}} + 2W_{\text{odd}} + \Delta D + 2W_{\text{even}} - 2W_{\text{odd}} + \Delta D \\
 &= D + 4W_{\text{even}} + 2\Delta D
 \end{aligned} \tag{7.1}$$

where W_{even} is the sum of the terms involving even powers of $\cos \phi$, W_{odd} is the sum of the terms involving odd powers of $\cos \phi$ (see Appendix 3), and ΔD is the difference in the lengths of the A and B paths for the principal ray.

To select the required combinations of double-pass beams, the beam from the collimator is polarized in the vertical plane, and a quarter-wave plate is introduced in the A path. The plane of polarization of the AB' and BA' beams is then rotated through 90° , while the plane of polarization of the AA' and BB' beams remains unchanged. Hence, when the analyzer is set with its axis vertical, the AB' and BA' beams are extinguished, and interference takes place between the AA' and BB'

beams. The path difference between these is

$$D_{BB'} - D_{AA'} = 4W_{\text{even}} + 2\Delta D. \tag{7.2}$$

Accordingly, symmetric errors such as change of focus, spherical aberration, and astigmatism are shown in this interferogram with doubled sensitivity, whereas features that depend on odd powers of $\cos \phi$ are eliminated.

When the analyzer is rotated so that its axis is horizontal, the AB' and BA' rays are isolated and made to interfere. The path difference between these rays is

$$D_{BA'} - D_{AB'} = 4W_{\text{odd}} \tag{7.3}$$

Only antisymmetric aberrations such as distortion and coma are shown in this interferogram, along with any tilt of the reference mirror (which is equivalent to the introduction of an additional “odd” term).

A typical set of interferograms obtained with an uncorrected lens is shown in Figure 7.2. Figure 7.2(a) is a normal Twyman–Green interferogram, and Figures 7.2(b, c) are double-pass interferograms, which show the “even” and “odd” components of the wavefront aberration. The patterns obtained by plotting the expressions for the most important aberration terms are shown below each interferogram.

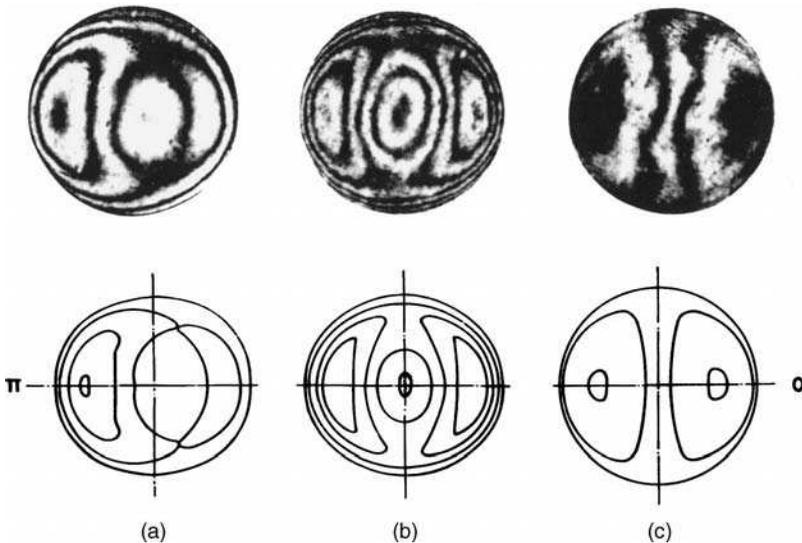


FIGURE 7.2. Interferograms obtained with the double-pass Twyman interferometer. The patterns obtained by plotting the expressions for the most important aberration terms are shown below each interferogram. (a) Normal interferogram: $2W = 0.25 + 1.85r \cos \phi - 3.99r^2 - 2.75r^2 \cos^2 \phi - 1.57r^3 \cos \phi + 4.33r^4 + 1.72r^4 \cos^2 \phi$. (b) Double-pass interferogram showing only the “even” terms: $4W_{\text{even}} = 0.50 - 7.98r^2 - 5.50r^2 \cos^2 \phi + 8.66r^4 + 3.44r^4 \cos^2 \phi$. (c) Double-pass interferogram showing only the “odd” terms: $4W_{\text{odd}} = 3.70r \cos \phi - 3.14r^3 \cos \phi$. (From Hariharan and Sen, 1961d.)

7.1.2. Reduction of Coherence Requirements

Residual wedge errors in plane parallel plates are commonly measured with the Fizeau interferometer. This has the disadvantage that the path difference between the interfering wavefronts is twice the optical thickness of the plate under test. It is therefore essential to use a well-collimated beam of monochromatic light when testing thick plates. Even with a laser source, a well-corrected objective must be used. However, this problem can be eliminated if the interferometer is double passed, so that superposition fringes are formed between the plate under test and its inverted image.

One method (Cagnet, 1954) is to use the wavefronts transmitted through the plate under test and reflect them back by means of an auxiliary afocal system. In this case, it is necessary to coat the surfaces of the plate so that their reflectivity R is fairly high. With a broadband source such as a high-pressure mercury vapor lamp, the fringes then have the same intensity distribution as those obtained with strictly monochromatic light in a conventional Fizeau interferometer with surfaces having reflectivity R^2 .

An alternative that does not require the surfaces to be coated and involves very little in the way of additional optics is to use a doubly reflected system of fringes (Sen and Puntambekar, 1965). The optical arrangement shown in Figure 7.3 is the same as that in a conventional Fizeau interferometer, except that the pinhole source is on the reflecting surface of a plane mirror. If this pinhole is shifted slightly off axis, the rays reflected back from the plate under test are brought to a focus at a point on this mirror on the other side of the axis, so that they are reflected back once more through the system. The double-pass rays are finally brought to a focus at the eye stop by a beam divider.

A ray traversing the plate under test at a point $P(x, y)$ on the first pass traverses it at $P'(-x, -y)$ on the second pass. Let $t(x, y)$ be the optical thickness of the plate at P , $\zeta = (2\pi/\lambda)[t(x, y) + t(-x, -y)]$, and $\psi = (2\pi/\lambda)[t(x, y) - t(-x, -y)]$, and assume that the spectral bandwidth of the source is such that no interference can take place for a phase difference of ζ , while the variation of ψ over this bandwidth can be

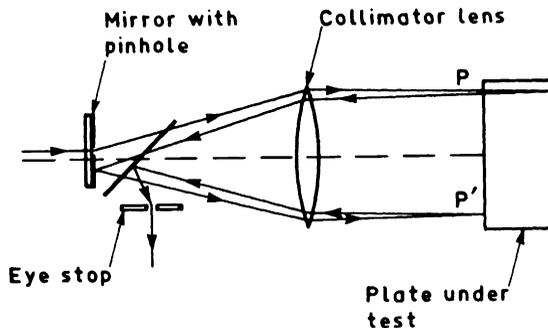


FIGURE 7.3. Double-pass Fizeau interferometer for testing plane parallel plates. (From Sen and Puntambekar, 1965.)

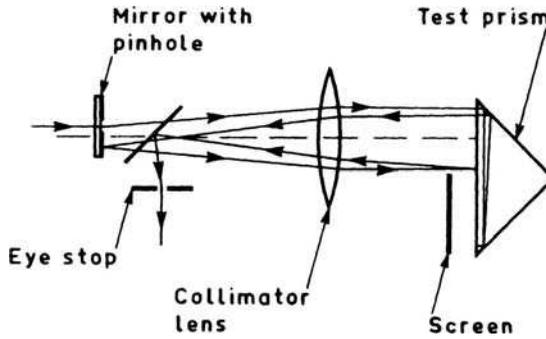


FIGURE 7.4. Double-pass Fizeau interferometer for testing reflecting prisms. (From Sen and Puntambekar, 1966.)

neglected. With an uncoated glass plate whose reflectivity R is small, the fringe irradiance at any point is then

$$I(x, y) = 2I_0R^2(2 + \cos 2\psi) \quad (7.4)$$

where I_0 is the irradiance of the original incident ray. This corresponds to the irradiance distribution in two-beam fringes with a visibility of 0.5, which is adequate for most measurements. The interferogram shows only the antisymmetrical (wedge) errors in the plate under test, for which its sensitivity is twice that of the normal Fizeau interferometer.

The double-pass Fizeau interferometer can also be used for testing right-angle prisms and cube corners (Sen and Puntambekar, 1966). The setup for this is shown in Figure 7.4. One half of the diagonal face of the prism is covered to isolate two reflected wavefronts, one formed by reflection from the uncovered half of the diagonal face and the other transmitted through the prism and reflected back by the covered half of the diagonal face. The angle between these wavefronts is doubled on the second pass, so that the sensitivity to angular errors is twice that obtainable with a Twyman–Green interferometer.

Another version of this interferometer, shown schematically in Figure 7.5, can be used to test concave surfaces (Puntambekar and Sen, 1971). With a thermal source, the distance between the reference surface and the concave mirror in a conventional Fizeau interferometer cannot be very large, so that only a limited range of curvatures can be tested with a given reference surface. This problem can be overcome if a laser source is used, but unwanted fringes due to reflection at other surfaces then appear. Double passing permits a wide range of curvatures to be tested with a single reference surface, using a thermal source.

In this case also, when the ray paths during the first and second passes are symmetrical about the optical axis of the system under test, only antisymmetrical errors appear in the interferogram. However, an additional shear can be introduced between the inverted wavefronts if the pinhole mirror is rotated about an axis on its surface passing through the common center of curvature; the symmetrical errors can

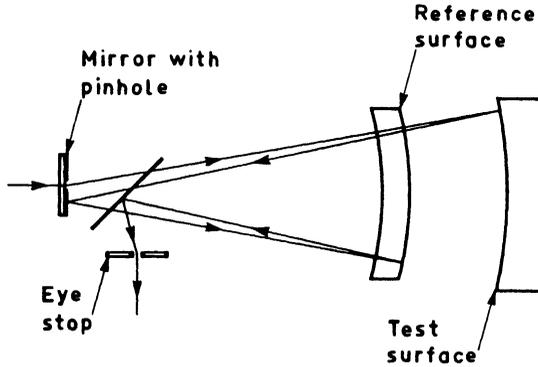


FIGURE 7.5. Double-pass Fizeau interferometer for testing concave surfaces. (From Puntambekar and Sen, 1974.)

then be evaluated from the resulting interferograms. Different parts of the reference surface are used for the sheared and unsheared interference patterns, but if the reference surface is accurately spherical, this does not introduce any error.

7.1.3. Double Passing for Increased Accuracy

The accuracy with which measurements can be made with a conventional two-beam interferometer is determined by the precision with which a local fringe displacement ΔQ can be estimated as a fraction of Q , the average interfringe spacing. This precision is limited by the sinusoidal irradiance distribution in two-beam fringes. One method to obtain greater accuracy, which was discussed in Chapter 6, is multiple-beam interferometry. Double-pass interferometry (Hariharan and Sen, 1960a, 1961b) can also be used to advantage in some cases.

Consider the double-pass Twyman-Green interferometer shown in Figure 7.6, and assume that the end mirrors M_1 and M_2 are normal to the original incident beams in the vertical plane, but make an angle equal to $(\pi/2) + \alpha$ with them in the horizontal plane, so that M_1 and M'_2 , the virtual image of M_2 in the beam divider, make an angle 2α with each other. In this case, the only aberration term present is the “odd” tilt term, and from Eq. (7.1) the total optical paths of the four double-pass rays emerging from the interferometer at a distance x from the center of the field in the horizontal plane are as follows:

$$\begin{aligned}
 D_{AA'} &= D \quad (\text{say}) \\
 D_{AB'} &= D + 4x\alpha + 2d \\
 D_{BA'} &= D - 4x\alpha + 2d \\
 D_{BB'} &= D + 4d
 \end{aligned}
 \tag{7.5}$$

where d is the distance between M_1 and M'_2 at the center of the field. If all the four beams are made to interfere, the complex amplitude A_p at this point can be

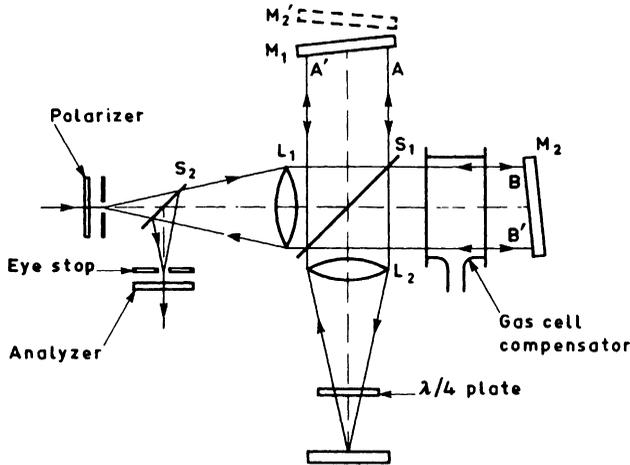


FIGURE 7.6. Twyman-Green interferometer double passed to obtain modulated fringes. (From Hariharan and Sen, 1960a.)

written as

$$A_p = A \exp - \left(\frac{2\pi i D}{\lambda} \right) [1 + \exp -2i(\eta - \xi) + \exp -2i(\eta + \xi) + \exp -4i\eta] \tag{7.6}$$

where $\xi = (2\pi/\lambda)2x\alpha$, $\eta = (2\pi/\lambda)d$, and A is the amplitude corresponding to a single beam. The irradiance I_p at this point, obtained by multiplying A_p by its complex conjugate, is then

$$I_p = 4I_0(\cos 2\xi + \cos 2\eta)^2 \tag{7.7}$$

where I_0 is the irradiance due to a single beam.

The irradiance distribution in the fringe system given by Eq. (7.7) is shown in Figure 7.7 for three different values of the phase difference η between the two paths at the center of the field. A change in η does not result in a displacement of the fringes but causes only a change in the irradiance distribution. When $\eta = 0$ or $n\pi$, where n is an integer, alternate fringes are suppressed, whereas when $\eta = (2n + 1)\pi/2$, all the fringes have the same irradiance. At the latter point, a very small change in the value of η results in an appreciable difference in irradiance between adjacent fringes. This effect can be used with a suitable compensator to achieve an accurate null setting, permitting measurements of the changes in the optical path of $\lambda/500$.

This technique is well adapted to measurement of the refractive index and thickness of thin films (Hariharan and Sen, 1961c). Even higher accuracy is possible by double passing a Fizeau interferometer (Hariharan and Sen, 1960b, 1961a).

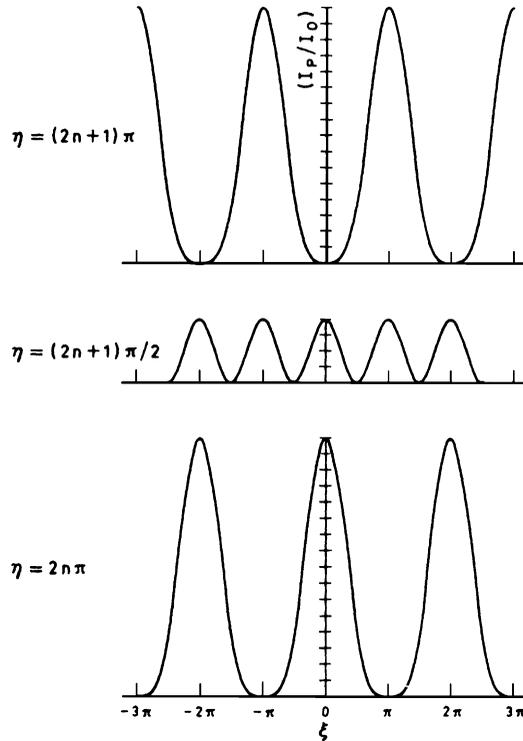


FIGURE 7.7. Irradiance distribution in the fringe pattern obtained with the double-pass Twyman–Green interferometer for various values of η , the phase difference between the two paths. (From Hariharan and Sen, 1960a.)

7.2. MULTIPASS INTERFEROMETRY

Direct visual presentation of wavefront errors with increased accuracy is obtained in multipass interferometry. In this technique, the fact that the total deformation of the wavefront is proportional to the number of times it is reflected from or transmitted through the optical system under test is utilized to increase the ratio of the fringe displacement ΔQ to the average interfringe spacing Q .

The optical system of the multipass Twyman–Green interferometer (Langenbeck, 1967) is shown schematically in Figure 7.8. An additional beam divider is inserted in one arm of the interferometer. If this makes a slight wedge angle with respect to the mirror under test, the multiply reflected beams formed between them give rise to a series of laterally separated images of the light source in the rear focal plane of lens L_2 . This permits selecting a beam that has undergone any desired number of reflections at the mirror under test by an aperture in the focal plane of L_2 . If the reference mirror is tilted so that the light reflected by it also passes through the same aperture, two-beam interference fringes are obtained between the reference wavefront and a wavefront that has undergone n reflections at the mirror under test and

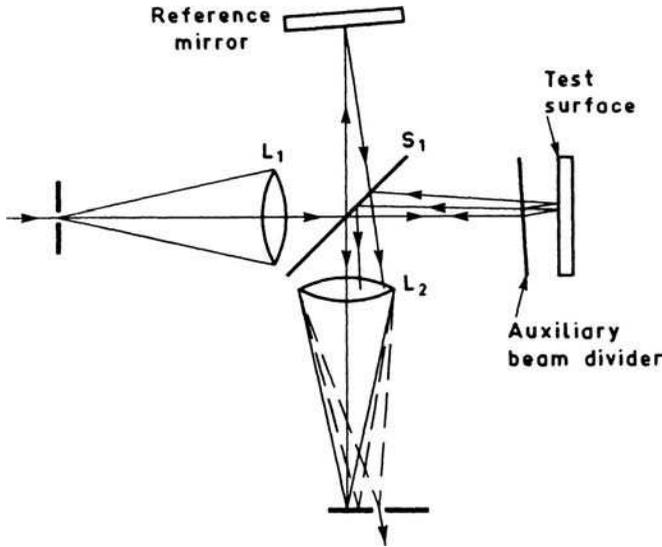


FIGURE 7.8. Multipass Twyman–Green interferometer. (From Langenbeck, 1969a.)

therefore exhibits a deformation of $2n \Delta t$, where Δt is the deviation of its surface from flatness.

In the resulting interferogram, $\Delta Q/Q = 2n \Delta t/\lambda$. Hence, the sensitivity is n times greater than that obtained with a normal Twyman–Green interferometer.

The irradiances of the two interfering beams can be made approximately equal and the visibility of the fringes optimized, if the auxiliary beam divider has a high transmittance and an uncoated glass surface is used for the reference mirror. A better solution (Langenbeck, 1969a) is to use a modified optical system in which the reference beam is reflected off an uncoated glass surface at Brewster’s angle. Rotation of the plane of polarization of the illuminating beam by a $\lambda/2$ plate makes it possible to control the attenuation of the reference beam.

A typical series of multipass interferograms of a flat surface with a square, $\lambda/8$ deep step at its center is shown in Figure 7.9. Figure 7.9(a) is the normal

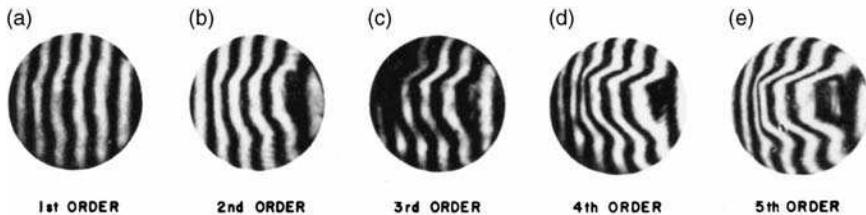


FIGURE 7.9. Interferograms obtained with the multipass Twyman–Green interferometer from a flat surface with a square, $\lambda/8$ deep step at its center. (a) Normal interferogram (first order) (b) Second order. (c) Third order. (d) Fourth order. (e) Fifth order. (From Langenbeck, 1967.)

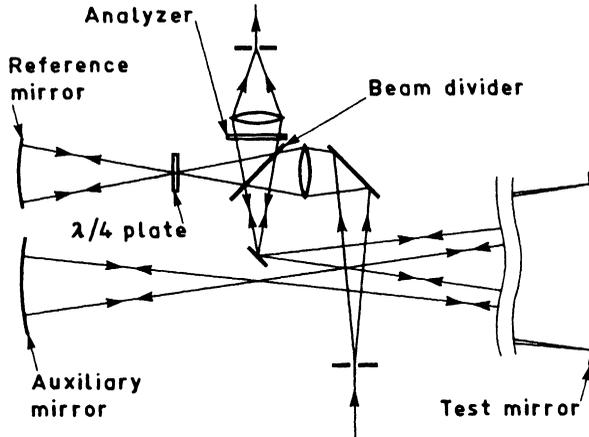


FIGURE 7.10. Multipass interferometer for testing large concave surfaces. (From Bubis, 1972).

Twyman–Green interferogram of the first-order reflection, and Figure 7.9 (b–e) are the second-, third-, fourth-, and fifth-order multipass interferograms, respectively, showing the increase in the relative fringe displacement $\Delta Q/Q$ with the order of the reflection.

A modified multipass interferometer for testing large concave surfaces has been described by Bubis (1972). This is essentially an unequal-path interferometer using a laser source in which an auxiliary concave mirror is introduced in the measuring arm.

As can be seen from Figure 7.10, the converging beam from the laser is split at the beam divider into two beams, one of which is reflected back from a small concave mirror to form the reference wavefront. The other beam undergoes a predetermined number of reflections in the system comprising the mirror under test, and the auxiliary concave mirror before it is reflected back to the beam divider to interfere with the reference wavefront. (The number of reflections that the beam undergoes can be controlled by the relative offsets of the centers of curvature of the mirror under test and the auxiliary mirror.) A mica $\lambda/4$ plate in the comparison arm and a Polaroid analyzer are used to equalize the amplitudes of the interfering beams. A fivefold increase in sensitivity has been obtained.

The Fizeau interferometer can also be operated in a multipass configuration (Langenbeck 1969a, 1969b) if the zero-order reflected beam and one of the higher order reflected beams are made to interfere. For this purpose, the wedge angle between the surfaces is made large enough that the individual orders of reflection can be separated in the focal plane. As shown in Figure 7.11, a stop with two openings is then used to select the zero-order and a suitable higher-order beam. An auxiliary lens whose front focal plane coincides with the stop gives interference fringes between nominally plane wavefronts.

The relative amplitudes of the interfering beams are equalized by a simple polarizing system. The zero-order beam is reflected from the uncoated part of a partially coated mirror at the Brewster angle, while the higher order beams are reflected from the adjacent metal-coated part. Rotation of the plane of polarization of

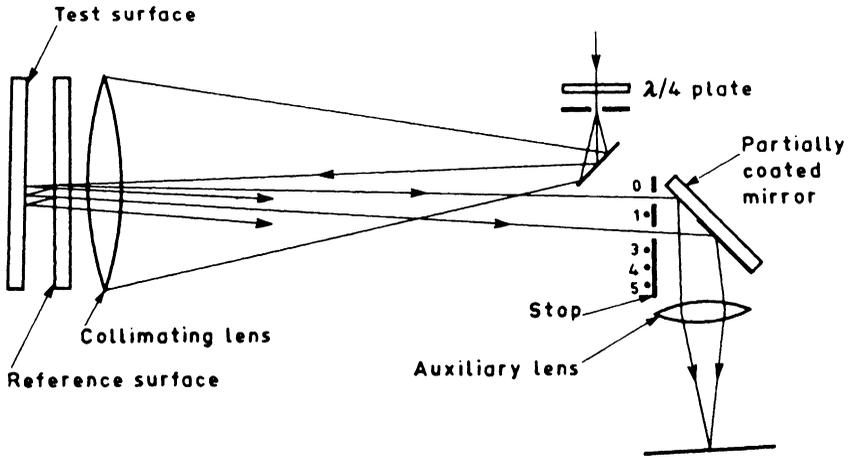


FIGURE 7.11. Multipass Fizeau interferometer. (From Langenbeck, 1969a.)

the illuminating beam with a $\lambda/2$ plate then permits attenuating the zero-order beam to any required degree.

Because of the relatively large angle between the zero-order beam and the n th-order beam, the resulting interference fringes are too narrow to be viewed directly, and the interference pattern is best observed by a moiré technique. For this, a grating with a line spacing equal to the average fringe spacing is introduced into the image plane. The moiré fringes seen have the same spacing; their spacing can be controlled by rotating the grating with respect to the interferogram.

Higher order moiré fringes can also be obtained by using a grating with a line spacing equal to $1/m$ the average fringe spacing. This gives m times as many moiré fringes across the field and a relative sensitivity m times as high, though at the expense of fringe contrast.

The most serious problem in multipass interferometry is beam “walk-off.” This integrates the errors over a finite region of the surface under test and can lead to fringe shifts even in areas where there is no deformation of the test surface. The extent of this walk-off is determined by the distance separating the multipass mirrors and by the wedge angle. It can be minimized by tilting the interferometer with respect to the incoming beam so that the beam, which is initially reflected toward the apex of the wedge in the test path, is finally reflected at normal incidence by the auxiliary beam divider and returns along the same path. However, in most cases, the order of reflection must be chosen so that the walk-off is small compared to the lateral dimensions of the defects under examination.¹

¹This limitation does not apply in some cases, such as thickness measurements of uniform thin films, for which specialized optical systems using a narrow beam and permitting as many as 50 passes have been described by Dupoisot and Lostis (1973).

Misinterpretation of multipass interferograms can also result from misalignment of the stop in the focal plane. If this stop accidentally cuts off low spatial frequencies in the image, strong diffraction effects are observed, as in the Foucault test.

Applications where multipass interferometry is useful include measurements of the homogeneity of glass plates and the precise determination of 90° or 45° angles on roof prisms and cube corners, where high accuracy is required (Langenbeck, 1969b). Another interesting application is in grazing incidence interferometry. Grazing incidence interferometry (Linnik, 1941; Saunders and Gross, 1959) makes it possible to test large surfaces with a system of relatively small aperture. In addition, it is possible to test many nonoptical surfaces, such as fine-ground glass and metal, from which a specular reflection of adequate intensity can be obtained at low angles of incidence (Birch, 1973; Hariharan, 1975; Murty, 1976).

With incidence at an angle α , a beam of diameter d covers a strip of length

$$L = \frac{d}{\cos \alpha} \quad (7.8)$$

However, a deviation Δt of the test surface from flatness results in a deviation of the fringes from straightness given by the relation

$$\frac{\Delta Q}{Q} = \frac{4\Delta t \cos \alpha}{\lambda} \quad (7.9)$$

where ΔQ is the local fringe displacement and Q is the average spacing of the fringes. As a result, the sensitivity of the interferometer is decreased by a factor proportional to the cosine of the angle of incidence. The sensitivity can be increased by a factor of 2, even with ground surfaces having a low reflectance, by using a double-pass system in which the reflectance of the beam splitter is adjusted to equalize the intensity of the two beams (Wilson, 1983). If the reflectivity of the test surface is adequate, it is even possible to use the multiple pass arrangement shown in Figure 7.12 (Langenbeck, 1969a). In this case, for an angle of incidence α , the deviation of the fringes from straightness in the fringe pattern obtained with the n th-order reflected beam is

$$\left(\frac{\Delta Q}{Q}\right)_n = \frac{4n \Delta t \cos \alpha}{\lambda}, \quad (7.10)$$

so that the loss in sensitivity due to the high angle of incidence can be regained by choosing a suitable higher-order beam.

Another interesting application of multiple-pass interferometry is accurate measurement of the radius of curvature of a concave surface with a very long radius of curvature, where the working distance available is limited (Gerchman and Hunter, 1979).

In this procedure, a reference plane surface and the test concave surface are set up to form a confocal cavity. Collimated light entering such a cavity through the plane

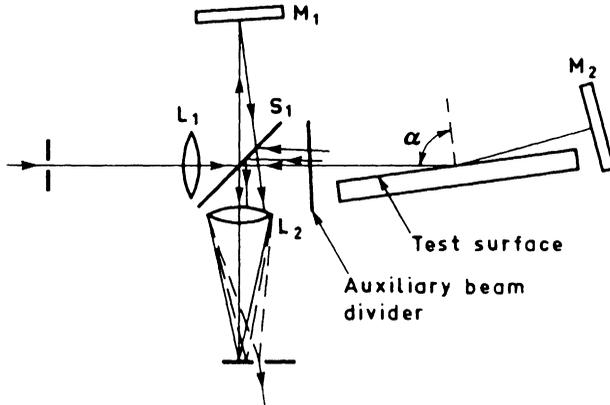


FIGURE 7.12. Grazing incidence multipass interferometer used for testing large surfaces with an optical system of small aperture. (From Langenbeck, 1969a.)

surface is brought to a focus as shown in Figure 7.13, on one or the other of the surfaces. In such a confocal configuration, a collimated beam that has undergone the appropriate number of reflections is reflected back as a collimated beam. The focus setting can therefore be determined accurately by observing the interference fringes formed by the reflected beam from the plane surface and the reflected beam from the confocal cavity, using a setup similar to a Fizeau interferometer.

Configurations in which the order number n is even bring the light to a focus on the plane surface, while configurations in which n is odd focus the light on the concave surface. The equations relating z_n , the separation of the surfaces in the n th-order configuration, and R , the radius of curvature of the concave surface, can be obtained from a paraxial analysis. This analysis shows that the expressions for the odd configurations are similar to the even-order Chebyshev polynomials of the first

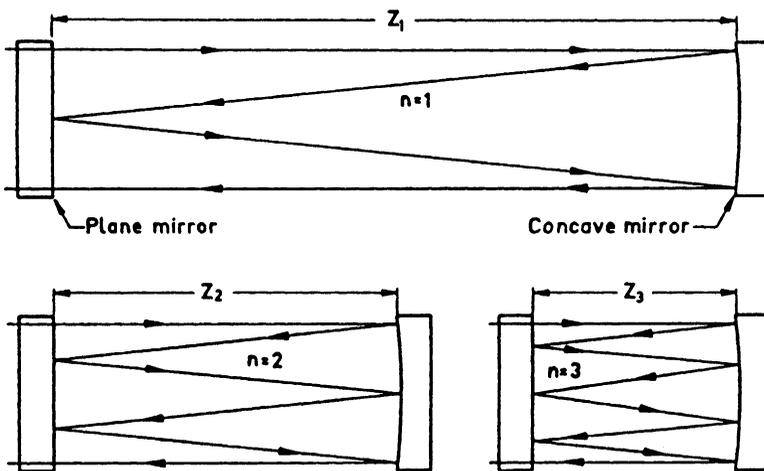


FIGURE 7.13. Ray paths in confocal cavities of orders 1, 2, and 3. (From Gerchman and Hunter, 1979.)

TABLE 7.1. Cavity length and radius of curvature
 $z_n = C_n R$.

n	C_n
1	0.5
2	0.25
3	0.146,446,6
4	0.095,491,5
5	0.066,987,3
6	0.049,515,6
7	0.038,060,3
8	0.030,153,7
9	0.024,471,7

kind, $T_{n+1}(x)$, ($n = 1, 3, 5, \dots$), where $x = z_n/R$. Similarly, the expressions for the even configuration are similar to the even-order Chebyshev polynomials of the second kind, $U_n(x)$, ($n = 2, 4, 6, \dots$). In both the cases, the solutions of interest are the first positive root.

Accordingly, if the change in the separation of the mirrors between two successive confocal configurations is measured, it is possible to calculate the radius of the concave surface. Table 7.1 lists the values of z_n , the cavity length for successive confocal configurations, as a function of R , the radius of curvature of the concave surface, while Table 7.2 lists $z_n - z_{n+1}$, the corresponding differences in the cavity length between successive configurations.

As can be seen, the separation of the surfaces in the higher order configurations is much smaller than the radius of curvature of the concave surface, while the distance through which it has to be moved between successive configurations is even smaller. As a result, surfaces with very long radii can be measured in a limited working space using a short measuring slide.

The accuracy with which measurements can be made is limited by two factors. One is the accuracy with which the position of best focus can be estimated; the other is the accuracy with which the movement of the mirror can be measured. A limit is also imposed on the highest order test configuration by the drop in the visibility of the fringes. To optimize the visibility for higher order configurations, the plane surface

TABLE 7.2. Difference in cavity lengths and radius of curvature.

$R = 4(z_1 - z_2)$
$R = 9.656,85(z_2 - z_3)$
$R = 19.625,12(z_3 - z_4)$
$R = 35.082,55(z_4 - z_5)$
$R = 57.235,25(z_5 - z_6)$
$R = 87.295,84(z_6 - z_7)$
$R = 126.477,41(z_7 - z_8)$
$R = 175.994,37(z_8 - z_9)$

should have a reflectivity of about 0.6, while the concave surface should have a highly reflecting coating. Acceptable fringe visibility can then be obtained even with an $n = 9$ configuration, and measurements of the radius of curvature can be made with an accuracy better than 0.01%.

REFERENCES

- Birch, K. G., "Oblique Incidence Interferometry Applied to Non-Optical Surfaces," *J. Phys. E.*, **6**, 1045 (1973).
- Bubis I. Y., "Multipass Interferometer for Surface Shape Inspection," *Sov. J. Opt. Technol.*, **39**, 411–413 (1972).
- Cagnet M., "Méthodes Interférométriques Utilisant les Franges de Superposition (voir Erratum, p. 552)," *Rev. Opt.*, **33**, 113 (1954).
- Dupoisot H. and P. Lostis, "Interférométrie à Passages Multiples et Mesure des Très Faibles Epaisseurs," *Nouv. Rev. Opt.*, **4**, 227 (1973).
- Gerchman M. C. and G. C. Hunter, "Differential Technique for Accurately Measuring the Radius of Curvature of Long Radius Concave Optical Surfaces," *Proc. SPIE*, **192**, 75 (1979).
- Hariharan P., "Improved Oblique-Incidence Interferometer," *Opt. Eng.*, **14**, 257 (1975).
- Hariharan P. and D. Sen, "Double-Passed Two-Beam Interferometers," *J. Opt. Soc. Am.*, **50**, 357–361 (1960a).
- Hariharan P. and D. Sen, "The Double-Passed Fizeau Interferometer," *J. Opt. Soc. Am.*, **50**, 999–1001 (1960b).
- Hariharan P. and D. Sen, "Double-Passed Fizeau Interferometer. II: Fringe Systems Formed by the Reflected Beams," *J. Opt. Soc. Am.*, **51**, 400–404 (1961a).
- Hariharan P. and D. Sen, "Fringes of Equal Inclination in the Double-Passed Michelson Interferometer," *J. Opt. Soc. Am.*, **51**, 617–619 (1961b).
- Hariharan P. and D. Sen, "Double-Passed Two-Beam Interferometers, II: Effects of Specimen Absorption and Finite Path Difference," *J. Opt. Soc. Am.*, **51**, 1212–1218 (1961c).
- Hariharan P. and D. Sen, "The Separation of Symmetrical and Asymmetrical Wave-Front Aberrations in the Twyman Interferometer," *Proc. Phys. Soc.*, **77**, 328 (1961d).
- Hopkins H. H., *Wave Theory of Aberrations*, Clarendon Press, Oxford, 1950, p. 48.
- Langenbeck P., "Multipass Twyman–Green Interferometer," *Appl. Opt.*, **6**, 1425–1426 (1967).
- Langenbeck P., "Multipass Interferometry," *Appl. Opt.*, **8**, 545–552 (1969a).
- Langenbeck P., in: *Optical Instruments and Techniques*, J. Home Dickson, Ed., Oriel Press, Newcastle-Upon-Tyne, 1969b.
- Linnik V. P., "An Interferometer for the Investigation of Large Plane Surfaces," *Computes Rendus (Doklady) Acad. Sci. URSS*, **32**, 189 (1941).
- Murty M. V. R. K. and Shukla R. P., "An Oblique Incidence Interferometer," *Opt. Eng.*, **15**, 461–463 (1976).
- Puntambekar P. N. and D. Sen, "A Simple Inverting Interferometer," *Opt. Acta*, **18**, 719–728 (1971).
- Saunders J. B. and F. L. Gross., "Interferometer for Large Surfaces," *J. Res. Natl. Bur. Stand.*, **62**, 137 (1959).

- Sen D. and P. N. Puntambekar, "An Inverting Fizeau Interferometer," *Opt. Acta*, **12**, 137-149 (1965).
- Sen D. and P. N. Puntambekar, "Shearing Interferometers for Testing Corner Cubes and Right Angle Prisms," *Appl. Opt.*, **5**, 1009-1014 (1966).
- Wilson I. J., "Double-Pass Oblique Incidence Interferometer for the Inspection of Non-optical Surfaces," *Appl. Opt.*, **22**, 1144-1148 (1983).

8

Foucault, Wire, and Phase Modulation Tests

J. Ojeda-Castañeda

8.1. INTRODUCTION

The knife-edge method, introduced by Foucault (1858, 1859), and, in general, all the Schlieren techniques—Töpler (1864, 1866, 1868), Ritchey (1904), Hartmann (1908), Platzek and Gaviola (1939), Wolter (1949 and 1956)—have proved to be extremely useful for testing optical surfaces. Rays may be moved from their expected trajectories (or wavefronts may be deformed) by optical aberrations, by diffraction, or by a deformed, unfinished surface. The basic idea behind the Schlieren techniques is to detect lateral displacements of rays by blocking out or modifying these displaced rays. The blocking or modification can be accomplished by placing screens in any of the planes of convergence of the light passing through, or being reflected from, the optical surface under test.

The main advantages of the Schlieren techniques are their high sensitivity and their simplicity, both in apparatus and in qualitative interpretation (at least from the geometrical point of view). Of course, to appreciate the relative merits of one method over others in this class, it is necessary to study its characteristics when it is used to detect the presence of (a) aberrations greater than the wavelength of the illuminating radiation (geometrical theory of image formation) and (b) aberrations smaller than the wavelength (physical theory of image formation).

The choice of test depends on the circumstances.

8.2. FOUCAULT OR KNIFE-EDGE TEST

8.2.1. Description

The knife-edge test may be considered, in general, as a method for detecting the presence of transversal aberrations. This is done by blocking out one part of a plane traversed by rays or diffracted light so that a shadow appears over the aberrated region as shown in Figure 8.1. This high simplicity of operation and interpretation

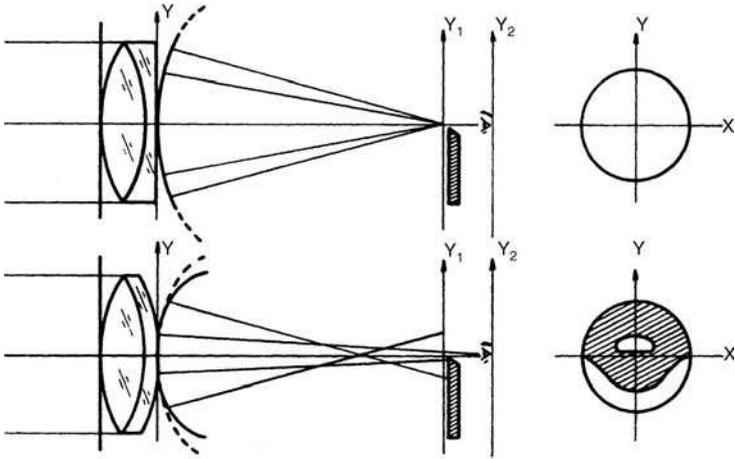


FIGURE 8.1. Foucault method for testing a lens.

makes the Foucault test unique among optical testing methods. This test may be well considered as the first optical shop test, from which many other optical tests were developed.

The Foucault test consists in placing an illuminated pinhole size source at one side of the center of curvature of a spherical concave mirror as shown in Figure 8.2. The image of this illuminated pinhole size source has the shape of a point source, localized on the other side of the center of curvature of the spherical mirror.

When a knife edge is introduced, cutting the illuminating beam that forms the image, an observer (placed behind the image) sees a shadow pattern appearing over the otherwise uniformly illuminated surface of the mirror (see Figures 8.1 and 8.3). In the case of a “perfect” spherical mirror, when the knife edge is introduced *inside the focus* (toward the mirror), the shadow pattern consists of a dark region and a bright region, sharply separated as shown in Figures 8.4 and 8.5. As the knife is moved, the dark region appears to move across the mirror in the *same direction* in which the knife edge moves. On the contrary, when the knife edge is introduced *outside the focus* (away from the mirror), the dark region of the shadow pattern

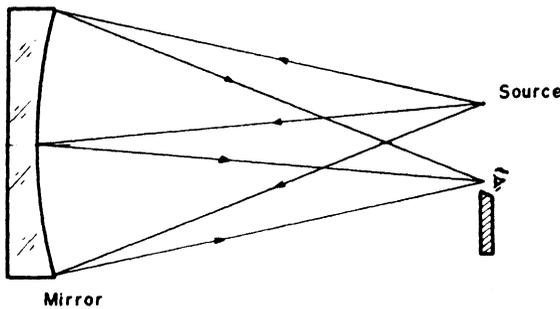


FIGURE 8.2. Knife-edge setup when testing a concave mirror.

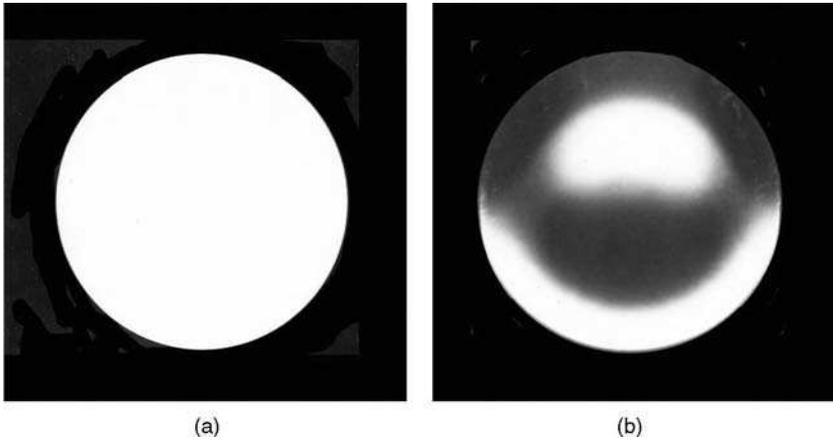


FIGURE 8.3. Foucault graphs of a parabolic mirror, (a) Before introducing the knife edge (b) After introducing the knife edge.

moves across the mirror in a direction *opposite* to that in which the knife edge moves. When the knife edge is introduced precisely *at the focus*, the mirror appears to *darken suddenly*, without any apparent motion of the shadow pattern. This is a very accurate way of finding the center of curvature.

Moreover, if the concave mirror is not spherical, each zone over the mirror has a different radius of curvature. In this case, when the knife edge is introduced, the corresponding zone darkens for different positions of the knife edge along the optical axis.

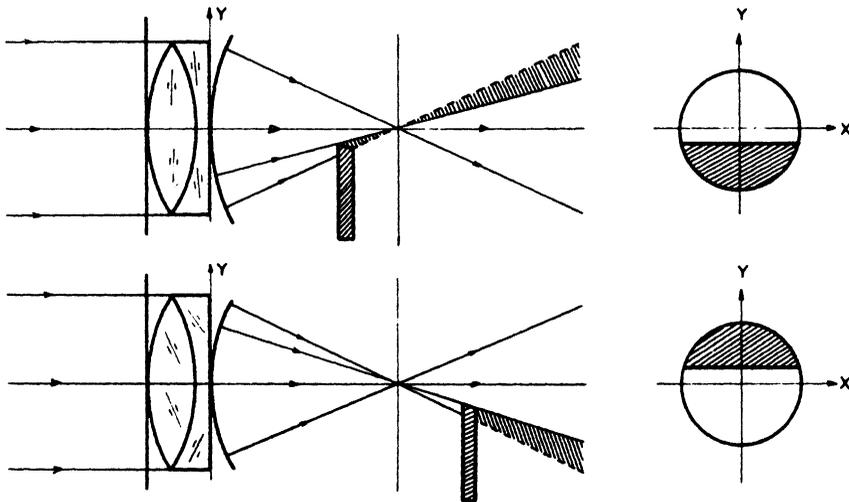


FIGURE 8.4. Knife-edge testing of an aberration-free lens.

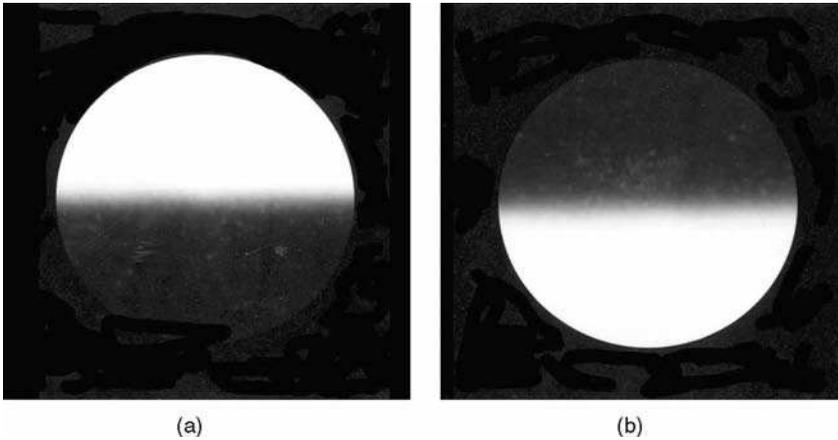


FIGURE 8.5. Foucault graphs of an aberration-free mirror with focus errors. (a) Inside of focus. (b) Outside of focus. Knife introduced from below.

For an irregular concave mirror there are many zones with different radii of curvature and different centers of curvature. In this case, when the knife edge is introduced, many different zones all over the mirror are darkened. An intuitive way to visualize how this shadow pattern is formed (given by Porter, 1953) is to imagine an illuminating source at one side of the mirror, at the opposite side of the knife edge, as shown in Figure 8.6. With this kind of illumination, it follows that the regions with positive slope are illuminated and those with negative slope are not illuminated or vice versa as shown in Figure 8.7.

To summarize, the Foucault test can be used (1) to measure the radius of curvature of each zone over the mirror and (2) as a *null test*, for checking that all the zones of a supposedly spherical wavefront have the same radius of curvature.

It has been found in practice that the Foucault test may be done perfectly by employing an illuminating slit instead of a point source as shown in Figures 8.8(b and c). The reason is that each point of the slit source produces an image with the same shape and at the same distance from the knife edge. Therefore, all of the points produce identical Foucault patterns, resulting in a tremendous gain in the brightness of the observed pattern. The patterns, however, must be exactly identical if the accuracy is to be preserved. This means that the slit and the knife edge must be exactly parallel. To avoid the problem of setting the knife edge exactly parallel to the slit, the arrangement in Figure 8.8d was suggested by Dakin (1967).

The use of an extended source, in the form of a slit, was previously used in flow visualization (Wolter, 1956; Stolzenburg, 1965) rather than in optical testing. In the previous field, it is also common to employ white-light sources for encoding optical path differences in terms of color variations; see for example Banerji (1918), Schardin (1942a, 1942b), North (1954), Wolter (1956), Weinberg (1961), and O'Hare (1969). The Foucault test in white light has been analyzed by Bescos and Berriel-Valdos (1986).

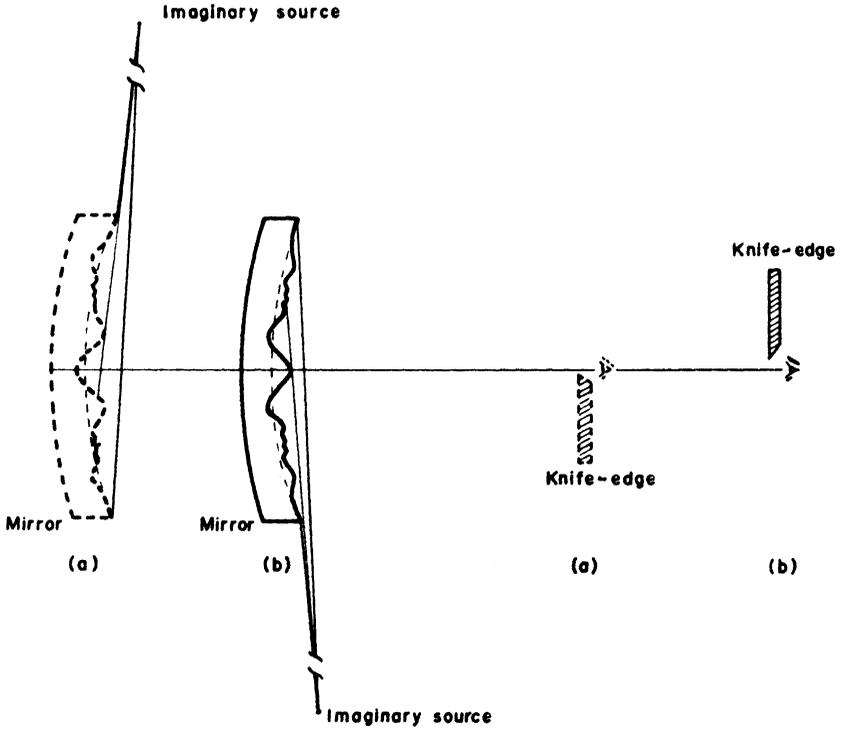


FIGURE 8.6. Visualization of the shadow patterns formed when using the knife-edge test. Notice that the imaginary light source and the knife edge are at opposite sides of the optical axis.

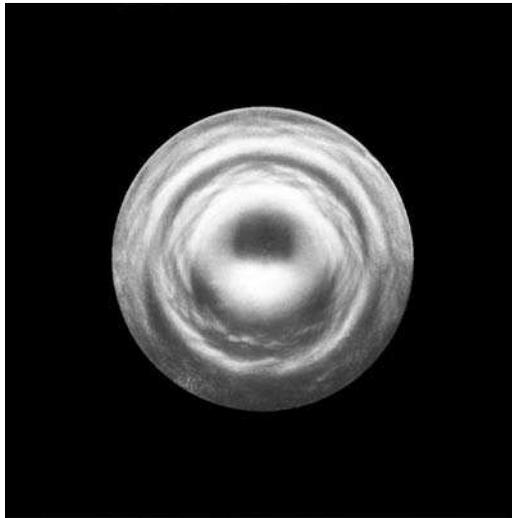


FIGURE 8.7. Foucault graph of an irregular mirror.

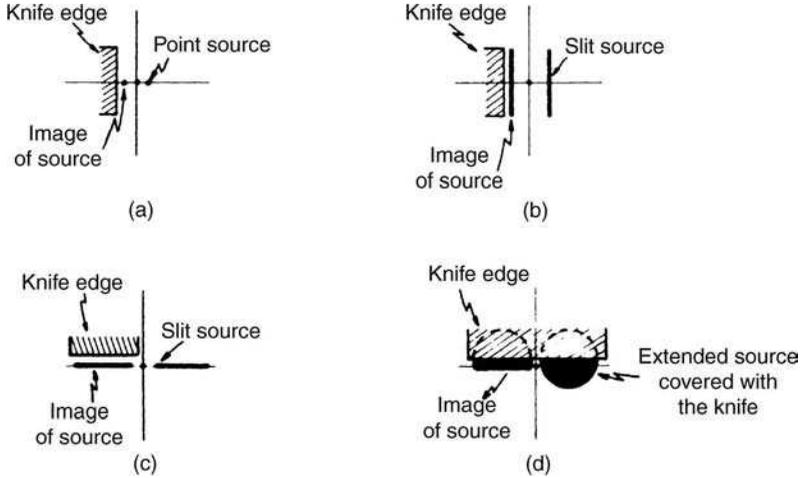


FIGURE 8.8. Different arrangements for the source and the knife edge.

Figures 8.9 and 8.10 show an instrument to perform the Foucault test. The instrument is mounted in a carriage capable of translating it, along the optical axis or perpendicularly to the optical axis, laterally or up and down.

8.2.2. Geometrical Theory

The border of the knife in the Foucault test is placed at a distance r_1 from the chief ray intersection (the origin of the X_1 - Y_1 plane), and an angle ϕ_1 is subtended between the Y_1 axis and the knife edge as shown in Figure 8.11. The angle ϕ_1 will be defined as positive if the slope of the knife edge is positive. The following equation defines the border:

$$x_1 \cos \phi_1 - y_1 \sin \phi_1 = r_1 \quad (8.1)$$

The transmittance over this plane may be expressed as

$$T(x_1, y_1) = \begin{cases} 1 & \text{if } x_1 \cos \phi_1 - y_1 \sin \phi_1 < r_1 \\ 0 & \text{if } x_1 \cos \phi_1 - y_1 \sin \phi_1 \geq r_1 \end{cases} \quad (8.2)$$

But since the X_1 - Y_1 plane defines the paraxial plane of convergence of the ideal wavefront W , any point (x_1, y_1) over this plane satisfies, approximately, the following property (Rayces, 1964):

$$(x_1, y_1) = \left(-R \frac{\partial W}{\partial x}, -R \frac{\partial W}{\partial y} \right) \quad (8.3)$$

where R is the distance between the $X = Y$ and the $X_1 = Y_1$ plane.

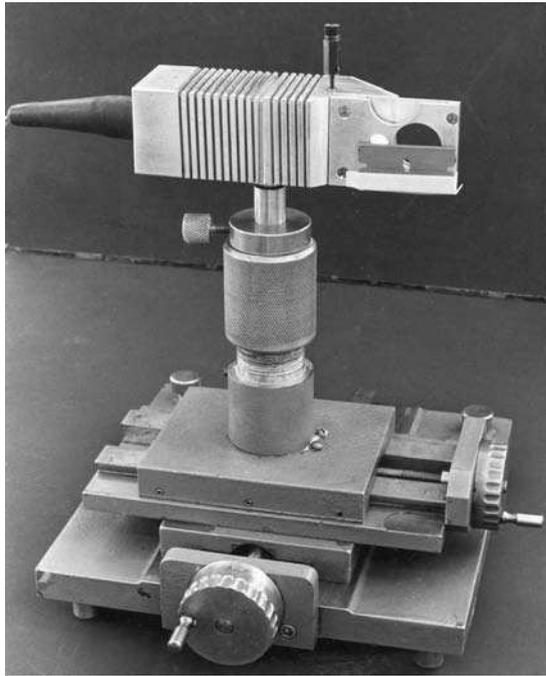


FIGURE 8.9. Photograph of the Foucault apparatus.

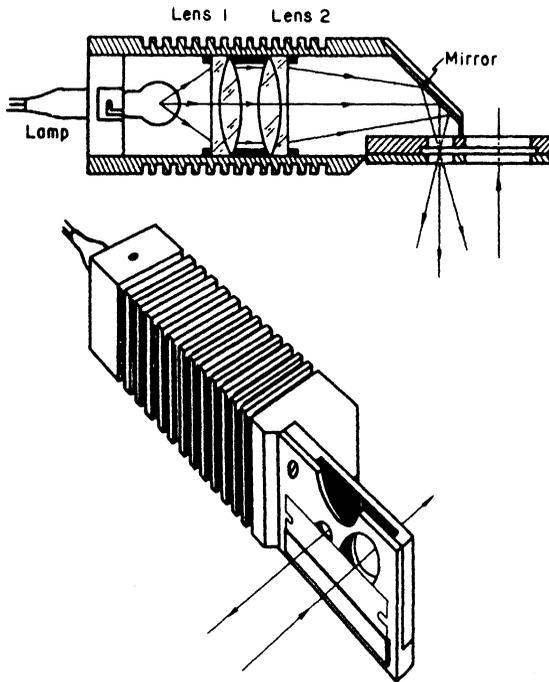


FIGURE 8.10. Diagrams of the Foucault apparatus.

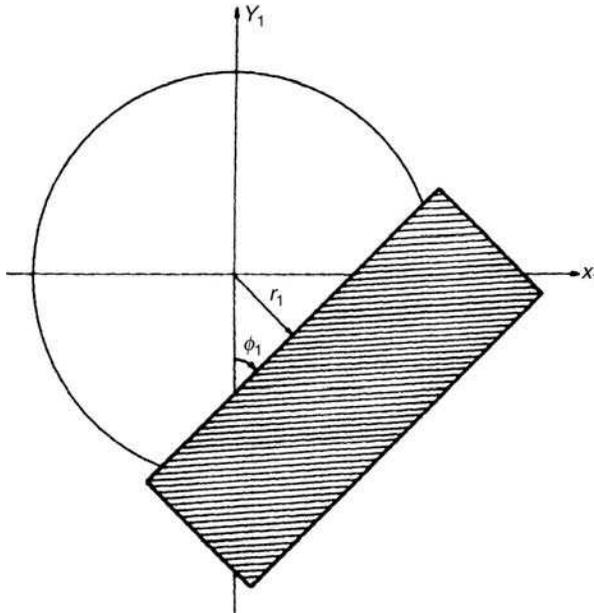


FIGURE 8.11. Knife-edge position projected over the entrance pupil plane of the viewing system.

Using this result in Eq. (8.2), the transmittance function becomes

$$T\left(\frac{\partial W}{\partial x}, \frac{\partial W}{\partial y}\right) = \begin{cases} 1 & \text{if } -\frac{\partial W}{\partial x} \cos \phi_1 + \frac{\partial W}{\partial y} \sin \phi_1 < \frac{r_1}{R} \\ 0 & \text{if } -\frac{\partial W}{\partial x} \cos \phi_1 + \frac{\partial W}{\partial y} \sin \phi_1 \geq \frac{r_1}{R} \end{cases} \quad (8.4)$$

For the particular case of aberrations with rotational symmetry, any axis over the X_1 - Y_1 plane may be used as reference to specify the position of the knife edge. It proves convenient, for our purposes, to choose the X axis, so that $\phi_1 = 90^\circ$. Thus Eq. (8.4) is reduced to

$$T\left(\frac{\partial W}{\partial y}\right) = \begin{cases} 1 & \text{if } \frac{\partial W}{\partial y} < \frac{r_1}{R} \\ 0 & \text{if } \frac{\partial W}{\partial y} \geq \frac{r_1}{R} \end{cases} \quad (8.5)$$

Focus Error. Let us consider our first example, where we have an aberration-free optical surface, but the knife is placed at some distance from the converging

X_1 - Y_1 plane. In other words, only focus error is present in the wavefront aberration function, that is,

$$W(x, y) = D(x^2 + y^2) \quad (8.6)$$

From Eq. (8.5) it follows that the boundary between dark and bright regions (bright region, $T = 1$ if $y_1 < r_1$) may be expressed as

$$y_1 = \frac{r_1}{2DR} \quad (8.7)$$

when $D \neq 0$. Hence, depending on whether the knife is placed inside the focus (D negative) or outside the focus (D positive), the shadow along the Y axis (i.e., parallel to the X axis) will change sides as shown in Figure 8.5, from the negative side of the Y axis to the positive side of the Y axis.

This situation is precisely the one described before when indicating on which side of a "perfect" spherical mirror the dark region appears when the knife edge is introduced.

A particular case occurs when the knife edge touches the optical axis. The shadow pattern will be a dark half-circle, the other half-circle will be bright. In other words, the boundary between the bright and the dark regions is at $y_1 = 0$ when $r_1 = 0$ for any $D \neq 0$.

Primary Spherical Aberration. When both primary spherical aberration and focus errors are present, the aberration function is given by

$$W(x, y) = A(x^2 + y^2)^2 + D(x^2 + y^2)$$

Thus, using Eq. (8.5), the equation of the borders of the shadow pattern may be written as

$$y^3 + \left(\frac{D}{2A} + y^2\right)y - \frac{r_1}{4AR} = 0 \quad (8.8)$$

It is to be noted that since the equation has dependence on x , the boundary of the shadow will not, in general, be limited by straight lines.

The boundaries of the shadow pattern along the Y axis for this case may be found by letting $x = 0$ in Eq. (8.8), which may then be written as

$$y^3 + \frac{D}{2A}y - \frac{r_1}{4AR} = 0 \quad (8.9)$$

Being a cubic equation, it has three roots. Of course, we are interested only in the roots that are real numbers. Defining the parameter Δ as

$$\Delta = \left(\frac{r_1}{8AR}\right)^2 + \left(\frac{D}{6A}\right)^3 \quad (8.10)$$

we have, from a known result from algebra, the following:

1. If $\Delta < 0$, there will be three real and unequal roots.
2. If $\Delta = 0$, there will be three real roots of which at least two are equal.
3. If $\Delta > 0$, there will be only one real root and two conjugate imaginary roots.

Consequently, the shadow pattern will show more than one dark region if conditions 1 and 2 are satisfied. This happens when the defocusing coefficient and the primary aberration coefficient have different signs. In other words, by letting $\Delta \leq 0$, we can find from Eq. (8.10) that

$$\left(\frac{r_1}{R}\right)^2 \leq \frac{\left(-\frac{2}{3}D\right)^3}{A} \quad (8.11)$$

Now, since the left-hand term is always positive, it follows that when the knife is inside the focus (D negative) and there is positive primary aberration, or, equivalently, when the knife is outside the focus and there is negative primary aberration, the shadow pattern will show two dark regions.

Again, a particular case of this result occurs when the knife edge is touching the optical axis. In this case $r_1 = 0$ and Eq. (8.9) becomes

$$y\left(y^2 + \frac{D}{2A}\right) = 0 \quad (8.12)$$

solutions of which are

$$y = 0 \quad (8.13a)$$

and

$$y = \pm \sqrt{-\frac{D}{2A}} \quad (8.13b)$$

It is, indeed, clear that D and A must have different signs in order to obtain real numbers as solutions of Eq. (8.13b). In Figures 8.1, 8.12, and 8.13, Foucault graphs for spherical aberration are shown. For different treatments for spherical aberration, the reader can refer to Conrady (1924) and Kingslake (1937).

Primary Coma. In the case of primary coma and focus error, the aberration function is given by

$$W(x, y) = By(x^2 + y^2) + D(x^2 + y^2) \quad (8.14)$$

Since primary coma does not have radial symmetry, we will consider the two cases in which the knife edge is being displaced along the X_1 axis and the Y_1 axis,

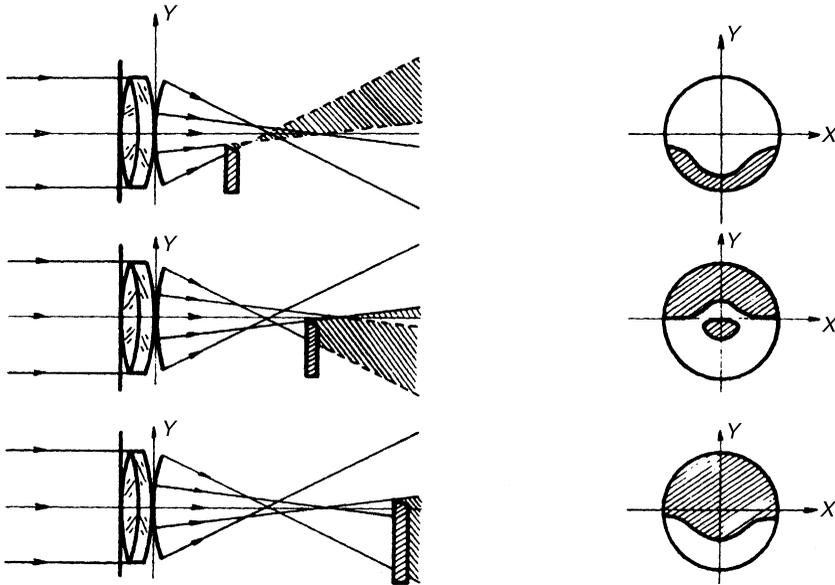


FIGURE 8.12. Knife-edge testing of a lens with spherical aberration.

respectively. Using the same procedure employed in the preceding examples, we have, in the case in which the knife edge is placed at a point on the X_1 axis at a distance r_1 from the origin (parallel to the Y_1 axis, $\phi_1 = 0$) that the shadow pattern satisfies the following equation:

$$2Bxy + 2Dx = \frac{-r_1}{R} \tag{8.15}$$

which may usefully be written as

$$\left(y + \frac{D}{B}\right)x = \frac{-r_1}{2RB} \tag{8.16}$$



FIGURE 8.13. Foucault graphs of an aspherical mirror.

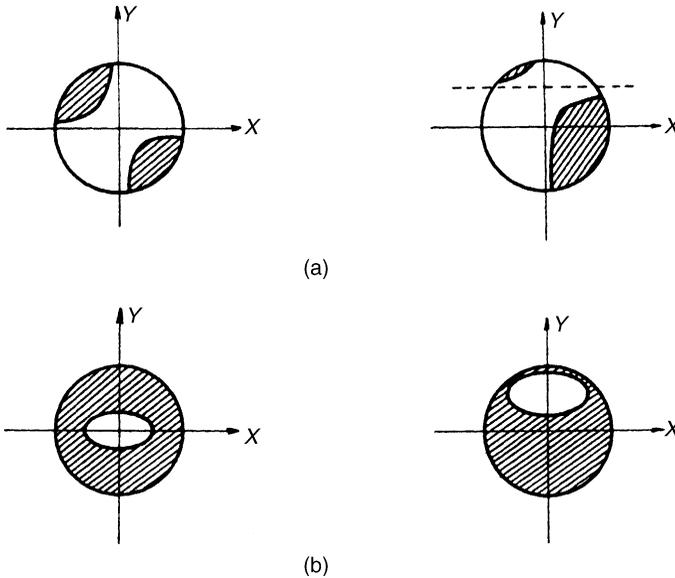


FIGURE 8.14. Diagrams showing the patterns associated with comatic aberrations. (a) Knife edge parallel to the y_1 axis. (b) Knife edge parallel to the x_1 axis.

It is clear, then, that the shadow pattern consists of rectangular hyperbolas, centered at the point $[O, -(D/B)]$. See Figure 8.14(a).

On the contrary, when the knife edge is introduced at a point on the Y_1 axis (parallel to the X_1 axis, $\phi_1 = 90^\circ$), the partial derivative of the wavefront aberration function with respect to y is

$$\frac{\partial W}{\partial y} = B(x^2 + 3y^2) + 2yD \quad (8.17)$$

Hence the borders of the shadow pattern may be obtained from the following equation:

$$x^2 + 3\left(y + \frac{D}{3B}\right) = \frac{r_1}{RB} + \frac{1}{3}\left(\frac{D}{B}\right)^2 \quad (8.18)$$

which is the equation for an ellipse centered at point $[O, -(D/3B)]$. The major axis of the ellipse will be parallel to the X axis and the minor axis will be along the Y axis. See Figure 8.14(b).

The common procedure employed in the knife-edge test is to place the knife in the paraxial plane. Thus, the shadow patterns will be centered at the origin of the optical surface under test. Under this condition, a scale is placed over the shadow pattern, and either the parameters of the rectangular hyperbola or the ellipse may be readily obtained. In other words, the knife edge is located at the paraxial plane,

$D = 0$. Hence, when the knife is oriented parallel to the Y_1 axis, in the Gaussian plane, we know from Eq. (8.16) that the shadow pattern will be described by

$$yx = \frac{-r_1}{2RB} \quad (8.19)$$

and since r_1 , the knife-edge position along the X_1 axis, and the distance R between the optical surface under test and the paraxial plane are known parameters, the value of B may be immediately obtained by measuring the ratio $r_1/2RB$.

In a similar manner, when the knife edge is oriented parallel to the X_1 axis, in the Gaussian plane, the shadow pattern in Eq. (8.18) may be written as

$$x^2 + 3y^2 = \frac{r_1}{RB} \quad (8.20)$$

Here, again, by knowing the parameters r_1 and R , and measuring the ratio r_1/RB , the amount of coma present may be determined.

Astigmatism. If the optical surface under test has astigmatism, and (in order to incorporate the position of the knife edge in any plane) we add focus error to the wavefront aberration function, we have

$$W(x, y) = C(x^2 + 3y^2) + D(x^2 + y^2) \quad (8.21)$$

As in the case of coma, since astigmatism does not have radial symmetry, we need to consider the effect of introducing the knife edge in a particular direction. In what follows, we consider the cases in which the knife edge is introduced along an axis that subtends an angle ϕ_1 with the Y_1 axis. Since the partial derivative of the wavefront aberration function with respect to x is

$$\frac{\partial W}{\partial x} = 2Cx + 2Dx \quad (8.22a)$$

and with respect to y is

$$\frac{\partial W}{\partial y} = 6Cy + 2Dy \quad (8.22b)$$

the border of the shadow pattern can be obtained from Eq. (8.4) to give

$$-(C + D)x \cos \phi_1 + (3C + D)y \sin \phi_1 = \frac{r_1}{2R} \quad (8.23)$$

It is clear, then, that an optical surface with astigmatism will have a straight line with slope ψ , given by

$$\tan \psi = \frac{3C + D}{C + D} \tan \theta_1 \quad (8.24)$$

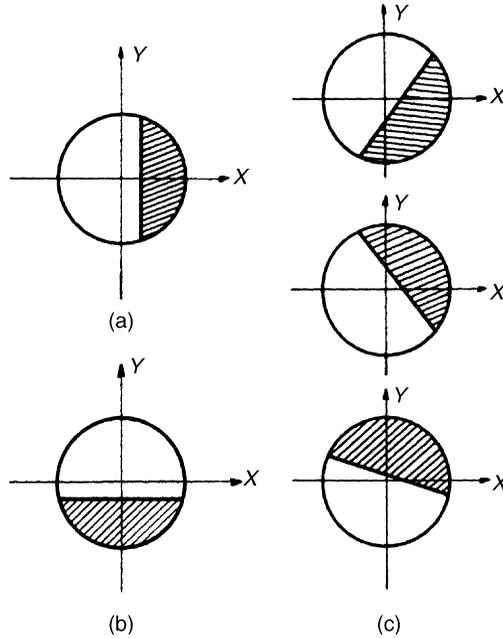


FIGURE 8.15. Diagrams showing the shadow patterns associated with astigmatism. (a) Knife edge parallel to the y_1 axis. (b) Knife edge parallel to the x_1 axis. (c) Knife edge along an axis making an angle ϕ with the y_1 axis.

From Eq. (8.24) it follows that, if the knife edge (oriented parallel to the Y_1 axis, i.e., $\phi_1 = 0^\circ$) is introduced along the X_1 axis, the boundary of the shadow pattern consists of a straight line parallel to the Y axis as seen in Figure 8.15(a), namely,

$$x = \frac{-r}{2R(C+D)} \quad (8.25)$$

and $\psi = 90^\circ$. On the contrary, when the knife edge (oriented parallel to the X_1 axis, i.e., $\phi_1 = 90^\circ$) is introduced along the Y_1 axis, the boundary of the shadow pattern consists also of a straight line parallel to the X_1 axis, as seen in Figure 8.15(b) namely,

$$y = \frac{r_1}{2R(3C+D)} \quad (8.26)$$

and $\psi = 90^\circ$. Employing Eq. (8.24), it proves convenient to write the slope associated with the angle subtended between the knife edge and the boundary of the shadow pattern as in the Ronchi test (see Chapter 9), that is,

$$\tan(\psi - \phi_1) = \frac{C \sin 2\phi_1}{D + C(2 - \cos 2\phi_1)} \quad (8.27)$$

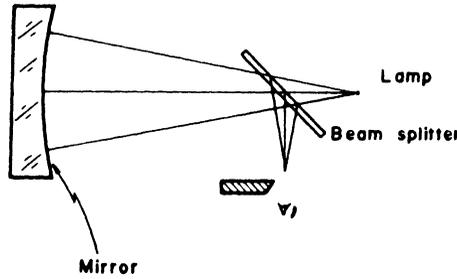


FIGURE 8.16. Foucault setup for testing on the optical axis.

From this it follows that in the paraxial plane ($D = 0$), the angle between the knife edge and the boundary of the shadow pattern is independent of the amount of astigmatism present in the wavefront, leaving the optical surface under test.

Moreover, as the knife is displaced along the optical axis (i.e., a variation in D), the slope of the boundary of the shadow pattern varies when ϕ_1 is different from 0° or 90° . Consequently, a rotation of the shadow pattern, as shown in Figure 8.15, is observed when the optical surface under test suffers from astigmatism.

It is worth noting that, when a concave mirror with a small f number is tested with a pinhole size source placed off axis (as shown in Figure 8.2), the mirror, even when perfect, appears to give rise to an astigmatic wavefront (see Chapter 18). To avoid this effect of false astigmatic error, the mirror can be tested on axis using a beam splitter as shown in Figure 8.16.

8.2.3. Physical Theory

As indicated in the introduction of this chapter, we restrict in terms of the diffraction theory to the case in which the aberrations of the illuminating radiation. A reader interested in a more general treatment is referred to the publications by Rayleigh (1917), Banerji (1918), Zernike (1934a, 1934b, 1942), Gascoine (1944), Linfoot (1945a, 1945b, 1946a, 1946b, 1946c, 1955), and Barakat (1969).

To find the observed Foucault pattern, let us first define the pupil function $A(x, y)$ describing the complex amplitude at the exit pupil of the system under test as

$$A(x, y) = A_0(x, y) \exp \left[i \frac{2\pi}{\lambda} W(x, y) \right] \approx A_0(x, y) \left(1 + i \frac{2\pi}{\lambda} W(x, y) \right) \quad (8.28)$$

where, assuming uniform illumination of the exit pupil of the system under test

$$A_0(x, y) = \begin{cases} 1 & \text{if } x^2 + y^2 \leq S_{max}^2 \\ 0 & \text{if } x^2 + y^2 > S_{max}^2 \end{cases} \quad (8.29)$$

S_{max} is the semidiameter of the exit pupil and the wavefront deformations $W(x, y)$ are measured with respect to a reference sphere with the center at the intersection

of the optical axis with the selected image plane (any selected plane near the paraxial image). We can see that if the wavefront deformations $W(x, y)$ are smaller than the wavelength of the light, the complex amplitude can be considered as the superposition of two wavefronts, (a) one with amplitude equal to one and phase zero as in a perfect aberration free system and (b) a wavefront with a variable amplitude directly proportional to $W(x, y)$ and a constant phase equal to $\pi/2$.

Let us now calculate the complex amplitude $B(f_x, f_y)$ at the observation plane, located at the center of curvature of the reference sphere, without the presence of the knife for the time being, which is given by the Fourier transform of the pupil function, as

$$B(f_x, f_y) = \mathfrak{F}\{A(x, y)\} = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} A(x, y) \exp[i2\pi(f_x x + f_y y)] dx dy \quad (8.30)$$

where r is the radius of curvature of the reference sphere and f_x and f_y are the spatial frequencies in the Fourier plane located at the center of curvature of the reference sphere. These spatial frequencies are related to the coordinates x_1 and y_1 in this plane by

$$f_x = \frac{x_1}{\lambda r} \quad \text{and} \quad f_y = \frac{y_1}{\lambda r} \quad (8.31)$$

The inverse Fourier transform of this complex amplitude will retrieve back the complex amplitude at the exit pupil of the system, which can be written as

$$\begin{aligned} A(x, y) &= \mathfrak{S}^{-1}\{B(f_x, f_y)\} = \mathfrak{S}^{-1}\{\mathfrak{F}\{A(x, y)\}\} \\ &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} B(f_x, f_y) \exp[-i2\pi(f_x x + f_y y)] df_x df_y \end{aligned} \quad (8.32)$$

However, if we introduce at the plane (x_1, y_1) a phase and amplitude modulator (filter) defined by a complex transmittance $M(f_x, f_y)$, the filtered complex amplitude $A'(x, y)$ at the exit pupil plane is

$$A'(x, y) = \mathfrak{S}^{-1}\{M(f_x, f_y)B(f_x, f_y)\} \quad (8.33)$$

Another method to calculate $A'(x, y)$ is by the convolution of the pupil function $A(x, y)$ which is the Fourier transform of $B(f_x, f_y)$ with the Fourier transform $G(x, y)$ of the filtering function $M(f_x, f_y)$, which is given by

$$G(x, y) = \mathfrak{S}^{-1}\{M(f_x, f_y)\} \quad (8.34)$$

Thus, by using the convolution theorem

$$A'(x, y) = A(x, y) * G(x, y) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} A(u, v) G(u - x, v - y) du dv \quad (8.35)$$

where the symbol $*$ represents the convolution. This expression as well as Eq. (8.33) are valid to obtain the complex amplitude at the image of the pupil of the system

under test, for any magnitude of the wavefront deformation and for any kind of filter at the plane of the Fourier transform of the pupil function.

In the Foucault test the filtering object is a knife edge, which is parallel to the x_1 axis. Thus, we may write the filtering function as

$$M(x_1, y_1) = \begin{cases} 1 & y_1 > -d \\ 0 & y_1 < -d \end{cases} \tag{8.36}$$

and hence $G(x, y)$ is given by

$$G(x, y) = \frac{1}{2} \left[\delta(x, y) - i \frac{1}{xy} \delta(x) \exp\left(-i \frac{2\pi yd}{\lambda r}\right) \right] \tag{8.37}$$

With these general tools we will now describe two particular interesting cases:

(a) *There are no aberrations and the knife edge is at the center of the Airy disc ($d = 0$).*

In this case, using Eqs. (8.28) and (8.37) in Eq. (8.35), the complex amplitude at the observation plane can be written as

$$A'(x, y) = \frac{1}{2} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} A_0(u, v) \left[\delta(u - x, v - y) - i \frac{1}{\pi(y - v)} \delta(u - x) \right] dudv \tag{8.38}$$

Using the properties of the Dirac function this expression reduces to

$$A'(x, y) = \frac{1}{2} A_0(x, y) - \frac{i}{2\pi} \int_{-\infty}^{\infty} A_0(x, v) \frac{1}{(y - v)} dv \tag{8.39}$$

The second term, which is Hilbert transform, contains the desired information, since $A_0(x, y)$ is equal to one inside of the pupil; after integration of this term,

$$A'(x, y) = \text{Const} \ln \left| \frac{y + (S_{\max}^2 - x^2)^{1/2}}{y - (S_{\max}^2 - x^2)^{1/2}} \right| \tag{8.40}$$

where \ln indicates the natural logarithm. From this expression we can conclude that a bright ring should be observed around the pupil of the system under test, known as the Rayleigh diffracting ring, which should not be interpreted as a turned edge or to be confused with any other real error. In practice, the brightness of this ring is attenuated by the finite size of the pupil of the observing system as shown by Welford (1970).

(b) *The aberrations are small, but sufficiently large so that the knife edge covers part of the image, although it is far from the optical axis so that the Airy disc size is not covered.*

If the wavefront is not spherical as assumed in the preceding section, but aberrated and the wavefront aberration $W(x, y)$ is small compared with the wavelength, then by using the approximation in Eq. (8.28) in Eq. (8.30), we can obtain $B(f_x, f_y)$ as

$$\begin{aligned} B(f_x, f_y) &= \Im\{A_0(x, y)\} + i\frac{2\pi}{\lambda}\Im\{A_0(x, y)W(x, y)\} \\ &= \Im\{A_0(x, y)\} + iB_2(f_x, f_y) \end{aligned} \quad (8.41)$$

where

$$B_2(f_x, f_y) = \frac{2\pi}{\lambda}\Im\{A_0(x, y)W(x, y)\} \quad (8.42)$$

Thus, from Eq.(8.33) and taking into account that the knife edge is not covering the central part of the image with a diameter equal to the Airy disc

$$A'(x, y) = A_0(x, y) + i\Im^{-1}\{M(f_x, f_y)B_2(f_x, f_y)\} \quad (8.43)$$

If we write this expression as

$$A'(x, y) = A_0(x, y) + i(W'_{real}(x, y) + iW'_{imag}(x, y)) \quad (8.44)$$

since the observed intensity over the Foucault pattern is given by

$$I(x, y) = A'(x, y)A'^*(x, y) \quad (8.45)$$

the observed intensity becomes

$$I(x, y) = A_0^2(x, y) + W'^2_{imag}(x, y) + W'^2_{real}(x, y) - 2A_0(x, y)W'_{imag}(x, y) \quad (8.46)$$

but since the values of W'_{imag} and W'_{real} are very small, we can approximate this expression by

$$I(x, y) \approx \text{Const} - 2A_0(x, y)W'_{imag}(x, y) \quad (8.47)$$

It is clear, the, that the intensity variations in the image plane show the presence of aberrations, provided that a modification is done over the Fourier spectrum of pupil function in the image plane, filtering with the knife edge, in order to obtain $W'_{imag}(x, y) \neq 0$. The contrast of the image intensity variations over the uniform background is

$$\gamma = \frac{4\pi}{\lambda} W'_{imag}(x, y) \quad (8.48)$$

The minimum contrast that the naked eye can detect is $\gamma = 2\%$, so it is possible to see that the theoretical limit to the sensitivity of the Foucault test is

$$\frac{\lambda}{200\pi} = W'_{imag}(x,y) \tag{8.49}$$

which is in good agreement with $\lambda/600$, as reported by Texereau (1957) when he checked the sensitivity of the method by comparing it with the Lyot (1946) phase contrast technique.

A detailed mathematical interpretation of how the image irradiance is related to the aberration over the wavefront under test has been published by Ojeda-Castañeda (1978, 1979, 1980a, 1980b).

8.3. WIRE TEST

From Section 8.2, it is clear that the knife-edge method is very sensitive in detecting zonal errors; however, the Foucault test is less sensitive in measuring the errors. Moreover, Foucault graphs are not very useful in testing nonspherical mirrors, since small details are lost, immersed on the mirror asphericity.

To measure the mirror asphericity, Ritchey (1904) suggested using a screen with zonal slots over any aspherical surface, when testing mirrors by the Foucault method. In this way it is possible to calculate the amount of transversal ray aberration of each zone, when it just becomes dark for a given position of the knife edge along the optical axis; see Figure 8.16. It is also possible to compare the illumination of the two slots of the zone, which are localized at opposite sides of the optical axis, to check whether the regions behind the slots are regular and to match between them.

Because parabolic surfaces were the most commonly used mirrors, a screen for testing these optical surfaces was designed to allow illumination in only the zones over the mirror: the central (paraxial) region, the rim (marginal region), and an intermediate region between them, which are associated with the two ends and the center of the caustic (see Appendix 1), respectively Figure 8.17(a). Once the screen

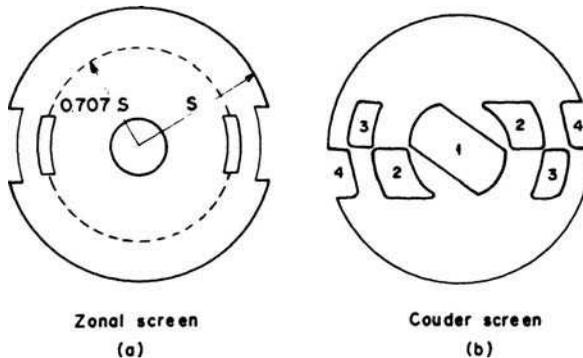


FIGURE 8.17. (a) Zonal and (b) Couder screens for zonal Foucault testing.

has been placed over the mirror, the knife edge can be used to find the places along the optical axis in which the rays, coming from two opposite ends of a given zone, converge (Gaviola, 1936). In other words, by finding each of the places where the three zones darken, we can localize the two extremes of the interval of the caustic as well as its center (Simon, 1976).

Since the amount of transverse aberration of the normals (associated with the presence of spherical aberration in a paraboloid tested at the center of curvature) is given by cS^2 , assuming that the light source and the knife move together, where again $S^2 = x^2 + y^2$ is the radius of any zone over the mirror and $c = 1/\text{radius of curvature}$ of the paraxial region of the paraxial region of the mirror, the radius of the intermediate region in the mask has to be $S = 0.707 S_{\max}$ if we want the ideal transverse aberration of the zone to be half the ideal transverse aberration at the edge, in order to intersect the optical axis at the center of the caustic.

The width of the slots over the screen must be found experimentally. The slots must be narrow enough to increase local precision but wide enough to avoid diffraction effects.

Couder (1932) designed a more versatile screen [see Figure 8.17(b)] in which only the width of the exterior zone (marginal region) has to be experimentally found; the radii and the width of the inner zones can be obtained from the facts that the zones immediately adjoin each other, and that the difference between the square of the radii of successive zones is a constant. In other words, if the experimental value of the radial width for the rim is δS , the difference between the outside and inside radii of this n -fold zone is

$$S_n^2 - S_{n-1}^2 = (\delta S)^2 \quad (8.50)$$

since the inside radius of the n -fold slot is the outside radius of the $(n - 1)$ -fold region. Consequently, the inside radius of the $(n - 1)$ -fold region, or the outside radius of the $(n - 2)$ -fold region, is given by

$$S_{n-1}^2 - S_{n-2}^2 = (\delta S)^2 \quad (8.51)$$

and in this way the radii of all the zones can be determined.

The advantage of this screen is that the whole mirror is observed and the overall shape of a surface can be tested. On the contrary, the disadvantage of any screen test is that the error in the measurement is the order of magnitude of the tolerance to which the optical surface is to be polished (Schroader 1953; Simon 1976). Therefore, a finer test for aspherical surfaces is required, and it is in these cases that the wire test and the Platzcek–Gaviola test (Section 8.4) can be usefully employed. The wire test is applicable only to aspherical surfaces, conical or not, and works in a manner similar to that of the zonal knife-edge test, by blocking out the deviated rays (or diffracted light) across a very narrow region. For this purpose, a thin opaque wire is placed sequentially near the intersections of the normals of various annular zones with the optical axis of the mirror or surface under test (see Fig. 8.18). In this way, a particular intersection of the normals to a particular annulus on the optical surface is

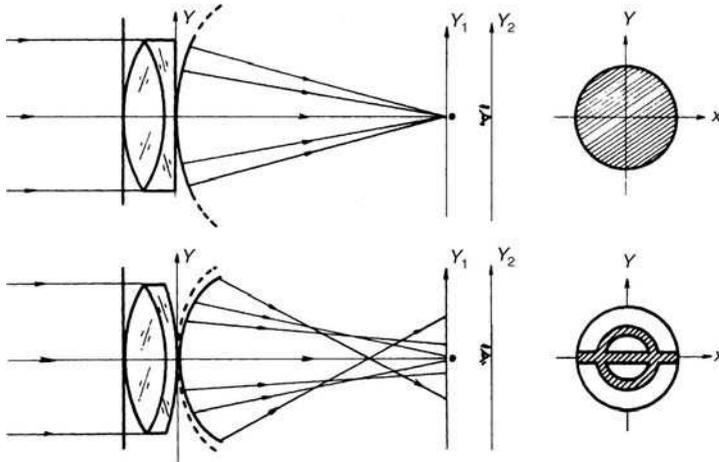


FIGURE 8.18. Wire setup for testing a lens.

found when the annulus in question appears dark against a uniformly bright background. In practice, this is done by displacing the wire together with the pinhole size source (i.e., displacing the plane in which both the wire and the pinhole size source are contained). The pinhole size source does not need to be monochromatic.

The main advantage of the wire test over the zonal knife-edge method, used when masking the optical surface under test, is the ability of the former to restrict the blocking-out action to a very narrow region at the annulus being tested. In addition, diffraction effects are symmetrical with respect to the annulus. As a result, measurement of the intersection of the normals with the optical axis is more precise, and in this way the deviations between the experimental values of the intersections and the theoretical values can be measured. In other words, the aberrations of the normals are determined (see Fig. 8.19).

Moreover, the wire test has the advantage that the measurements of the intersections of the normals with the optical axis are referred to the theoretical value of the intersection of the first or paraxial annulus. This feature can prove to be extremely useful when testing an optical blank with a central hole.

Since the region covered by the wire corresponds only to the edge of the knife in the Foucault test, the shadow patterns obtained with the wire test consist only of thin, dark contours as shown in Figure 8.20. Compare these wire patterns with the corresponding Foucault patterns in Figure 8.21.

The wire test proves to be a good technique for testing an aspherical optical surface that is being figured; that is, it is an effective method of checking the longitudinal ray aberration, from which it is possible to obtain the wavefront aberration by using a numerical integration technique (Smith, 1966; Fell, 1968; Loomis, 1968; Buchroader et al., 1972a, 1972b). The main disadvantage of the wire test is that the width of the shadow pattern sometimes masks small displacement of the whole shadow while the wire is displaced along the optical axis.

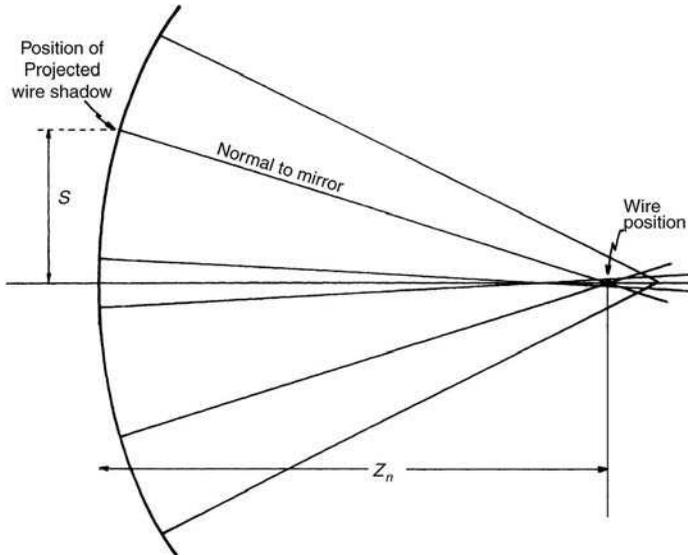


FIGURE 8.19. Diagram showing the wire method.

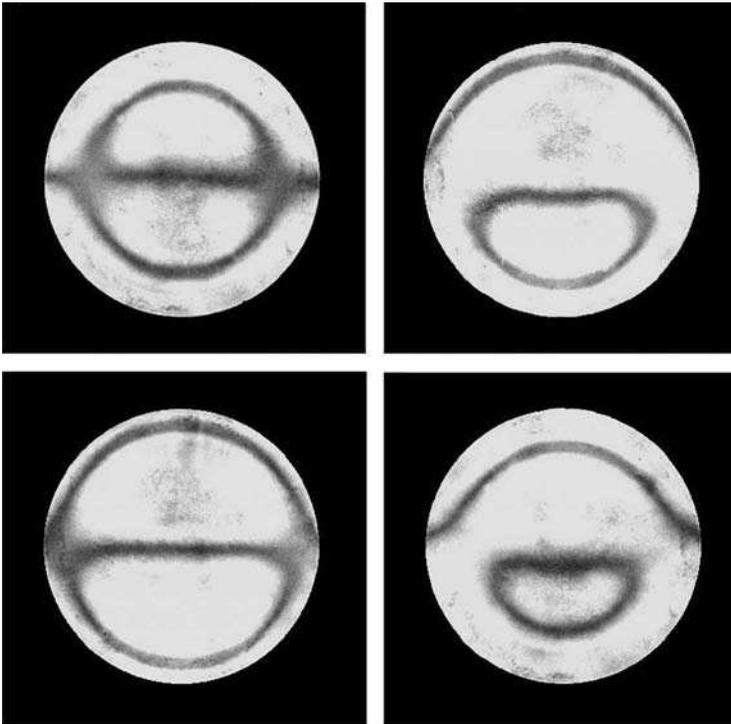


FIGURE 8.20. Shadow patterns obtained with the wire test.

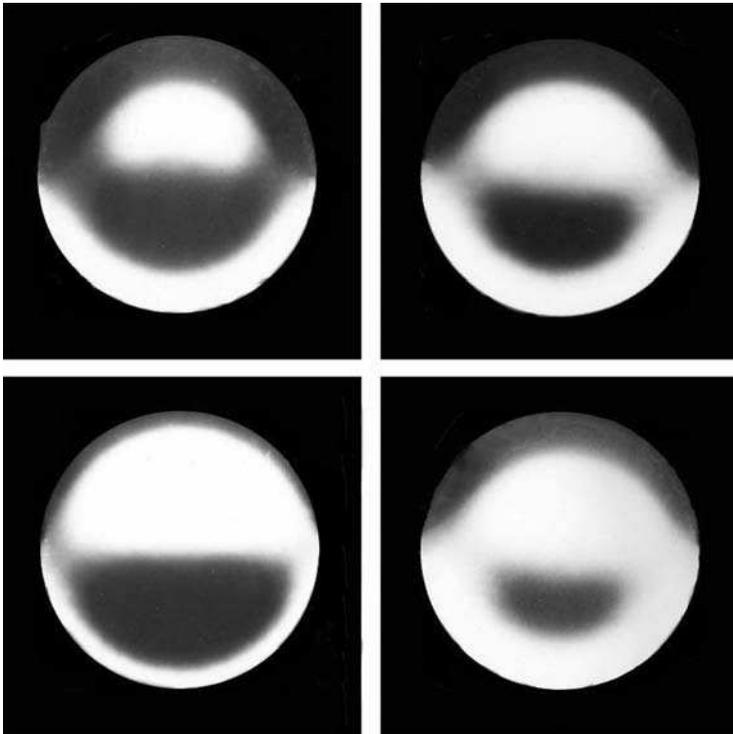


FIGURE 8.21. Shadow patterns obtained with the knife-edge test. Notice the similarity to the patterns in Figure 8.20.

Meinel (1968) indicated that the wire test can also be used for testing astigmatism by determining whether or not the annular shadows are circular.

8.3.1. Geometrical Theory

If the position of the central part of the wire is denoted by r_1 and the thickness of the wire by $2\delta r$, the amplitude transmittance in the paraxial plane is given by

$$T(x_1, y_1) = \begin{cases} 1 & \text{if } |x_1 \cos \phi_1 - y_1 \sin \phi_1| > \delta r \\ 0 & \text{if } |x_1 \cos \phi_1 - y_1 \sin \phi_1| < \delta r \end{cases} \quad (8.52)$$

The center of the dark contour obtained with the wire test coincides with the boundary between the dark and bright regions of the shadow pattern obtained with the Foucault test. Of course, since the wire does not have the covering extension of the knife, the dark areas are reduced to dark lines along the boundary positions.

In this section, we describe the procedure used for obtaining the wavefront aberration of the optical surface under test from the experimental measurement of the intersection of the normals with the optical axis. Let us denote the experimental

and theoretical values of the intersection of the normal of the optical surface under test with the optical axis by Z'_n and Z_n , respectively. The transverse ray aberrations are then given by

$$TA = (Z'_n - Z_n)Sc \quad (8.53)$$

where c is the inverse of the radius of curvature of the central zone and $S^2 = x^2 + y^2$. The wavefront aberration is obtained from

$$W = c^2 \int_{s=0}^s (Z'_n - Z_n)S dS. \quad (8.54)$$

This integral may be solved numerically by employing the trapezoidal rule (dividing the region of integration into M subintervals) to give

$$W_M = \frac{c^2}{2} \sum_{m=1}^{M-1} [(Z'_n - Z_n)_{m+1}S_{m+1} + (Z'_n - Z_n)_m S_m] \times (S_{m+1} - S_m). \quad (8.55)$$

This formula can then be used to calculate the wavefront aberration from the experimental values Z'_n, S_m, S_{m+1} , and from the theoretical value Z_n .

The procedure is as follows:

1. The theoretical value for the intersection of the normal with the optical axis is (see chap. 18)

$$Z_n = \frac{1}{c} - Kz \quad (8.56)$$

where $K = e^2$ (e is the eccentricity of the conic surface) and z is found by calculating Eq. (18.1) for the experimental values of S_m .

2. The terms $(Z'_n - Z_n)_m S_m$ can then be evaluated to obtain finally the wavefront aberration W_M over the optical surface under test.

8.4. PLATZECK–GAVIOLA TEST

The so-called caustic test, introduced by Platzeck and Gaviola (1939), is based on the fact that the center of curvature of any off-axis segment over a “perfect” parabolic mirror lies not on the optical axis but off-center on a curve termed the caustic, as was pointed out by Wadsworth (1902). To prove this point Platzeck and Gaviola performed a beautiful experiment, masking a parabolic mirror except in two symmetrical zones of elliptical shape. The mirror was then illuminated by a slit source and the images formed with this optical system were recorded for different positions of the illuminating slit.

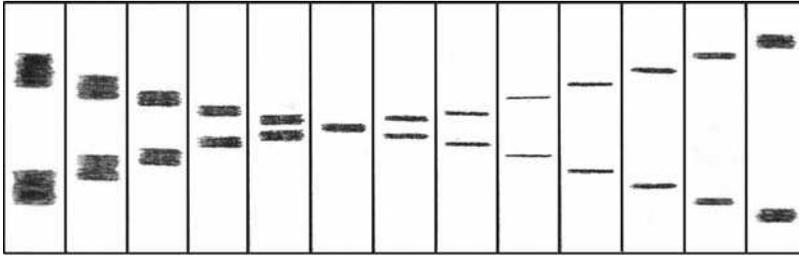


FIGURE 8.22. Sketch showing the appearance of the photographs taken by Platzeck–Gaviola using two slits over the mirror.

A series of photographs (showing the images of the illuminating slit) taken from different positions inside and outside of the focus of the mirror clearly indicated (see Fig. 8.22), which is a sketch of the photographs taken by Platzeck and Gaviola that when both beams cross the optical axis, the image is not sharp. But it shows the diffraction structure associated with the images obtained when the optical system has focus errors. Furthermore, in the same series of photographs it appears that behind the region where the two images cross the optical axis, the focal plane for these images is found.

Hence, Platzeck and Gaviola concluded that to obtain accurate measurements, it was necessary to use this focal plane with the images outside the optical axis as the modifying plane. Also, a thin opaque wire would be more suitable as the modifying screen since the wire would coincide with the image of the illuminating slit.

The method itself has proved to be useful and accurate. As Schroeder (1953) has pointed out, it has the following main advantages:

1. It is possible to observe one hole of the mask at a time.
2. Holes considerably wider than those used in testing zonal errors with the amplitude knife-edge test may be perforated in the mask.
3. The criteria for knowing the center of curvature of the portion of a mirror exposed by one hole in the mask, that is, the minimum of illumination that passes an opaque wire or strip are more reliable than the criteria for matching the irradiance distribution within the two holes, when using the knife edge on axis.

8.4.1. Geometrical Theory

In this section, we first show how to calculate the theoretical center of curvature for a section of any conic of revolution (Cornejo and Malacara, 1978) and then indicate the procedure employed in measuring the shape of the mirror from the experimental and theoretical data for the centers of curvature of the several sections over the optical surface under test.

We define a new set of coordinates (η, ξ) , also called caustic coordinates, such that the origin of the plane is at the center of curvature of the central zone,

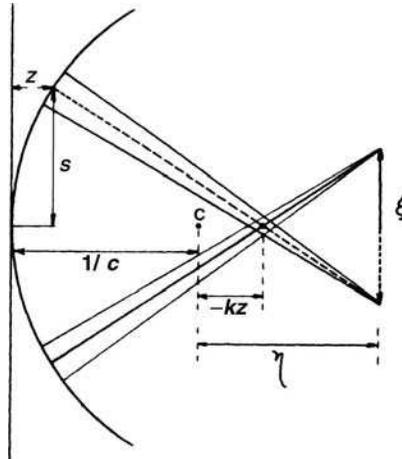


FIGURE 8.23. Geometry of the caustic.

as shown in Figure 8.23. The center of curvature of any zone will be referred to this origin.

The coordinates (η, ξ) for the center of curvature of any conic of revolution can be found from the following formula:

$$\frac{\xi}{2(\eta + Kz)} = \frac{S}{(1/c) - (K + 1)z} \quad (8.57)$$

for the triangles in dashed lines in Figure 8.23. From the same figure, the distance R between a point on the mirror with coordinates (z, S) and the center of curvature at $(1/c + \eta, \xi/2)$ can be obtained from

$$R^2 = \left(S + \frac{\xi}{2}\right)^2 + \left(\frac{1}{c} + \eta - z\right)^2 \quad (8.58)$$

Of course, R is the radius of curvature, which can also be found by using a well-known formula of calculus for the radius of curvature of any function S , that is,

$$R = \frac{[1 + (dz/dS)^2]^{3/2}}{d^2z/dS^2} \quad (8.59)$$

Now, the first derivative of z with respect to S is given by (see chap. 18)

$$\frac{dz}{dS} = cS[1 - (K + 1)c^2S^2]^{-1/2} \quad (8.60)$$

from which the second derivative can be calculated to give

$$\frac{d^2z}{dS^2} = c[1 - (K + 1)c^2S^2]^{-3/2} \quad (8.61)$$

If these two results are employed in Eq. (8.59), we have

$$R = \frac{1}{c}(1 - Kc^2S^2)^{3/2} \quad (8.62)$$

If, however, the value of $\xi/2$ in Eq. (8.57) is used in Eq. (8.58), we have

$$R^2 = \left(\frac{1}{c} + \eta - z\right)^2 \left\{ 1 + \frac{S^2}{[(1/c) - (K + 1)z]^2} \right\} \quad (8.63)$$

Now, from formulas (8.62) and (8.63) we can obtain the value of η , namely,

$$\eta = \frac{(1/c)(1 - Kc^2S^2)^{3/2}}{\left\{ 1 + \frac{S^2}{[(1/c) - (K + 1)z]^2} \right\}^{1/2}} + z - \frac{1}{c} \quad (8.64)$$

which, to eliminate the dependence on S , can be rewritten by using the expression for z as

$$\eta = -Kz\{3 + cz(K + 1) [cz(K + 1) - 3]\} \quad (8.65)$$

This value can be used in Eq. (8.57) to obtain the value of ξ , that is,

$$\xi = -2ScKz \left\{ \frac{2 + cz(K + 1)[cz(K + 1) - 3]}{1 - cz(K + 1)} \right\} \quad (8.66)$$

From Eq. (8.66), it is possible to calculate the theoretical position of the center of curvature $(\eta, \xi/2)$ of a particular zone of the mirror with coordinates (z, S) .

The usual procedure when employing the caustic method is as follows (Schroeder, 1953):

1. Measure c from the radius of curvature of the central zone of the mirror.
2. Cut a mask with an odd number of zones in order to have a zone at the center.
3. Determine the center of curvature of the central zone, with the shadow pattern over the mirror while the wire is displaced along the optical axis.
4. Calculate the theoretical value of (η, ξ) and the experimental values of (η', ξ') for the center of curvature of a particular zone, as the wire blocks out the light coming from each slot.

It is to be noted that only in this method each slot is, for a particular zone, tested separately. Consequently, the method may be particularly useful for testing errors that are not symmetrically distributed.

5. Calculate the theoretical minus the experimental values. If the difference is positive, the particular zone is overcorrected and vice versa.

The diffraction theory of this test has been considered by Simon (1971), Platzeck and Simon (1974), and Simon et al. (1979). An interesting variation of this method is considered by Teocaris and Gdoutos (1976).

8.5. PHASE MODULATION TESTS

Zernike (1934a) analyzed the Foucault test in terms of diffraction theory. At the same time he suggested an improved form of the knife-edge test, which he named “phase contrast” (Linfoot, 1946b). Essentially the same method was discovered quite independently by Lyot (1946). A further application of the phase modulation technique was later developed by Wolter (1956).

8.5.1. Zernike Test and its Relation to the Smartt Interferometer

The point diffraction interferometer has been previously described in Chapter 3. The aberrated wave to be measured is focused to produce an image, at a diffracting plate. This diffracting plate has a small central circular zone of radius a , with a smaller amplitude transmission than the surrounding annular zone as illustrated in Figure 8.24.

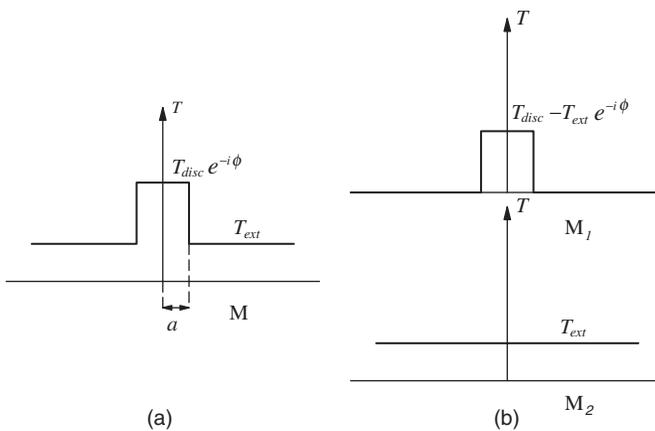


FIGURE 8.24. Transmission in diffracting plate of a Smartt Interferometer. (a) Total transmission and (b) transmission separated into two components.

The annular region is coated with an amplitude transmission T_{ext} , typically smaller than one, while the central disc is made with an amplitude transmission T_{disc} , typically equal to one. To produce the desired 90° phase difference between the two interfering wavefronts, the central zone is coated with a film with optical thickness ϕ , so that the phase optical path through the center is greater than the phase optical path through the annular region. It has been shown in Section 3.8 that the phase difference γ between the wavefront under test and the reference wavefront generated by diffraction on the central zone is given by

$$\tan \gamma = \frac{T_{disc} \sin(\phi)}{T_{disc} \cos \phi - T_{ext}} \quad (8.67)$$

but since we need $\gamma = \pi/2$ we require that

$$\cos \phi = \frac{T_{ext}}{T_{disc}} \quad (8.68)$$

To produce an interferogram with good contrast, the two interfering wavefronts should have the same amplitudes at the observing plane. This is controlled by means of the transmittances T_a and T_b as well as by the central disk diameter as described by Golden (1977a and 1977b).

The phase contrast method has the following advantages for testing an optical surface with aberrations smaller than the wavelength of the illuminating radiation:

1. The aberrations appear as a linear superposition term over a uniform background intensity.
2. The contrast of the aberrations, over this uniform background intensity can be increased by decreasing the intensity in the background illumination, that is, by decreasing the optical transmittance over the $\lambda/4$ region. The mathematical formulation associated with this method is given in the following discussion.

The modifying function employed in the phase contrast test may be described as a disk with a complex amplitude transmittance equal to $TA_{disc} \exp(i\phi)$ and a radius a nearly equal to the radius of the Airy disk. Outside this disk the complex amplitude transmittance is equal to T_{ext} . Thus, the filtering plate has an amplitude transmission

$$\begin{aligned} M(\sigma) &= T_{disc} \exp(i\phi) & \text{if } \sigma < \frac{a}{\lambda r} \\ M(\sigma) &= T_{ext} & \text{if } \sigma > \frac{a}{\lambda r} \end{aligned} \quad (8.69)$$

with

$$V > \frac{a}{\lambda r} \quad (8.70)$$

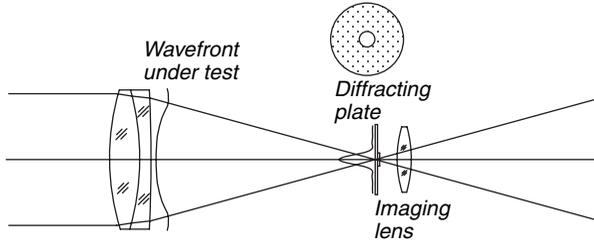


FIGURE 8.25. Testing a lens with a Smartt interferometer.

Equivalently, as illustrated in Figure 8.25, we can write $M(\sigma) = M_1(\sigma) + M_2(\sigma)$, where

$$\begin{aligned} M_1(\sigma) &= T_{disc} \exp(i\phi) - T_{ext} & \text{for } 0 < \sigma < \frac{a}{\lambda r} \\ M_2(\sigma) &= T_{ext} & \text{for } 0 < \sigma < \infty \end{aligned} \quad (8.71)$$

The observed complex amplitude at the image of the observed pupil of the system under test, after being filtered by the diffraction plate is given by

$$A'(x, y) = \mathfrak{S}^{-1}(M_1(f_x, f_y) \mathfrak{S}(A(x, y))) + \mathfrak{S}^{-1}(M_2(f_x, f_y) \mathfrak{S}(A(x, y))) \quad (8.72)$$

which, after using the value of the filter functions in Eq. (8.55)

$$A'(x, y) = T_{ext} A(x, y) + \mathfrak{S}^{-1}(M_2(f_x, f_y) \mathfrak{S}(A(x, y))) \quad (8.73)$$

The phase contrast method is frequently presented as the result of two-beam or multiple-beam interference (Françon, 1952; Françon and Nomarski, 1952; Golden, 1977a; Smartt, 1975) The first term represents the wavefront under test with an amplitude T_{ext} and the second term represents the reference wavefront. The complex amplitude B_0 at the center of the image (optical axis) given by the central ordinate theorem is

$$B_0 = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} A(x, y) dx dy \quad (8.74)$$

If we assume that the central disk is small enough so that the amplitude and phase of the image over this disk is a constant equal to B_0 , the second term in Eq. (8.56) can be written as

$$2\pi a^2 (T_{disc} \exp(i\phi) - T_{ext}) B_0 \frac{J_1(2\pi\rho a)}{(2\pi\rho a)} \quad (8.75)$$

so that the amplitude of the reference wavefront at the optical axis is equal to

$$2\pi a^2(T_{disc} \exp(i\phi) - T_{ext})B_0 \quad (8.76)$$

The condition for maximum visibility of the fringes is that the wavefront under test and the reference wavefront should have the same intensity. Hence,

$$2\pi a^2(T_{disc} \exp(i\phi) - T_{ext})B_0 = T_{ext} \quad (8.77)$$

If the wavefront aberrations are small compared with the wavelength and the phase difference between the two interfering wavefronts is $\pi/2$, the image intensity variations are proportional to the aberration function. Of course, if the aberration is a smoothly varying function fairly distributed over the object, its diffraction pattern may be passing by the central disk. Therefore, it will remain invisible in the image intensity distribution. The contrast can be increased by decreasing the background of the image intensity, that is, by increasing the absorption of the central disk, as it was pointed out before.

To make more clear how the two interfering waves interact in this interferometer, let us represent the complex amplitude leaving the surface under test by

$$A(x, y) = A_0(x, y) \exp\left(i \frac{2\pi}{\lambda} W(x, y)\right) \quad (8.78)$$

where again $W(x, y)$ is the wavefront deformation. The complex amplitude after passing the diffracting plate, as described in Figure 3.23 in chapter 3, may conveniently be written as

$$\begin{aligned} A'(x, y) &= T_{ext}A_0(x, y) \exp\left(i \frac{2\pi}{\lambda} W(x, y)\right) \\ &+ K(T_{disc} \exp(i\phi) - T_{ext})A_0(x, y) \end{aligned} \quad (8.79)$$

where the first term is the wavefront under test and the second term is the “ideal” spherical reference wavefront. From Eq. (8.75), the factor K is

$$K = 2\pi a^2 B_0 \frac{J_1(2\pi\rho a)}{(2\pi\rho a)} \quad (8.80)$$

8.5.2. Lyot Test

Finally, as Texereau (1957) pointed out, although the phase contrast technique has proved to be extremely useful in microscopy, its value in optical testing seems to be limited. The problem is that the test is photometric, that is, it is based on irradiance variations, and the interpretation and application of the results obtained are not

straightforward tasks. This is especially true when a clear understanding, in terms of physical optics, of the properties of the test is lacking. The most common application in optical testing of the phase contrast method involves the detection of surface defects that are small in area and amplitude.

As mentioned before, the Zernike test usually is performed using a point source (white or monochromatic) and a phase plate with a small circular disk, smaller than or equal to the Airy disk. Another alternative, used with modern phase contrast microscopes, is to employ a circular, or annular, extended light source and a phase plate whose retarding zone exactly coincides with the geometrical image of the source. The advantages of this procedure are that more light is used and that it is easier to manufacture a larger phase plate. Using this principle, Lyot (1946), independently of Zernike, developed a technique in which he used a relatively wide slit (100–200 μm) as an illuminating light source. This slit is imaged by the optical surface under test upon a phase-retarding slit with an optical thickness of $\lambda/4$. To increase the contrast, the transmittance of the retarding slit is reduced to about $T = 10^{-D}$, with D being the optical density, which ranges from 1.5 to 3.0. See for example Françon and Nomarski (1952).

Since the light source is a slit and not circular, the light diffracted in the direction of the slit will not alter its phase relative to the undiffracted light and thus, the corresponding errors will not be detected. This means that only errors with variations in the direction perpendicular to the slit will be detected. This preferential orientation of the errors is shown in Figure 8.26, taken with a vertical slit. This detected pseudo-orientation of the error can be ignored, assuming that the real errors are randomly oriented.

The Lyot version of the phase contrast method is so sensitive that it is possible to see defects that deviate from the mean surface by 1 \AA with a contrast equal to 15%.

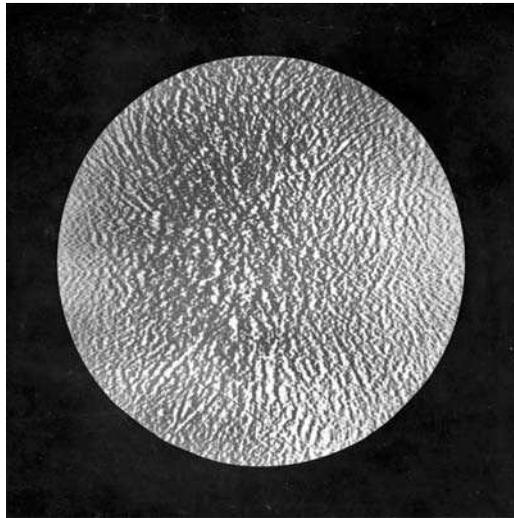


FIGURE 8.26. Lyotgraph taken with a vertical phase slit. (From Françon, 1963.)

When a polychromatic light source is employed, the phase-modulating screen can have different optical thicknesses for the different wavelengths present, and the image irradiance variations can appear in color.

Information about the experimental procedures for making phase disks or phase slits can be found in the articles of Burch (1934), Zernike (1942), Françon and Nomarski (1952), and Golden (1977b). This test can also be implemented in polarized light (Kastler and Momtarnal, 1948).

8.5.3. Wolter Test

The $\lambda/2$ phase-edge or phase-knife test, proposed by Wolter (1949), and independently by Kastler (1950), can be regarded as an improvement of the wire test or the zonal Foucault method. The idea is to employ usefully the diffracted light at one side, instead of blocking it out as when the knife edge is used.

The method is based on reversing the sign of the complex amplitude across a half-plane, while the other half remains unaltered. This is achieved by introducing a transparent half-plane with an optical thickness equal to $\lambda/2$. This method is also known as the Hilbert transform method (Lowenthal and Belvaux, 1967; Belvaux and Vareille, 1970, 1971) because mathematically the image is obtained as the Hilbert transform of the object.

Since in this test more diffracted light is collected by the eye of the observer, this method would be expected to produce a shadow pattern of higher contrast. Therefore, the $\lambda/2$ phase-edge test is expected to be more sensitive and accurate, in locating on the aspherical surface under test the projected shadow of the phase edge, than the normal knife-edge test. Here, again, as in the phase contrast method, it is not possible to explain in terms of geometrical optics the irradiance variations over the image plane when a $\lambda/2$ phase edge is introduced near the paraxial focal plane. However, Landgrave (1974) made the additional assumption that only the rays passing through the border of the $\lambda/2$ phase edge interfere destructively.

Although the later approach may appear primitive, it is in accordance with the fact that the Fresnel diffraction images of a $\lambda/2$ phase edge have zero irradiance values of the image at the border (Kastler, 1950; Hopkins, 1952; Ojeda-Castañeda, 1976), as shown in Figure 8.27. Indeed, the Wolter test proves to be extremely useful since the $\lambda/2$ phase edge is considered to be equivalent to the wire test. In other words, the

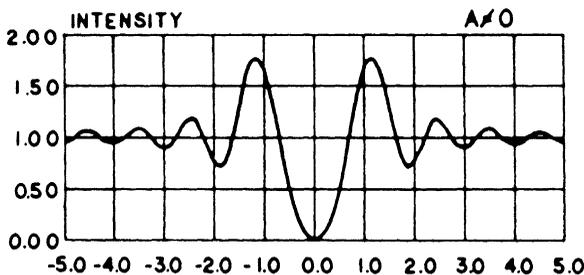


FIGURE 8.27. Fresnel diffraction image of a $\lambda/2$ phase edge observed over the surface under test.



FIGURE 8.28. Shadow patterns obtained when using a $\lambda/2$ phase edge for testing a lens with comatic aberrations.

border of the $\lambda/2$ phase edge may be thought of as the border of the opaque knife edge, and the “reversal” side of the $\lambda/2$ phase edge may be considered to have a transmittance equal to unity, as on its “clear” side. Therefore, the shadow patterns associated with the $\lambda/2$ phase edge are, from the point of view of geometrical optics, considered to be identical with the shadow patterns associated with the wire test. Upon analysis the boundaries of the shadow patterns are, then, identical to those presented in Section 8.2.2. Notice the similarity between the Foucault test drawings in Figure 8.9 and the phase-edge photographs of a comatic lens (kindly provided by Landgrave, 1974) in Figure 8.28.

It has been reported (Vasil’ev, 1971; Landgrave, 1974) that the dark lines of the shadow pattern associated with the $\lambda/2$ phase edge are better contrasted against the background irradiance intensity than the dark lines associated with both the knife-edge test and the wire test. Thus, the sensitivity of the $\lambda/2$ phase-edge test is greater than the sensitivity obtained when using an opaque knife edge or the wire test. The wire test photographs in Figures 8.29(a,b) should be compared with the phase-edge photographs, Figures 8.31(c,d), respectively, taken with the same aspherical mirror.

Obviously, a formal analysis of the sensitivity is, indeed, more complicated, and it is not considered here. However, it is possible to show (Ojeda-Castañeda, 1976, 1979) that for aberrations smaller than the wavelength of the incident radiation, the $\lambda/2$ phase edge will produce images with higher contrast than the images produced with the opaque knife edge. The treatment is repeated here.

The modulating function for the phase-edge method is

$$M(x_1, y_1) = \begin{cases} -1, & x_1 < -r_1 \\ 1, & x_1 > -r_1 \end{cases} \quad (8.81)$$

The complex amplitude leaving the plane of modulation and traveling toward the image plane, after being modulated by function (8.81), is again obtained by $M(x_1, y_1)a(x_1, y_1)$. The complex amplitude distribution over the image plane can be calculated by taking the inverse Fourier transform, giving

$$A'(x_2, y_2) = 1 + i \frac{2\pi}{\lambda} W'(x_2, y_2) \quad (8.82)$$

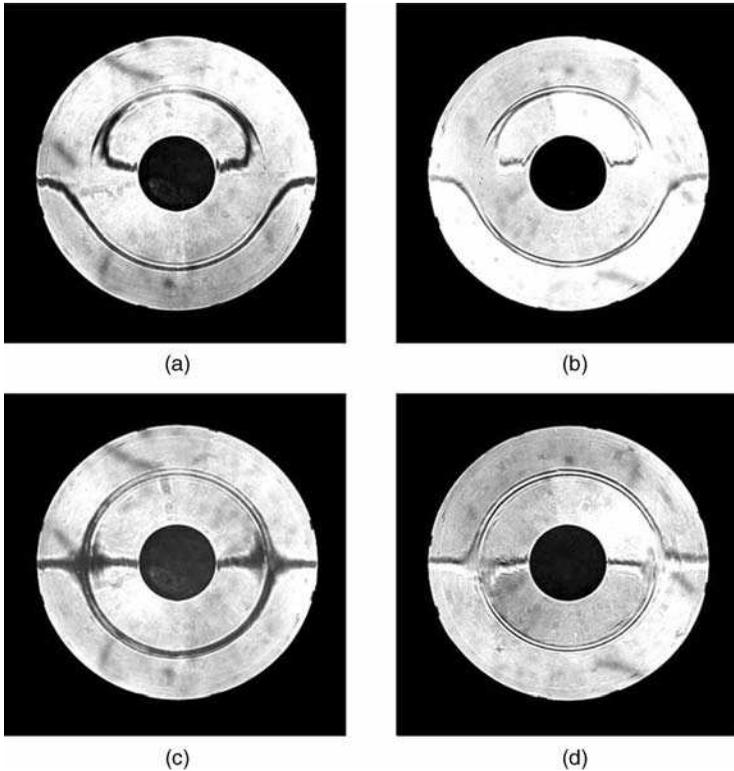


FIGURE 8.29. Shadow patterns of a hyperbolic mirror. (a) Off-axis wire. (b) Off-axis phase edge. (c) On-axis wire. (d) On-axis phase edge.

where again $W'(x_2, y_2)$ denotes the inverse Fourier transform of $M(x_1, y_1)w_1(x_1, y_1)$. When calculating the image intensity, the same approximation, neglecting second-order terms in λ^{-1} , may be made. Under this approximation, as was pointed out in the sections dealing with the diffraction theory of the knife-edge and wire tests, only the real-odd component of the modulating function gives rise to the visible term in the image irradiance variations. Since in this case the function is

$$g(x_1, y_1) = \begin{cases} -1, & x_1 < -r_1 \\ 0, & |x_1| < r_1 \\ 1, & x_1 > r_1 \end{cases} \quad (8.83)$$

we do not have to calculate the image irradiance, for we know that the result will be the same as that obtained when the image irradiance is calculated using the knife edge. The only exception to this statement that should be noted is that function (8.83) has double amplitude transmittance. Therefore, the contrast of the image irradiance

variations when the phase edge is used is twice that obtained with the knife edge, that is,

$$\gamma_{pe} = 2\gamma_{ke} = \frac{8\pi}{\lambda} W_i(x_2, y_2) \quad (8.84)$$

The other characteristics associated with the Wolter test are the same as those listed for the knife-edge and the wire tests.

8.6. RITCHEY-COMMON TEST

According to Ritchey (1904), this test was suggested by a Mr. Common as a way of testing large optical flats (Shu, 1983). For this purpose a good spherical mirror is used, with the flat acting also as a mirror, to obtain an apparatus similar to the one used for the knife-edge test. See Figure 8.30.

Any small spherical concavity or convexity of the surface under test appears as an astigmatic aberration in the image of the illuminating point source. The two focal planes associated with the sagittal and tangential foci can be accurately found (by employing the knife-edge test as indicated in Section 8.1), and with this information the radius of curvature of the surface under test can be calculated. Of course, any localized deformation or deviation from flatness can also be observed.

The advantage of the Ritchey-Common test over other methods is that a reference flat is not needed as in the Newton or Fizeau interferometer. The autocollimation method requires a paraboloid as large as the largest dimension of the flat under test. On the contrary, the auxiliary mirror used in the Ritchey-Common test has a spherical shape, which is the easiest to make and test. An additional advantage is that the spherical mirror needs to be only as large as the minor diameter, if it is elliptical, like many diagonal mirrors used in telescopes or other instruments.

If we measure the astigmatism, we can obtain the magnitude of the concavity or convexity of the "flat" mirror, but it is assumed that this mirror is spherical, not

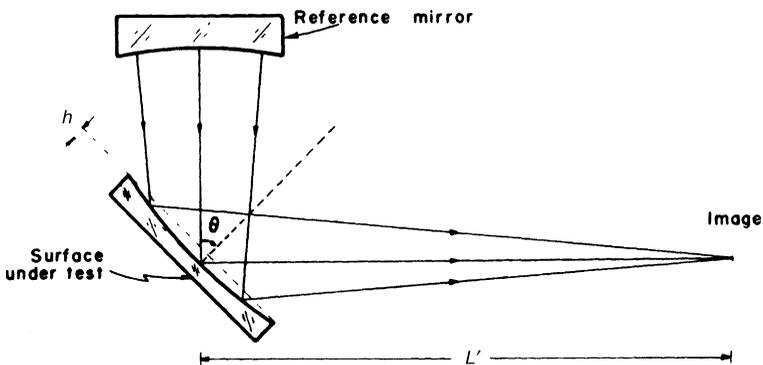


FIGURE 8.30. Geometry of the Ritchey-Common test.

toroidal, in shape. If the mirror is not circular but elliptical, the possibility that it will develop a toroidal shape while being polished is very high. This toroidal shape introduces an additional astigmatism that complicates the analysis. Interested readers can consult the work of Tatian (1967), Silvernail (1973), and Shu (1983).

The mathematical description given here assumes that the mirror under test is not toroidal and follows the approach used by Couder (1932) and Strong (1938), that is, finding the curvature from the Coddington equations for astigmatism. In what follows, it is assumed that the pinhole size source and the wire travel together in the same plane. Hence, when evaluating the transversal ray aberrations, it is necessary to multiply the value obtained in this way by a factor of 0.5 to get the real value of the transverse ray aberrations. However, since the aberrations of the wavefront are duplicated in the Ritchey-Common test because of the double reflection on the optical flat under test, we can conveniently omit the double reflection factor of 2 and factor 0.5 when calculating the value of the transversal ray aberrations.

Applying the well-known Coddington equations for astigmatism to a reflecting surface, we find that the sagittal focus of a converging beam incident on an optical reflecting surface is given (see Figure 8.31) by

$$\frac{1}{s'} = \frac{1}{s} + \frac{2 \cos \theta}{r} \tag{8.85}$$

and the tangential focus by

$$\frac{\cos^2 \theta}{t'} = \frac{\cos^2 \theta}{t} + \frac{2 \cos \theta}{r} \tag{8.86}$$

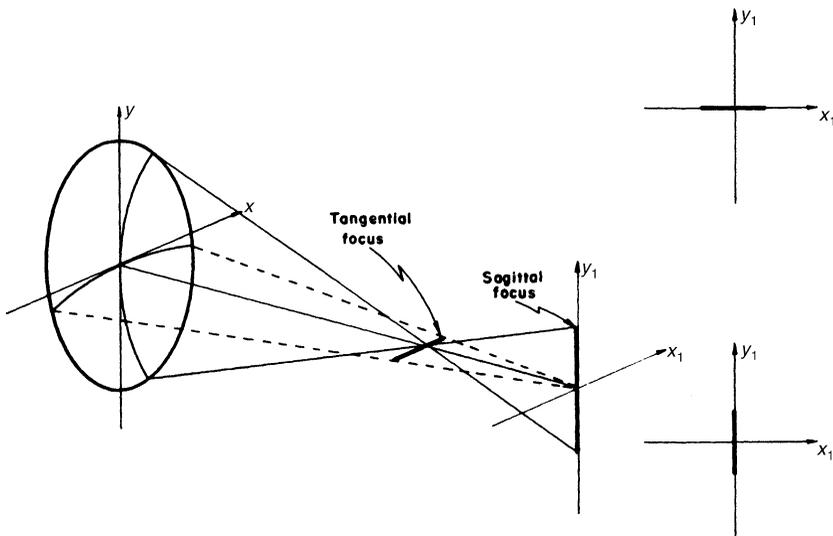


FIGURE 8.31. An astigmatic wavefront and its associated focal images. It is to be noted that the orientations of the images are changed when the setup in Figure 8.30 is used. That is, the sagittal image is a horizontal line and the tangential image is a vertical one.

where θ is the angle of incidence at the surface under test and r is the radius of curvature of the optical surface under test.

If we assume that the incident beam is free from astigmatism, that is, $s = t$, Eqs. (8.85) and (8.86) can be used to obtain

$$\frac{s' - t'}{s't'} = \frac{2 \sin^2 \theta}{r \cos \theta} \quad (8.87)$$

Now, by denoting $s' = L' + \delta L'$ and $t' = L'$, and substituting these values in Eq. (8.87), we have

$$r = 2L' \left(1 + \frac{L'}{\delta L'} \right) \frac{\sin^2 \theta}{\cos \theta}. \quad (8.88)$$

Since $|L'/\delta L'| > 1$, it follows from Eq. (8.88) that, when the sagittal focus is longer than the tangential focus ($\delta L' > 0$), the surface under test has a positive radius of curvature, that is, the surface is convex. On the contrary, when the sagittal focus is shorter than the tangential focus ($\delta L' < 0$), the surface has a negative radius of curvature and is concave. The sagitta h of a surface with a diameter D and a radius of curvature r can be approximated by

$$h = \frac{D^2}{8r} \quad (8.89)$$

Using Eq. (8.88) in (8.81) we have

$$h = \frac{D^2 \cos \theta}{16L'(1 + L'/\delta L') \sin^2 \theta} \quad (8.90)$$

For $L' \gg \delta L'$ ($|L'/\delta L'| \gg 1$), Eq. (8.90) reduces to

$$h = \frac{D^2 \delta L' \cos \theta}{16L'^2 \sin^2 \theta} \quad (8.91)$$

which gives the value for the sagitta of the optical surface under test for any angle of incidence θ . For the particular case of $\theta = 45^\circ$, the sagitta is given by

$$h = \frac{D^2 \delta L'}{11.3L'^2} \quad (8.92)$$

Couder (1932 in Texereau 1957) has indicated that an optimum angle for testing optical flats is $\theta = 54^\circ 45'$. Of course, in practice it is more convenient to set $\theta = 60^\circ$. It is also common practice to use any eyepiece, instead of the knife edge, to determine visually the positions of the sagittal and tangential foci, by locating the positions of the patterns shown to the right of Figure 8.31.

8.7. CONCLUSIONS

As mentioned in the introduction of this chapter, the examples examined demonstrate the versatility and usefulness of the Schlieren techniques and the related phase modulation tests. Of course, the choice of any particular test depends on the circumstances.

REFERENCES

- Belvaux Y. and J. C. Vareille, "Contrôle de L'état de Surface ou d'Homogénéité de Matériaux Optiques par Contraste de Phase à Déphasage Quelconque," *Optics Comm.*, **2**, 101 (1970).
- Belvaux Y. and J. C. Vareille, "Visualisation d'Objets de Phase par Transformation de Hilbert," *Nouv. Rev. Opt.*, **2**, 149 (1971).
- Banerji S., "On Some Phenomena Observed in the Foucault Test," *Astrophys. J.*, **48**, 50 (1918).
- Barakat R., "General Diffraction Theory of Optical Aberration Tests, from the Point of View of Spatial Filtering," *J. Opt. Soc. Am.*, **59**, 1432 (1969).
- Bescós J. and L. R. Berriel-Valdos, "Foucault Test in White Light Imagery," *Opt. Commun.*, **57**, 10 (1986).
- Buchroader R. A., L. H. Elmore, R. V. Shack, and P. N. Slater, "The Design Construction and Testing of the Optics for a 147 cm Aperture Telescope," Technical Report No. 79, University of Arizona, Tucson, AZ, 1972a.
- Buchroader R. A., L. H. Elmore, R. V. Shack, and P. N. Slater, Optical Sciences Center Technical Report No. 79, 1972b.
- Burch C. R., "On the Phase-Contrast Test of F. Zernike," *Mon. Not. R. Astron. Soc.*, **94**, 384 (1934).
- Conrady H. G., "Study of the Significance of the Foucault Knife-edge Test When Applied to Refracting Systems," *Trans. Opt. Soc.*, **25**, 219 (1924).
- Cornejo-Rodriguez A. and D. Malacara, "Caustic Coordinates in Platzeck-Gaviola Test of Conic Mirrors," *Appl. Opt.*, **17**, 18 (1978).
- Couder A., *Recherches sur les Déformations des Grands Miroirs Employés aux Observations Astronomiques*, Thesis Faculté des Sciences Paris, *Bull. Astron.*, 7 (1932).
- Dakin R. K., "An Improved Foucault Testing Device," *Sky Telesc.*, **33**, 45 (1967).
- Fell B., "Optical Surfaces: Computer Program Facilitates Work in Optical Shop," *Opt. Sci. Newslett.*, **2**, 5 (1968).
- Foucault L. M., "Description des Procédés Employés pour Reconnaître la Configuration des Surfaces Optiques," *C. R. Acad. Sci. (Paris)*, **47**, 958 (1858).
- Foucault L. M., "Mémoire sur la Construction des Télescopes en Verre Argente," *Ann. Obs. Imp. Paris.*, **5**, 197 (1859).
- Françon M., *Modern Applications of Physical Optics*, Interscience, New York, 1963.
- Françon M. Ed., *Étude Théorique Comparative du Contraste de Phase et de Méthodes Interférentielles, Contraste de Phase et Contraste par Interférométrie*, Edition de la Revue d'Optique Théorique et Instrumentale, Paris, 1952, p. 48.

- Françon M. and G. Nomarski, Eds., *Les l'Ames de Phase par Reflexion Contraste de Phase et Contraste par Interferometrie*, Edition de la Revue d'Optique Theorique et Instrumentale, Paris, 1952, p. 136.
- Gascoine S. C. B., "The Theory of the Foucault," *Test. Mon. Not. R. Astron. Soc.*, **104**, 326 (1944).
- Gaviola E., "On the Quantitative Use of the Foucault Knife-Edge Test," *J. Opt. Soc. Am.*, **26**, 163 (1936).
- Golden L. J., "Zernike Test. 1: Analytical Aspects," *Appl. Opt.*, **16**, 205 (1977a).
- Golden L. J., "Zernike Test. 2: Experimental Aspects," *Appl. Opt.*, **16**, 214 (1977b).
- Hartmann J., "An Improvement of the Foucault Knife-Edge Test in the Investigation of Telescope Objectives," *Astrophys. J.*, **27**, 254 (1908).
- Hopkins H. H., "Phase Structures Seen in the Ordinary Microscope," in *Contraste de Phase et Contraste par Interferometrie*, M. Françon, Ed., Rev. d'Opti. Theor. et Inst, Paris, 1952, p. 142.
- Kastler A. "Un Systeme de Franges de Diffraction a Grand Contraste," *Rev. Opt.*, **29**, 307 (1950).
- Kastler A. and R. Momtarnal, "Phase Contrast in Polarized Light," *Nature*, **161**, 357 (1948).
- Kingslake R., "The Knife-Edge Test for Spherical Aberration," *Proc. Phys. Soc.*, **49**, 376 (1937).
- Landgrave J. E. A., *Phase Knife Edge Testing*, M.Sc. Report Imp. Coll. Sc. and Tech., London, 1974.
- Linfoot E. H., "Astigmatism Under the Foucault Test," *Mon. Not. R. Astron. Soc.*, **105**, 193 (1945a).
- Linfoot E. H., "A Contribution to the Theory of the Foucault Test," *Mon. Not. R. Astron. Soc.*, **105**, 193 (1945b).
- Linfoot E. H., "A Contribution to the Theory of the Foucault Test," *Proc. Roy. Soc. Ser. A*, **186**, 72 (1946a).
- Linfoot E. H., "On the Zernike Phase-Contrast Test," *R. Astron. Soc.*, **58**, 759 (1946b).
- Linfoot E. H., "On the Interpretation of the Foucault Test," *Proc. Roy. Soc. Ser. A*, **193**, 248 (1946c).
- Linfoot E. H., "The Foucault Test," in *Recent Advances in Optics*, Oxford University Press, Dover, 1955, Chap. II.
- Loomis D. A., "Testing Techniques Arouse Interest," *Opt. Sci. Newslett.*, **2**(3), 73 (1968).
- Lowenthal S. and Y. Belvaux, "Observation of Phase Objects by Optically Processed Hilbert Transform," *Appl. Phys. Lett.*, **2**, 49 (1967).
- Lyot B., "Procedes Perme Hand d'Etudier les Irregularities d'une Surface Optique Bien Polie," *C. R. Acad. Sci. (Paris)*, **222**, 765 (1946).
- Meinel A. B., "Quantitative Reduction of a Wire Test (Azimuthal is Explained)," *Opt. Sci. Newlett.*, **2**(5), 134 (1968).
- North R. J., "A Colour Schlieren System Using Multicolor Filters of Simple Construction," National Physics Laboratory, England, Aero, 1954, p. 26.
- O'Hare J. E. and J. D. Trolinger, "Holographic Colour Schlieren," *Appl. Opt.*, **8**, 2047 (1969).
- Ojeda-Castañeda J., *Images of Certain Type of Phase Objects*, Ph.D. Thesis, University of Reading, England, 1976.

- Ojeda-Castañeda J., "Nota Sobre el Microscopio de Contraste de Modulación," *Bol. Inst. Tonantzintla*, **2**, 293 (1978).
- Ojeda-Castañeda J., "Necessary and Sufficient Conditions for Thin Phase Imagery," *Opt. Acta*, **27**, 905 (1980a).
- Ojeda-Castañeda J., "A Proposal to Classify Methods Employed to Detect Thin Phase Structures under Coherent Illumination," *Opt. Acta*, **27**, 917 (1980b).
- Ojeda-Castañeda J. and L. R. Barriol-Valdos, "Classification Scheme and Properties of Schlieren Techniques," *Appl. Opt.*, **18**, 3338 (1979).
- Platzcek R. P. and E. Gaviola, "On the Errors of Testing a New Method for Surveying Optical Surfaces and Systems," *J. Opt. Soc. Am.*, **29**, 484 (1939).
- Platzcek R. P. and J. M. Simon, "The Method of the Caustic for Measuring Optical Surfaces," *Opt. Acta*, **21**, 267 (1974).
- Porter R. W., "Mirror Making for Reflecting Telescopes," in: *Amateur Telescope Making*, vol. 1, A. G. Ingalls, Ed., Scientific American, New York, 1953, p. 1.
- Rayces J. L., "Exact Relation Between Wave Aberration and Ray Aberration," *Opt. Acta*, **11**, 85 (1964).
- Rayleigh Lord, "On Methods for Detecting Small Optical Retardations, and on the Theory of Foucault's Test," *Philos. Mag.*, **33**, 161 (1917).
- Ritchey G. W., "On the Modern Reflecting Telescope and the Making and Testing of Optical Mirrors," *Smithson. Contrib. Knowl.*, **34**, 3 (1904).
- Schardin H., "Die Schlierenverfahren und ihre Anwendungen," *Ergebn Exakt Naturw.*, **20**, 303 (1942a).
- Schardin H. and G. Stamm, "Prüfung von Flachglass mit Hilfe Eines Farbigen Schlierenverfahrens," *Glastechn. Ber.*, **20**, 249 (1942b).
- Schroader I. H., "The Caustic Test," in: *Amateur Telescope Making*, vol. 3, A. G. Ingalls, Ed., Scientific American, New York, 1953, p. 429.
- Shu K. L., "Ray-Trace Analysis and Data Reduction Methods for the Ritchey-Common Test," *Appl. Opt.*, **22**, 1879 (1983).
- Silvernail C. J., "Extension of the Ritchey Test," *Appl. Opt.*, **12**, 445 (1973).
- Simon J. M., "Diffraction Theory of the Method of the Caustic for Measurement of Optical Surfaces," *Opt. Acta*, **18**, 369 (1971).
- Simon J. M., "Difracción en el Método de Foucault por Zonas," Physics 01/76, University of Buenos Aires, Argentina, 1976.
- Simon M. C. and J. M. Simon, "Testing of Flat Optical Surfaces by the Quantitative Foucault Method," *Appl. Opt.*, **17**, 132 (1978).
- Simon M. C., J. M. Simon, and E. L. Zenobi, "Testing Optical Surfaces by the Method of the Caustic Using a Phase-Stripe as Spatial Filter," *Appl. Opt.*, **18**, 1463 (1979).
- Smartt R. N. and W. H. Steel, "Theory and Application of Point Diffraction Interferometers," Proceedings of the ICO Conference on Optical Methods in Scientific and Industrial Measurements, Tokyo, 1974; *Jap. J. Appl. Phys.*, **14**, Suppl. I, 351 (1975).
- Smith W. J., *Modern Optical Engineering*, Mc-Graw Hill, New York, 1966, p. 439.
- Stolzenburg W. A., "The Double Knife Edge Technique for Improved Schlieren Sensitivity in Low Density Hypersonic Aerodynamic Testing," *J. SMPTE*, **74**, 654 (1965).
- Strong J., *Procedures in Experimental Physics*, Prentice-Hall, New York, 1938.

- Tatian B., An Analysis of the Ritchey–Common Test for Large Plane Mirrors, Internal Report, ITEK Corporation, Massachusetts, 1967.
- Texereau J., *How to Make a Telescope* (Translation and Adaptation from *La Construction du Telescope d'Amateur*), American Museum of Natural History, Anchor Books Doubleday, New York, 1957.
- Teocaris P. S. and E. E. Gdoutos, “Surface Topography by Caustics,” *Appl. Opt.*, **15**, 1629 (1976).
- Töpler A., *Beobachtungen Nach Einer Neuen Optischen Methode*, Max. Cohen. und Sohn, Bonn, 1864.
- Töpler A., “Beobachtungen Nach Einer Neuen Optischen Methode,” *Pogg. Ann.*, **127**, 556 (1866).
- Töpler A., “Beobachtungen Singender Flammenmit clem Schlierenapparat,” *Pogg. Ann.*, **128**, 126 (1866); also in 131, 33 and 180 (1867) and 134, 194 (1868).
- Vasil’ev L. A., *Schlieren Methods*, Israel Prog. for Scient. Transl., New York, 1971.
- Wadsworth F. L. O., “Some Notes on the Correction and Testing of Parabolic Mirrors,” *Pop. Astron.*, **10**, 337 (1902).
- Weinberg F. J., *Optics of Flames*, Butterworth, London, 1961.
- Welford W. T., “A Note on the Theory of the Foucault Knife-Edge Test,” *Opt. Commun.*, **1**, 443 (1970).
- Wolter H., Minimum Strahl Kenn Zeichnung, German Patent No. 819925, assigned to E. Leitz, Wetzlar, 1949.
- Wolter H., “Schlieren-Phase Kontrast und Lichtschnittverfahren,” in: *Handbuch der Physik*, vol. 24, Springer-Verlag, Berlin, 1956, p. 582.
- Zernike F., “Diffraction Theory of Knife-Edge Test and Its Improved Form, the Phase Contrast,” *Mon. Not. R. Astron. Soc.*, **94**, 371 (1934a).
- Zernike F., “Begungstheorie des Schneidenver-Fahrans und Seiner Verbesserten Form, der Phasenkontrastmethode,” *Physica*, **1**, 44 (1934b).
- Zernike F., “Phase Contrast, a New Method for the Microscopic Observation of Transparent Objects,” *Physica*, **9**, 686 (1942).

9

Ronchi Test

A. Cornejo-Rodriguez

9.1. INTRODUCTION

Since its discovery and application to the testing of optical surfaces, the Ronchi test has been used widely either in a qualitative or in a quantitative way. Also, because it is simple to accomplish and easy to interpret the experimental observations, the Ronchi test has almost always been conceptually interpreted from the point of view of geometrical optics. Sections 9.2 and 9.5 reflect this point of view. However, the author believed that a more comprehensive presentation of the two aspects of the Ronchi test would provide a fuller view of the subject. These two aspects are (a) the need for a quantitative analysis of the data and (b) the development of the test from the point of view of the principles of physical optics and a comparison to the geometrical approach (Toraldò di Francia, 1941a). Sections 9.3 and 9.4 are devoted to these areas of study, respectively, and the need for mathematical treatment in these two sections was unavoidable. Perhaps some readers would like to skip these two sections (as can be done without losing continuity); others seeking a more comprehensive treatment of the test will find this material interesting. Section 9.6 is a brief review of some tests that can be related to the classical Ronchi test in various ways.

With respect to the previous editions of this book, in the present one, the author has added some corrections to the original equations, figures, and additional applications and new references have been included.

9.1.1. Historical Introduction

The Ronchi test is one of the simplest and most powerful methods to evaluate and measure the aberrations of an optical system. The Italian physicist Ronchi (1923a) discovered that when a ruling was placed near the center of curvature of a mirror, the image of the grating was superimposed on the grating itself, producing a kind of Moiré pattern that he called combination fringes. Since the shape of these combination fringes depended on the aberrations of the mirror, he immediately thought of

applying the phenomenon to the quality testing of mirrors. However, the combination fringes proved extremely difficult to interpret. The Ronchi test in the form that we now know appeared when Ronchi (1923b) published his thesis in order to obtain his final diploma at the Scuola Normale Superiore di Pisa.

One of the first applications of the method was to measure the aberrations of the telescope made by Galileo (Ronchi, 1923c) and a lens made by Toricelli (Ronchi, 1923d). The first serious application of this test to astronomical telescopes was made by Anderson and Porter (1929), and since then it has been very widely used by professional and amateur astronomers (King, 1934; Strong, 1938; Kirkham, 1953; Porter, 1953; Phillips, 1963).

A few months after Ronchi's invention, Lenouvel (1924a, 1924b, 1925a, 1925b) published an extensive study of this test in France, along the same lines as those followed by Ronchi.

General descriptions of the Ronchi method can be found in many review articles (Ronchi, 1925; Morais, 1958; Adachi, 1962a; Wehn, 1962; Murty, 1967; Briers, 1972; Briers, 1979; Briers and Cochrane, 1979; Cornejo-Rodriguez, 1983 and Rosenbruch, 1985). The history of the test has been wonderfully described by Ronchi (1962a, 1962b, 1964) himself.

The name Ronchigrams for the patterns observed with a Ronchi ruling was coined by Schulz (1948). It is interesting that an attempt has been made to use this method for the measurement of the optical transfer function of lenses (Adachi, 1962b) and even their chromatic aberration (Toraldo di Francia, 1942b; Malacara and Cornejo, 1971). Salzmann (1970) used it to evaluate the quality of laser rods. Stoltzmann (1978) has applied this test to evaluate the quality of large-aperture flat mirrors, Assa et al. (1977) measured slope and curvature contours of flexed plates, Brookman et al. (1983) measured Gaussian beam diameters, and Kasana et al. (1984) have measured glass constants. Ronchi rulings have also been used in nontraditional configurations to evaluate optical components; for example, Patorski (1979) measured the wavefront curvature of small-diameter laser beams, using the Fourier imaging phenomena, and Kessler and Shack (1981) performed dynamic optical tests of a high-speed polygon. More recently, Hegeman et al. (2001) applied this test for testing components for EUV lithography; Xu et al. (2004) performed measurements of cylinders and Agureev (1995) studied gradient index profiles.

New analysis and proposals to improve the sensitivity and range of applications of the Ronchi test have recently been done by several authors; for example, Patorski (1984) described the reversed path Ronchi test, Patorski (1986) described a method using spatial filtering techniques, and Lin et al. (1990) described a quantitative three-beam Ronchi test.

9.2. GEOMETRICAL THEORY

The Ronchi test has two equivalent models; one is geometrical, interpreting the fringes as shadows of the ruling bands, and the other is physical, interpreting the fringes as shadows due to diffraction and interference. It will be shown in this chapter

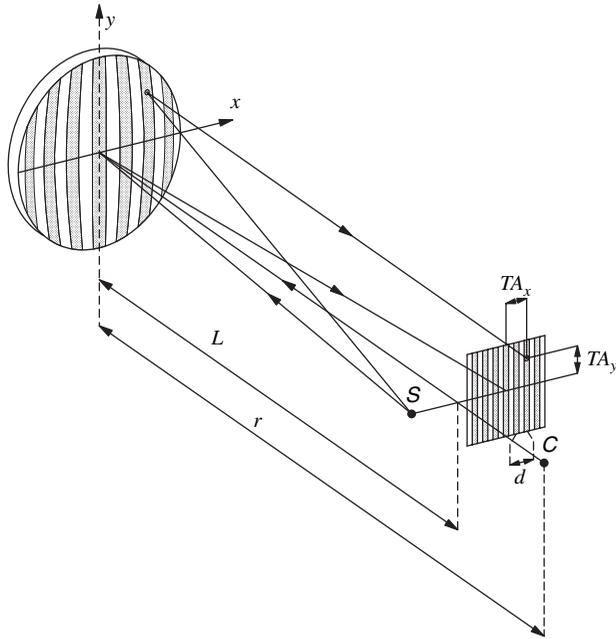


FIGURE 9.1. Geometry of the Ronchi test.

that, when the frequency of the ruling is not very high, the two models arrive at the same result. Both the geometrical and physical models were described by Ronchi in his original paper. A good treatment of the geometrical model was developed by Jentsch (1928), and it is briefly explained by Martin (1960).

As explained by Malacara (1965c), the Ronchi test really measures the transverse aberration TA in a direct way as shown in Figure 9.1. In this figure both the object and the image are on the optic axis, so the transverse aberration is measured from the axis and can be seen to include defocusing as well as other aberrations.

The wave aberration is defined at the exit pupil of the optical system under test, using a formula given by Rayces (1964), as

$$\frac{\partial W}{\partial x} = -\frac{TA_x}{r - W}; \quad \frac{\partial W}{\partial y} = -\frac{TA_y}{r - W} \tag{9.1}$$

For all practical purposes very accurate results can be obtained if we write formula (9.1) as equal to

$$\frac{\partial W}{\partial x} = -\frac{TA_x}{r}; \quad \frac{\partial W}{\partial y} = -\frac{TA_y}{r} \tag{9.2}$$

where r is the distance from the exit pupil or surface under test to the Ronchi ruling. If the ruling defocusing is small, this distance can be approximated by the radius of

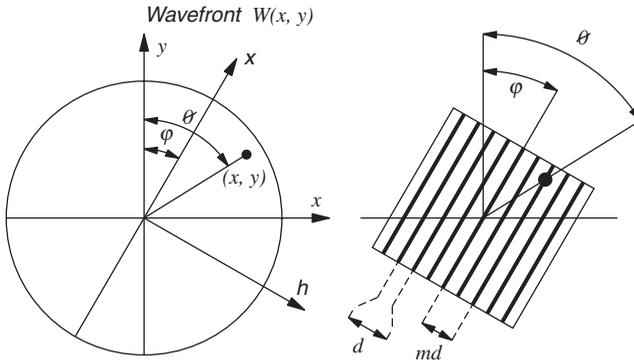


FIGURE 9.2. Wavefront and ruling orientation.

curvature of the wavefront. Thus, if we assume a Ronchi ruling with spacing d between the slits for a point (x, y) on the m th fringe, we may write, in general,

$$\frac{\partial W}{\partial x} \cos \varphi - \frac{\partial W}{\partial y} \sin \varphi = -\frac{md}{r} \quad (9.3)$$

where it is assumed that the ruling lines are inclined at an angle φ with respect to the y axis, as shown in Figure 9.2. This is the basic formula for the geometrical model of the Ronchi test.

9.2.1. Ronchi Patterns for Primary Aberrations

Ronchi patterns obtained with primary aberrations have been studied by many authors, as we will see later on. General treatments of them have been published by Schulz (1928), Crino (1933), Toraldo di Francia (1947, 1954) and Adachi (1960a, 1960b).

The wavefront of a system with primary aberrations can be written (see Appendix 3) as

$$W = A(x^2 + y^2)^2 + B y(x^2 + y^2) + C(x^2 + 3y^2) + D(x^2 + y^2) \quad (9.4)$$

where A , B , and C are the spherical aberration, coma, and astigmatism coefficients, respectively. The last coefficient D is the defocusing, given by the distance l' from the Ronchi ruling to the paraxial center of curvature, as

$$D = \frac{l'}{2r^2} \quad (9.5)$$

No tilt terms are included since the Ronchi test is insensitive to them. If we substitute Eq. (9.4) in Eq. (9.3), we obtain

$$\begin{aligned} &4A(x^2 + y^2)(x \cos \varphi - y \sin \varphi) + B[2xy \cos \varphi - (3y^2 + x^2) \sin \varphi] \\ &+ 2C(x \cos \varphi - 3y \sin \varphi) + 2D(x \cos \varphi - y \sin \varphi) = -\frac{md}{r} \end{aligned} \quad (9.6)$$

In the study of each of the aberrations, it will often be convenient to apply a rotation ψ to this expression by means of the relations

$$\begin{aligned} x &= \eta \cos \psi + \zeta \sin \psi \\ y &= -\eta \sin \psi + \zeta \cos \psi \end{aligned} \tag{9.7}$$

where η and ζ are the new coordinate axes.

Defocusing. By applying the rotation $\psi = \varphi$ to the defocusing term, we obtain

$$2D\eta = -\frac{md}{r} \tag{9.8}$$

Hence we obtain straight, equidistant bands, which are parallel to each other and to the ruling slits. The separation S between these bands on the wavefront under study is

$$S = \frac{d}{2Dr} \tag{9.9}$$

These bands are illustrated in Figure 9.3; an interesting reference is the work by Wan and Chang (1993).

Spherical Aberration with Defocusing. Spherical aberration patterns were first studied by Bocchino (1943) and Scandone (1933). By applying the rotation $\psi = \varphi$, we obtain for the spherical aberration and defocus terms

$$4A(\eta^2 + \zeta^2)\eta + 2D\eta = -\frac{md}{r} \tag{9.10}$$

where we can see that the axis of symmetry of the patterns is parallel to the ruling slits, as shown in Figure 9.4. The lines are cubics in η . This Ronchi pattern without defocusing is identical to the Twyman–Green interferogram for coma. Adding defocusing in this Ronchigram is equivalent to adding tilt in the Twyman–Green interferogram.

In the absence of defocusing, the central fringe is very broad and for this reason the paraxial focus is called the “fusiform” or uniform focus (Crino, 1939; Di Jorio, 1939a; Ricci 1939).

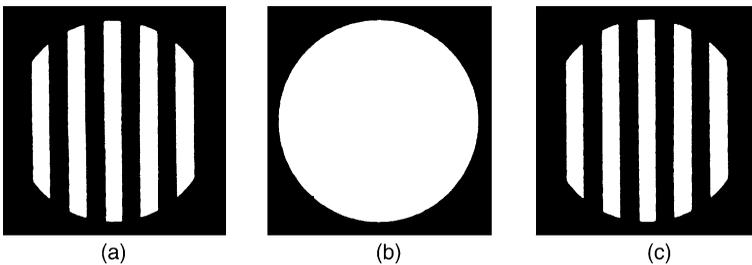


FIGURE 9.3. Ronchigrams with defocusing. (a) Outside of focus. (b) In focus. (c) Inside of focus.

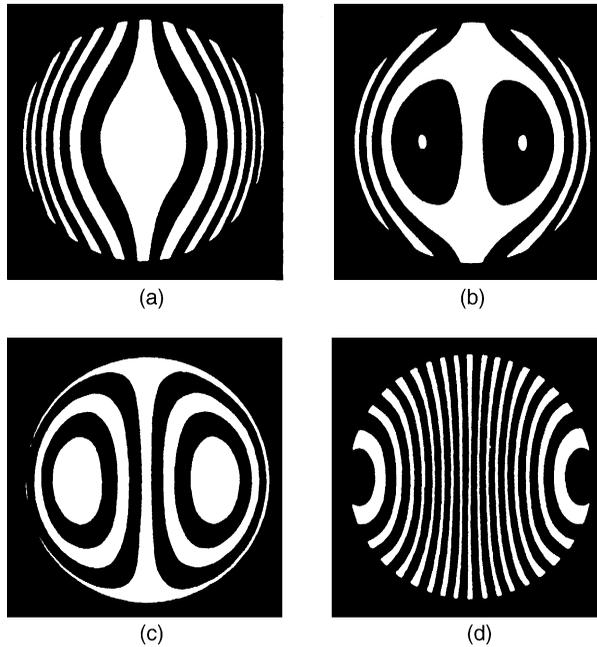


FIGURE 9.4. Ronchigrams with spherical aberration ($A = -20$). (a) Paraxial focus. (b) Medium focus. (c) Marginal focus. (d) Outside of focus.

Fifth-order spherical aberration Ronchigrams were studied by Bocchino (1940), Erdős (1959), and Scandone (1930).

Coma. Ronchigrams for coma were studied by Brusciaglioni (1932b), Villani (1930), and Villani and Brusciaglioni (1932). Applying a rotation $\psi = \varphi/2 + \pi/4$ to the coma term in Eq. (9.6), we obtain

$$B[-\eta^2(1 + 2 \sin \varphi) + \xi^2(1 - 2 \sin \varphi)] = -\frac{md}{r} \quad (9.11)$$

Different figures are obtained depending on the value of φ (ruling inclination with respect to the meridional plane) as follows (see Fig. 9.5):

$\varphi = 0^\circ$	Hyperbolas
$\varphi = 90^\circ$	Ellipses with semiaxes in the ratio $\sqrt{3}$ to 1
$0^\circ < \varphi < 30^\circ$	Hyperbolas inclined at an angle ψ
$\varphi = 30^\circ$	Straight bands
$30^\circ < \varphi < 90^\circ$	Ellipses inclined at an angle ψ

If we rewrite Eq. (9.11) as

$$B[-2(1 + \sin \varphi)(\eta^2 + \xi^2) + (\eta^2 + 3\xi^2)] = -\frac{md}{r} \quad (9.12)$$

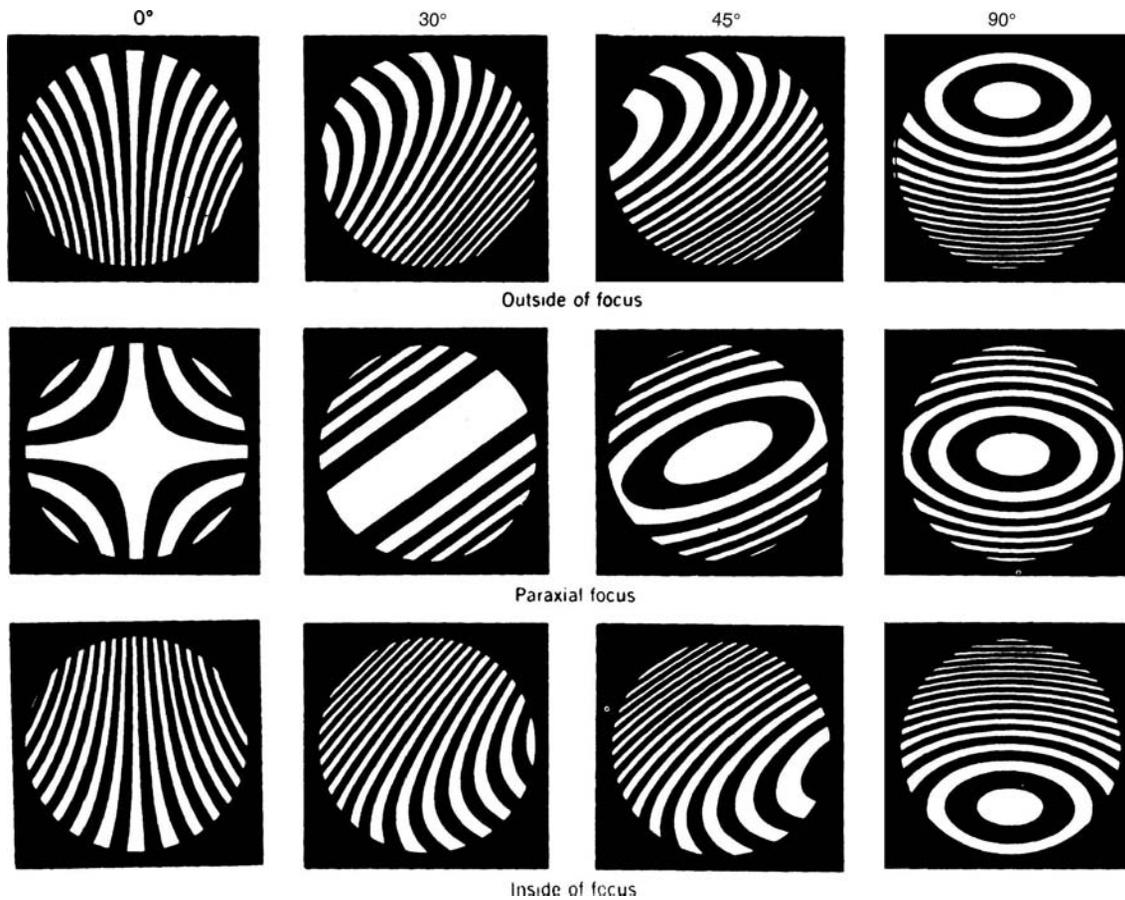


FIGURE 9.5. Ronchigrams with coma ($B = -30$).

we can see that this pattern is identical to that of a Twyman-Green interferogram for astigmatism with defocusing, where the magnitude of this apparent defocusing is given by the angle φ .

Astigmatism with Defocusing. Astigmatism Ronchigrams were extensively studied by Villani (1930), Scandone (1931a), Villani and Brusciaglioni (1932), Brusciaglioni (1932b), and Calamai (1938). By applying a rotation $\psi = \varphi$ to the corresponding terms in Eq. (9.6), we obtain

$$2C[\eta(2 - \cos 2\varphi) - \zeta \sin 2\varphi] + 2D\eta = -\frac{md}{r} \quad (9.13)$$

We can see that the Ronchigram is formed by straight, equidistant, parallel bands whose inclination α with respect to the ruling slits is given (see Fig. 9.6) by

$$\tan \alpha = \frac{C \sin 2\varphi}{D + C(2 - \cos 2\varphi)} \quad (9.14)$$

The intersections of the bands with the ζ axis are fixed, independently of the focusing term D , at equally separated points whose separation $\Delta\zeta$ is

$$\Delta\zeta = \frac{d}{2rC \sin 2\varphi} \quad (9.15)$$

The value of α changes with the focus term D (ruling position the optical axis), making the bands rotate as the ruling is moved along the optical axis. The effect, called “capriola,” was studied in detail by Brusciaglioni (1932a) and Scandone (1931b). The bands become perpendicular to the ruling slits when

$$D + C(2 - \cos 2\varphi) = 0. \quad (9.16)$$

a condition that occurs for a value of D / C between -1 (at the sagittal focus, when $\varphi = 0^\circ$) and -3 (at the tangential focus, when $\varphi = 90^\circ$). The precise values of D / C equal to -1 and -3 are excluded because at those points $\sin 2\varphi = 0$ and hence $\Delta\zeta \rightarrow \infty$, rendering the test insensitive to astigmatism. The maximum sensitivity is obtained when $\Delta\zeta$ is as small as possible with respect to the wavefront diameter. Thus the optimum angle for measuring the astigmatism is $\varphi = 45^\circ$, and then the bands become perpendicular to the ruling slits ($\alpha = 90^\circ$) when $D/C = -2$.

Spherical Aberration with Astigmatism and Defocusing. The patterns obtained with this combination of aberrations were studied by Scandone (1931a, 1931b). Applying a rotation $\psi = \varphi$ to the corresponding terms in Eq. (9.6), we obtain

$$4A\eta(\eta^2 + \zeta^2) + 2\eta[D + C(2 - \cos 2\varphi)] - 2C\zeta \sin 2\varphi = -\frac{md}{r} \quad (9.17)$$

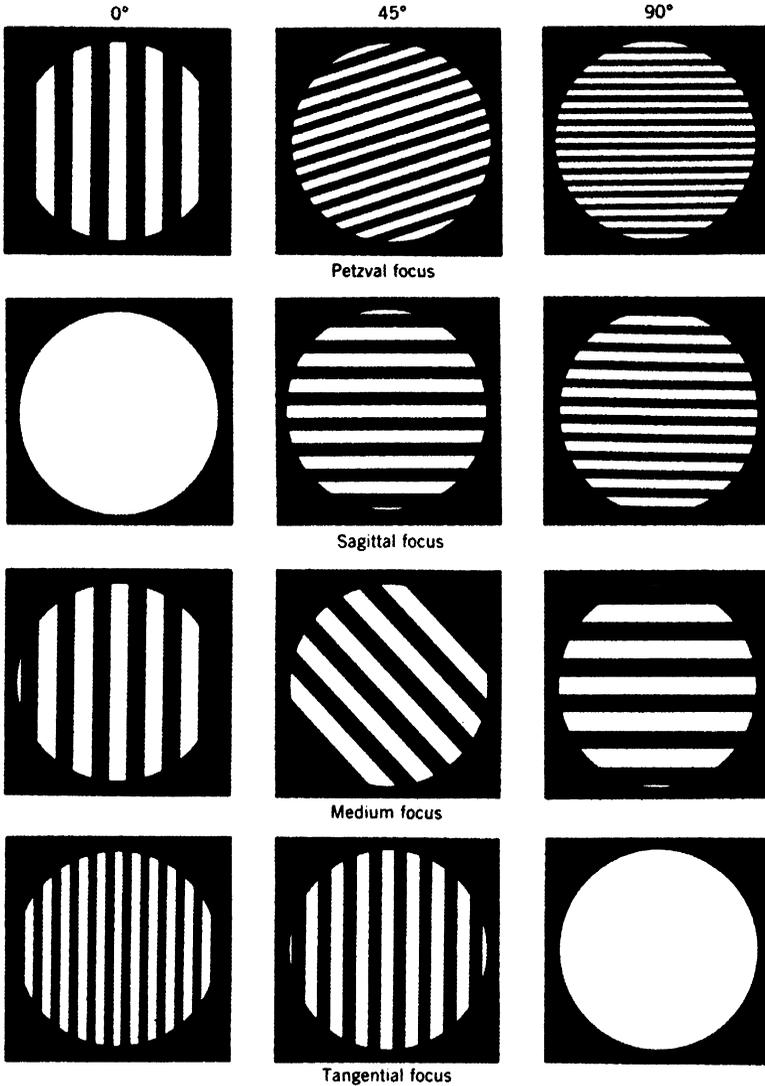


FIGURE 9.6. Ronchigrams with astigmatism ($C = -20$).

This combination of aberrations produces the serpentine fringes described by Scandone and shown in Figure 9.7. The first term comes from the spherical aberration and is identical to the Twyman–Green coma term. The second term comes from the defocusing D (ruling position) and the astigmatism C and is identical to a tilt term about the ξ axis in a Twyman–Green interferometer. The third term comes from the astigmatism and is identical to a tilt term about the η axis in a Twyman–Green interferometer.

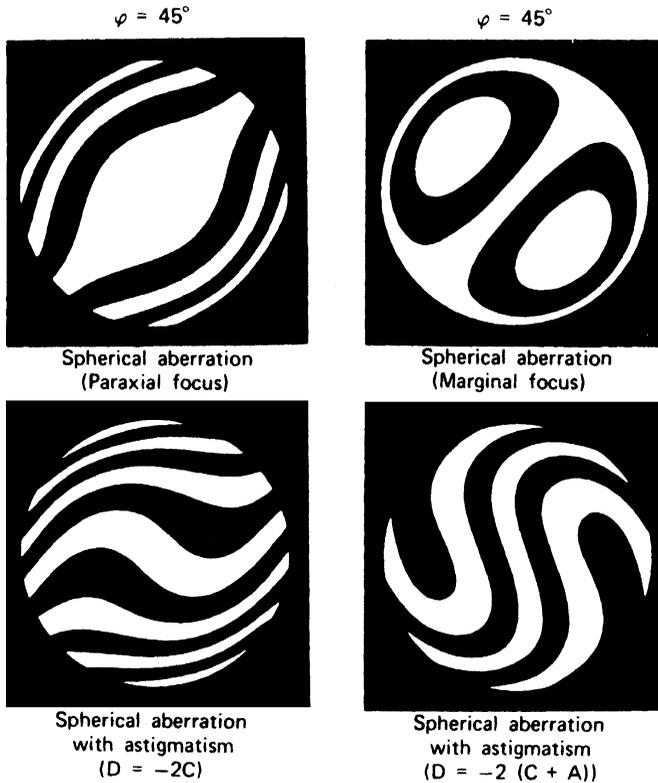


FIGURE 9.7. Ronchigrams with spherical aberration ($A = -10$) alone top combined with astigmatism ($C = -10$) (bottom).

Using this analogy between astigmatism in the Ronchi test and tilt in the Twyman–Green, Murty (1971) suggested adding an apparent tilt about the η axis to the Ronchi test by introducing astigmatism by means of a cylindrical lens in the beam, with its axis at 45° with respect to the ruling slits. To have apparent tilt only about the η axis, we need

$$\begin{aligned} D + C(2 - \cos 2\varphi) &= 0 \\ C \sin 2\varphi &\neq 0 \end{aligned} \tag{9.18}$$

for the particular case of $\varphi = 45^\circ$,

$$D = -2C \tag{9.19}$$

Expression (9.19) assumes that the defocusing term is just enough to cancel the effect of the astigmatism in producing zero tilt about the ξ axis. To produce an apparent tilt about the η axis of the Ronchigram at the marginal focus, we must introduce an additional defocusing equal to $-2A$, thus obtaining

$$D = -2(C + A). \tag{9.20}$$

9.2.2. Ronchi Patterns for Aspherical Surfaces

The Ronchi test is very useful for testing aspherical surfaces, including large mirrors for astronomical telescopes (Popov, 1972). Using Eq. (9.3) with $\varphi = 0^\circ$, we can compute the ideal Ronchigram for any aspherical surface, as defined in Appendix 1, by assuming valid the approximate relation

$$z(x, y) - z_0(x, y) = 2W(x, y) \tag{9.21}$$

where z is the sagitta for the aspheric surface and z_0 is for the osculating sphere. This method is, however, only approximate, not exact. An alternative and more accurate procedure is to trace rays by using the law of reflection. It is interesting that two early attempts to use the Ronchi test for aspherical surfaces were made by Waland (1938) and Schulz (1948).

Sherwood (1958) and Malacara (1965a, 1965b) independently showed (by two different methods) that, when the configuration in Figure 9.1 is used, the transverse aberration TA at the Ronchi ruling plane is given by

$$TA(S) = \frac{(l + L - 2z) \left[1 - \left(\frac{dz}{dS} \right)^2 \right] + 2 \frac{dz}{dS} \left[S - \frac{(l - z)(L - z)}{S} \right]}{\frac{l-z}{S} \left[1 - \left(\frac{dz}{dS} \right)^2 \right] + 2 \frac{dz}{dS}} \tag{9.22}$$

where S is the distance from the optical axis to the point on the mirror. In Figure 9.1, we can also see that

$$TA = \frac{md}{\sin \theta} \tag{9.23}$$

The Ronchigram is obtained by assigning many values of S and then calculating θ for different values of m . In general, one wishes to obtain the Ronchigram over a flat surface, parallel to and near the mirror, since this is what is obtained when a photograph is taken. In most cases the error introduced by considering the fringes over the mirror surface is very small; but when the radius of curvature of the mirror is small with respect to its diameter, the error becomes important.

The error is compensated for if, when plotting the Ronchigram, S_p , which is given by

$$S_p = S \left\{ 1 - \left[\frac{z(S_{\max}) - z(S)}{l - z(S)} \right] \left(1 - \frac{TA}{S} \right) \right\} \tag{9.24}$$

is used instead of the calculated S (Malacara, 1965b) as shown in Figure 9.8.

Ideal Ronchigrams for paraboloids have been computed for the use of amateur astronomers by Lumley (1960, 1961) and Sherwood (1960). Recently Cordero-Dávila et al. (2002), developed a computer program for arbitrary optical systems.

For easy interpretation of the Ronchigram, it is desirable to avoid any closed-loop fringes. This is possible only outside the caustic limits described in Appendix 1 as

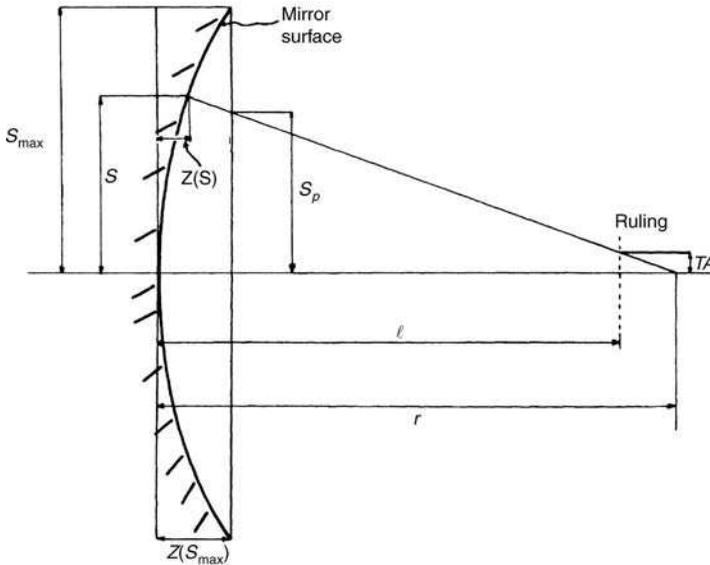


FIGURE 9.8. Projection of fringes over a plane.

shown in Figure 9.9, where it is evident that when the ruling is at the marginal focus, the only closed-loop fringes are inside the mirror and the outer one closes near the edge of the mirror.

It is also interesting to see that much information about the center of the mirror is lost when the fringes are curved, as at the paraxial focus. Therefore, to obtain the maximum information, the Ronchigram should be taken at the paraxial focus if $K > 0$ or at the end of the caustic if $K < 0$.

9.2.3. Null Ronchi Rulings

We have shown that, when the surface under test is aspherical, the Ronchi fringes are not straight but curved. The surface errors can then be computed from the deviations of the fringes from the desired shape. This procedure can give very good results when the data are processed in a computer, as will be shown in Section 9.3. If a computer is not used, it is easier for a technician to detect deviations from a straight line than from a curved line. Also, when the fringes are curved, the diffraction effects tend to diffuse the fringes, making the measurements even more difficult.

These disadvantages can be overcome by means of a special kind of Ronchi ruling with curved lines, the curvature of which compensates for the asphericity of the surface in order to produce straight fringes of constant widths. This idea was first qualitatively suggested by Pastor (1969) and was later developed quantitatively by Popov (1972) and independently by Malacara and Cornejo (1974c) using a ray-tracing program and then solving five linear equations. An approximate method to produce these null Ronchi rulings for paraboloids was devised by

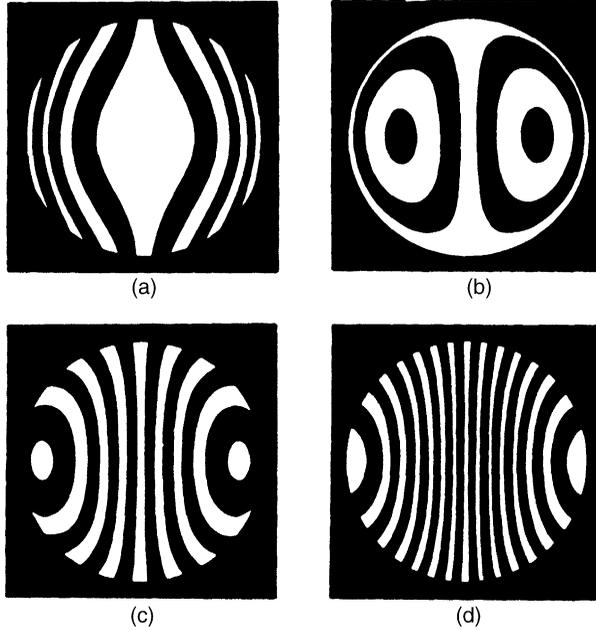


FIGURE 9.9. Ronchigram of an aspherical surface at different ruling positions on the caustic. (a) Paraxial focus. (b) Marginal focus. (c) Between marginal focus and end of caustic. (d) End of caustic.

Mobsby (1973, 1974). More recently Hopkins and Shagan (1977) used spot diagrams for designing the null gratings.

An alternative exact procedure uses Eq. (9.22) to compute $TA(S)$ for different values of S from the center to the edge of the mirror. Then we take the intersections of the desired straight fringe on the aspheric mirror with the circles on which the $TA(S)$ was computed. For each of those intersections there is an angle θ , which is the same on the ruling to be computed. The radial coordinate on the ruling is computed $TA(S)$ for that circle with radius S on the mirror.

Still another alternative and approximate method is based on the computation of the third-order transverse spherical aberration (Malacara and Cornejo, 1976b). If \bar{x} and $\bar{\rho}$ are the x and radial coordinates of a point over a fringe in the null Ronchi ruling, respectively, the ruling is computed with

$$\bar{\rho}^2 - \frac{r^2 \bar{x}^3}{Km^3(\Delta x)^3} - \frac{r \Delta r \bar{x}^2}{Km^2(\Delta x)^2} \tag{9.25}$$

where r is the vertex radius of curvature of the mirror, K is the conic constant, m is the fringe order number, Δx is the fringe separation on the mirror, and Δr is the displacement of the ruling from the paraxial focus, being positive if it is displaced outside the focus.

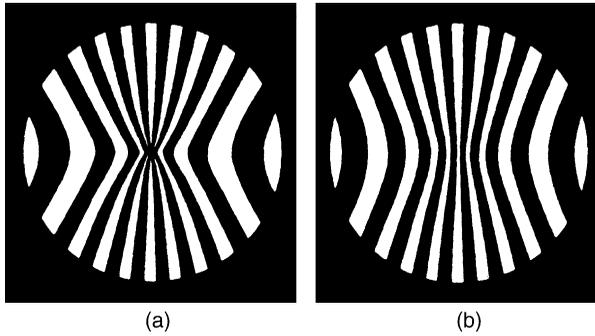


FIGURE 9.10. Two null Ronchi rulings, (a) At paraxial focus, (b) Inside paraxial focus.

The ruling is computed by assigning values of \bar{x} and then computing $\bar{\rho}$, starting in \bar{x}_{\min} , given by

$$\bar{x}_{\min} = -\frac{Km^3(\Delta x)^3}{r^2} - \frac{m(\Delta x)(\Delta r)}{r} \quad (9.26)$$

Here it is assumed that the light source is at the center of curvature. This approximate method has proved to be very accurate for most practical purposes.

Examples of computed null Ronchi rulings are given in Figure 9.10. Figure 9.11 shows a normal and a null Ronchigram. It is evident that the fringes, in addition to being straight, are much better defined in the null test [Fig. 9.11(b)].

A disadvantage of the null test, as compared to the normal test, is that the ruling must be positioned very precisely on the calculated place along the optical axis when the test is made. The reason is that the asphericity compensation depends very critically on that position. To draw a circle on the ruling so that its projection on the mirror coincides with the outer edge of that surface is of great help.

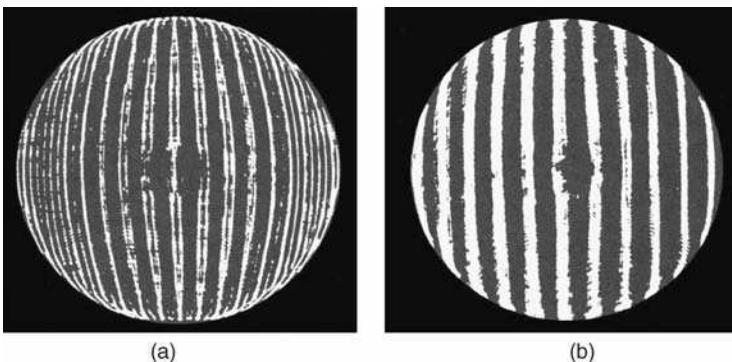


FIGURE 9.11. Normal and null Ronchigrams of an aspheric surface. (From Malacara and Cornejo, 1974c.)

Another important restriction is that the light source cannot be extended with a ruling covering it, as in the normal Ronchi test, but must be a point source. An interesting application of null rulings are the papers by Diaz-Uribe and Campos (2000), applied in a different configuration for convex surfaces, and Diaz-Uribe (2000) for off-axis parabolic mirrors as it will be described in Chapter 10.

9.3. WAVEFRONT SHAPE DETERMINATION

The wavefronts or mirror deformations can be determined from the Ronchigram. Probably Pacella (1927) was the first to attempt this and to identify the type of aberration. A qualitative idea about the mirror deformations can be very easily obtained (De Vany, 1965, 1970) just by observing the pattern. Figure 9.12 shows some possible patterns and their corresponding surface deviations are shown in Figure 9.13. Table 9.1 indicates the relationship between the Ronchigrams and the deformations. The identification and measurement of the primary aberration was treated in detail by Adachi (1960a, 1960b).

Following a combination of shearing diffraction interferometer theory and some geometrical equations of the Ronchi test, Briers did a thorough study about the theory and experiment of the Ronchi test in order to have semiquantitative results (Briers, 1979; Briers and Cochrane, 1979). Also De Vany extended his previous work and established a relation between observed patterns and surface defects (De Vany, 1978, 1980, 1981).

A more general treatment for surface with any kind of deformations is given in the following section.

TABLE 9.1. Relation between the ronchigrams of Figure 9.12 and the surface profile deformations shown in Figure 9.13.

Ronchigram	Surface			
	Reflecting		Refracting	
	Ruling outside of focus	Ruling inside of focus	Ruling outside of focus	Ruling inside of focus
1	B	A	A	B
2	A	B	B	A
3	D	C	C	D
4	C	D	D	C
5	F	E	E	F
6	E	F	F	E
7	H	G	G	H
8	G	H	H	G
9	J	I	I	J
10	I	J	J	I

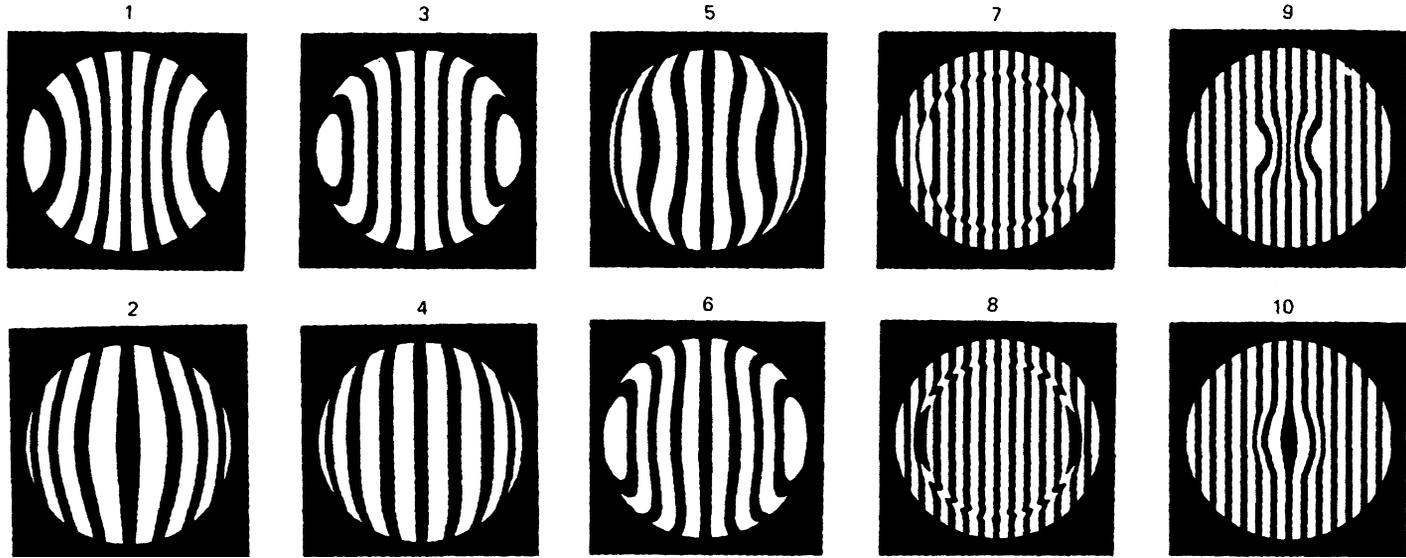


FIGURE 9.12. Ronchigrams for some typical surface deformations. (From Malacara and Cornejo, 1974a.)

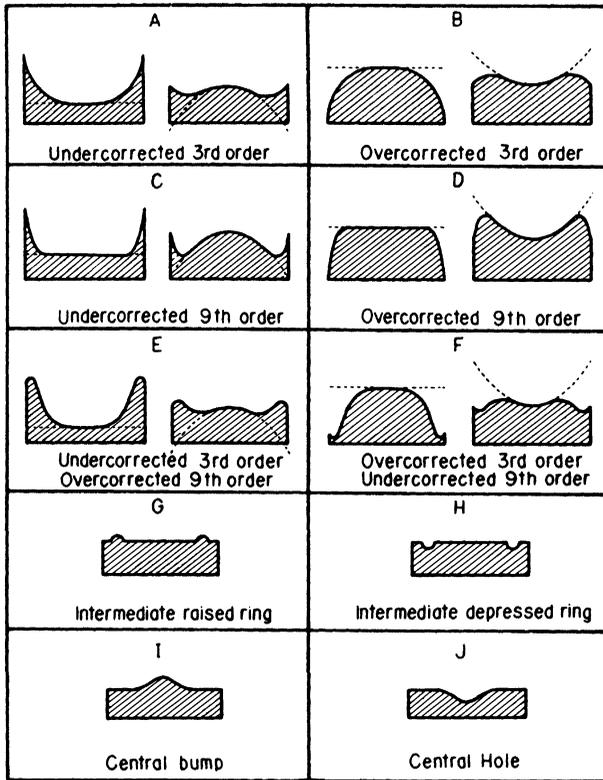


FIGURE 9.13. Surface deformations corresponding to Ronchigrams in Figure 9.12.

9.3.1. General Case

The only assumption we make in considering a surface without any symmetry is that it is smooth enough to be represented by the following two-dimensional polynomial of k th degree (Cornejo-Rodriguez and Malacara, 1976):

$$W(x, y) = \sum_{i=0}^k \sum_{j=0}^i B_{ij} x^j y^{i-j} \tag{9.27}$$

Then it can be shown that the partial derivatives of $W(x, y)$ with respect to x and y are

$$\frac{\partial W}{\partial x} = \sum_{i=0}^{k-1} \sum_{j=0}^i (j+1) B_{i+1, j+1} x^j y^{i-j} \tag{9.28}$$

and

$$\frac{\partial W}{\partial y} = \sum_{i=0}^{k-1} \sum_{j=0}^i (i-j+1) B_{i+1, j} x^j y^{i-j} \tag{9.29}$$

But from the fundamental relation for the Ronchi test in Eq. (9.3), we can write

$$\frac{\partial W}{\partial x} \cos \varphi - \frac{\partial W}{\partial y} \sin \varphi = -\frac{[m(x, y) - m_0(x, y)]}{r} d \quad (9.30)$$

where $m(x, y)$ is the measured value of m at point (x, y) in the real Ronchigram and $m_0(x, y)$ is the computed value of m at point (x, y) in the perfect theoretical Ronchigram using Eqs. (9.22) and (9.23).

We now define $m_x(x, y)$ and $m_y(x, y)$ as the values of $m(x, y)$ when the ruling orientations $\varphi = 0^\circ$ and $\varphi = 90^\circ$, respectively, are used. These are obtained with two Ronchigrams having mutually perpendicular rulings. Then, for these two patterns we may write

$$\frac{\partial W}{\partial x} = -\frac{[m_x(x, y) - m_0(x, y)]}{r} d \quad (\varphi = 0^\circ) \quad (9.31)$$

and

$$\frac{\partial W}{\partial y} = -\frac{[m_y(x, y) - m_0(x, y)]}{r} d \quad (\varphi = 90^\circ) \quad (9.32)$$

The difference function ($m_y - m_0$) can be fitted to a two-dimensional polynomial of $(k - 1)$ th degree by means of a least-squares procedure to find

$$\frac{\partial W}{\partial x} = \sum_{i=0}^{k-1} \sum_{j=0}^i C_{ij} x^j y^{i-j} \quad (9.33)$$

and

$$\frac{\partial W}{\partial y} = \sum_{i=0}^{k-1} \sum_{j=0}^i D_{ij} x^j y^{i-j} \quad (9.34)$$

If we compute Eqs. (9.28) and (9.29) with these two relations, we can see that

$$B_{ij} = \frac{C_{i-1, j-1}}{j} \quad \text{for } \begin{cases} i = 1, 2, 3, \dots, k \\ j = 1, 2, 3, \dots, i \end{cases} \quad (9.35)$$

and

$$B_{ij} = \frac{D_{i-1, j}}{i-j} \quad \text{for } \begin{cases} i = 1, 2, 3, \dots, k \\ j = 0, 1, 2, \dots, (i-1) \end{cases} \quad (9.36)$$

For the value of $m = 0$ only Eq. (9.36) can be used; therefore,

$$B_{i0} = \frac{D_{i-1, 0}}{i} \quad \text{for } i = 1, 2, 3, \dots, k \quad (9.37)$$

For the value $m = n$ only Eq. (9.35) can be used; therefore,

$$B_{ii} = \frac{C_{i-1,i-1}}{i} \quad \text{for } i = 1, 2, 3, \dots, k \tag{9.38}$$

For all other combinations of n and m , either Eq. (9.35) or Eq. (9.36) can be used. Therefore, to increase the accuracy, we take the average of both values:

$$B_{ij} = \frac{1}{2} \left(\frac{C_{i-1,j-1}}{j} + \frac{D_{i-1,j-1}}{i-j} \right) \quad \text{for } \begin{cases} i = 2, 3, \dots, k \\ j = 1, 2, 3, \dots, (i-1) \end{cases} \tag{9.39}$$

Once the coefficients B_{ij} have been found, the wavefront deviations $W(x, y)$ can be computed with Eq. (9.27) and then the mirror surface deviations can be found with Eq. (9.21).

Given the developments of PC and software in many respects, different ways for the analysis of the Ronchi test as a lateral shearing interferometer have been done more recently, for example by Wan and Lin (1990), Harbers et al. (1996), Seiichi and Nomura (2000), Elster and Weingartner (1999), Elster (2000), Van Brug (1997), Páez et al. (2000), Fischer and Stahl (1994), Servin et al. (1994), and Lei et al. (1993).

9.3.2. Surfaces with Rotational Symmetry

If we assume that the surface has a rotational symmetry, one Ronchigram with $\varphi = 0^\circ$ is enough to completely determine the shape of the surface. One approach is to measure the fringe intersections with the x axis (Malacara, 1965a, 1965b) to determine the transverse aberration $TA(S)$ at those points. The value of S is the distance from the center of the pattern to the fringe intersection with the x axis and TA is equal to md . Then, these measurements are fitted to a polynomial with only odd powers of S , using a least-squares procedure, and the wavefront is computed by a simple integration of $TA(S)$.

This method, however, does not provide enough data because no information is obtained for zones between the fringe intersections, and for this reason a polynomial interpolation is used. Unfortunately, with this procedure the surface profile is sometimes smoothed more than is desired.

Using a procedure by Malacara and Cornejo (1975), we may assume that the ideal fringes on the mirror are not necessarily straight but can be curved as shown in Figure 9.14. The solid lines are the actual shape of the fringes on an imperfect mirror, while the dotted lines show the ideal shape for a perfect surface. The residual transverse aberration $TA(S)$ can be defined by

$$TA(S) = TA_A(S) - TA_0(S) \tag{9.40}$$

where $TA_A(S)$ is the actual total transverse aberration for the imperfect surface and $TA_0(S)$ is the computed ideal transverse aberration. It can be seen in Figure 9.14 that

$$\frac{S_0}{x_0} = \frac{S}{x} = \frac{TA_A(S)}{TA_x} \tag{9.41}$$

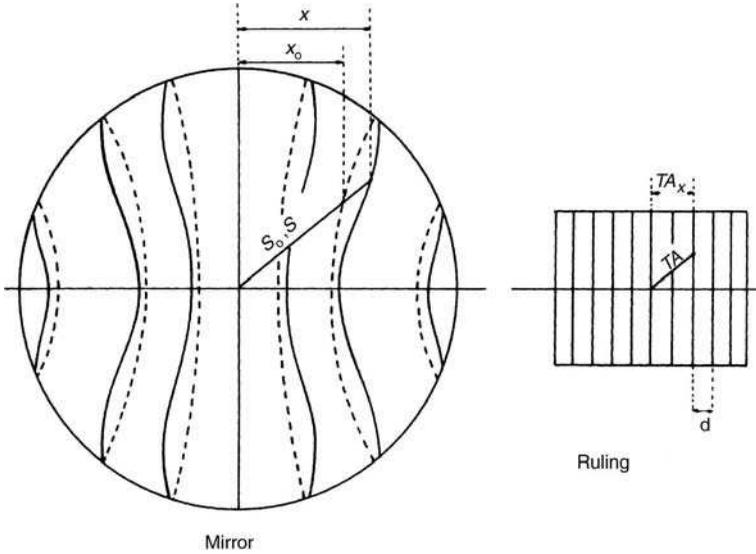


FIGURE 9.14 Ronchigram formation in a surface with rotational symmetry.

Since the points on the actual and ideal fringes are considered to be aligned with the center of the surface, they must correspond to the same point on the ruling. Therefore,

$$TA_0(S_0) = TA(S) \tag{9.42}$$

Now, since the lines on the ruling are straight, we can write $TA_x = md$, which is obtained from Eq. (9.41)

$$TA(S) = \frac{md}{x} S \tag{9.43}$$

which, substituted in Eq. (9.40), gives

$$TA(S) = \frac{md}{x} S - TA_0(S) \tag{9.44}$$

To determine $TA(S)$, we need to measure one value of x for each value of S . There are, however, in general, several values of x , one for each fringe, for a given value of S . An average value of x/m can be shown to be

$$\left(\frac{x}{m}\right)_{\text{average}} = \frac{\sum_{i=1}^N x_i m_i}{\sum_{i=1}^N m_i^2} \tag{9.45}$$

where N is the number of data points on the circle with radius S . The actual shape of the mirror is then found by numerical integration without the need for a polynomial fitting to the data.

9.4. PHYSICAL THEORY

As Ronchi (1923b) pointed out in his first paper, the Ronchi test can be considered from a physical point of view as an interferometer. At the beginning, many attempts were made to make a good physical theory model (Ronchi, 1924a, 1924b, 1926a, 1926b, 1927, 1928), from which a reasonably good explanation was produced that assumed that the Ronchi ruling really acted as a diffraction grating, producing many diffracted orders, each one giving a laterally sheared image of the pupil as shown in Figure 9.15. The theory was further developed by Di Jorio (1939a, 1939b, 1939c, 1939d, 1942, 1943), Pallotino (1941), and Toraldo di Francia (1941a, 1941b, 1942a, 1943a, 1943b, 1946), who obtained the exact shape of the fringes with the physical theory and demonstrated their similarity to the shadow fringes. They also investigated many small details related to this theory. A good account of the history of these developments has been given by Ronchi (1962a, 1962b, 1964).

9.4.1. Mathematical Treatment

A mathematical treatment using Fourier theory was originally developed by Adachi (1963) and later extended by Barakat (1969). It is assumed that the wavefront at the exit pupil in the plane $x_0 - y_0$ is represented by a complex function $F_0(x_0, y_0)$, which is zero outside the limits imposed by the aperture. If the system is evenly illuminated,

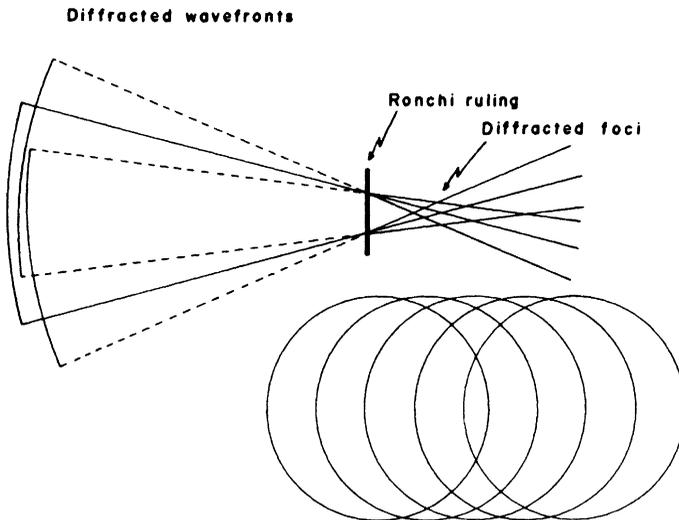


FIGURE 9.15. A physical model for the Ronchi test.

this function $F_0(x_0, y_0)$ inside the free portions of the exit pupil is given by $\exp [i2\pi W(x_0, y_0)/\lambda]$, where $W(x_0, y_0)$ is the wavefront deformation function. The phase deviations (wavefront deformations) given by $F_0(x_0, y_0)$ are measured with respect to a sphere with its center at the Ronchi ruling, in the plane $x_r - y_r$. If r is the radius of curvature of the reference wavefront, the field $U(x_r, y_r)$ at the ruling plane will be

$$U(x_r, y_r) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} F_0(x_0, y_0) \exp \left[-i \frac{2\pi}{\lambda r} (x_r x_0 + y_r y_0) \right] dx_0 dy_0 \quad (9.46)$$

As suggested by Barakat (1969), we can use the concept of spatial filtering to treat the effect of the Ronchi ruling and thus, consider it as a filtering mask in the Fourier transform plane $x_r - y_r$. Then, if the observation plane $x_1 - y_1$ is an image of the pupil plane $x_0 - y_0$, the amplitude in that plane will be given by

$$G(x_1, y_1) \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} U(x_r, y_r) M(x_r, y_r) \left\{ \exp \left[i \frac{2\pi}{\lambda r} (x_r x_1 + y_r y_1) \right] \right\} dx_r dy_r \quad (9.47)$$

where $M(x_r, y_r)$ is the ruling function acting as a filtering or modulating device. If we substitute the value of $U(x_r, y_r)$ from Eq. (9.46) in Eq. (9.47), we obtain

$$G(x_1, y_1) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} F_0(x_0, y_0) dx_0 dy_0 \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} M(x_r, y_r) \times \left[\exp \left\{ i \frac{2\pi}{\lambda r} [(x_1 - x_0)x_r + (y_1 - y_0)y_r] \right\} \right] dx_r dy_r \quad (9.48)$$

This expression is valid for any kind of modulating function. Now, however, we assume that the ruling is formed by straight, parallel, equidistant bands, and so we have

$$M(x_r) = \sum_{n=-\infty}^{\infty} B_n \exp \left(i \frac{2\pi n}{d} x_r \right) \quad (9.49)$$

where d is the ruling period, as usual. Then, substituting this expression in Eq. (9.48), we obtain

$$G(x_1, y_1) = \sum_{n=-\infty}^{\infty} B_n \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} F_0(x_0, y_0) dx_0 dy_0 \times \left\{ \int_{-\infty}^{\infty} \exp \left[i \frac{2\pi}{\lambda r} \left(x_1 - x_0 + \frac{\lambda r n}{d} \right) x_r \right] dx_r \right\} \times \left\{ \int_{-\infty}^{\infty} \exp \left[i \frac{2\pi}{\lambda r} (y_1 - y_0) y_r \right] dy_r \right\} \quad (9.50)$$

which, when the definition for the Dirac δ function,

$$\int_{-\infty}^{\infty} \exp[i(k - k_0)x] dx = \delta(k - k_0) \tag{9.51}$$

is used, it becomes

$$G(x_1, y_1) = \sum_{n=-\infty}^{\infty} B_n \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} F_0(x_0, y_0) \delta\left(x_1 - x_0 + \frac{\lambda rn}{d}\right) \delta(y_1 - y_0) dx_0 dy_0 \tag{9.52}$$

This can be shown to be equal to

$$G(x_1, y_1) = \sum_{n=-\infty}^{\infty} B_n F_0\left(x_1 + \frac{\lambda rn}{d}, y_1\right) \tag{9.53}$$

If the ruling is of the form

$$M(x_r) = 1 + \cos\left(\frac{2\pi n x_r}{d}\right) \tag{9.54}$$

we have $B_0 = B_1 = B_{-1} = 1$, and thus we have three laterally sheared images of the pupil. In general, however, we have many image laterally sheared by an amount $\lambda rn/d$ as in Figures 9.15 and 9.16. If the ruling has a periodic squarewave profile as in Figure 9.17, we can show, using Fourier theory, that the coefficients B_n are given by

$$B_n = (-1)^n \frac{\sin n\pi k_0}{n\pi} \tag{9.55}$$

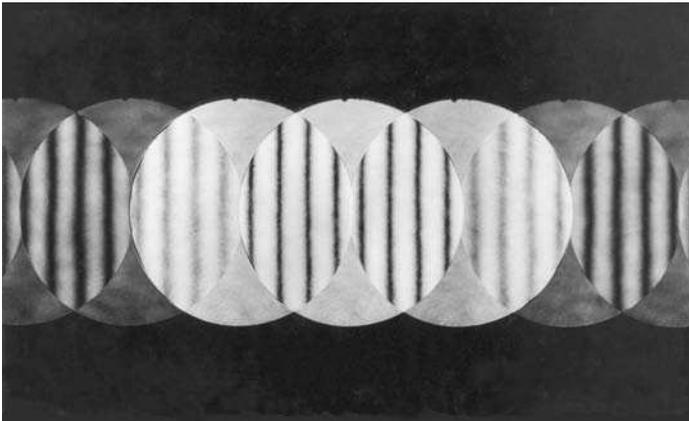


FIGURE 9.16. Interference fringes between different diffracted orders in a Ronchi ruling.

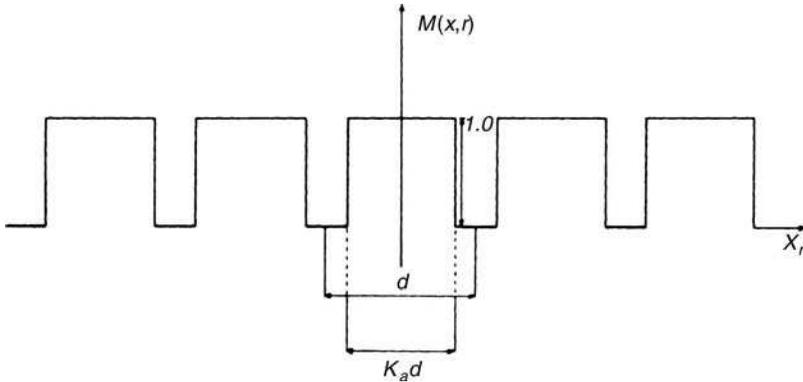


FIGURE 9.17. Square-wave profile in a Ronchi ruling.

where k_0 is the ratio between the width of the clear slits in the ruling and its period d as shown in Figure 9.17. If the width of the clear and dark bands is the same ($k_0 = \frac{1}{2}$), all even orders are found missing. Substituting Eq. (9.55) into (9.53) gives

$$G(x_1, y_1) = \sum_{n=-\infty}^{\infty} (-1)^n \frac{\sin n\pi k_0}{n\pi} F_0\left(x_1 + \frac{\lambda r n}{d}, y_1\right) \quad (9.56)$$

Let us now consider the case of a perfect wavefront with a radius of curvature R , different from the reference wavefront with radius r . The defocusing Δr is then equal to $R - r$. Then $F_0(x_1, y_1)$ may be represented by

$$F_0(x_1, y_1) = \begin{cases} \exp\left[i\frac{\pi}{\lambda r R}(x_1^2 + y_1^2)\right] \Delta r & \text{for } x_1^2 + y_1^2 < S_{\max}^2 \\ 0; & \text{for } x_1^2 + y_1^2 > S_{\max}^2 \end{cases} \quad (9.57)$$

where S_{\max} is the semidiameter of the aperture. Let us assume that the lateral shear $\lambda r n/d$ is very small compared with the semiaperture S_{\max} , so that the summation can be extended to large values of n . To find the amplitude profile of the fringes, we also take $y_1 = 0$; then we can show that

$$G(x_1) = \left[k_0 + 2 \sum_{n=-\infty}^N (-1)^n \frac{\sin n\pi k_0}{n\pi} \cos\left(\frac{2\pi n \Delta r}{Rd} x_1\right) \right. \\ \left. \times \exp\left(i\frac{\pi n^2 \lambda r \Delta r}{Rd^2}\right) \right] - \left[\exp\left(i\frac{\pi \Delta r}{\lambda R r} x_1^2\right) \right] \quad (9.58)$$

where N is the number of wavefronts that overlap in the region under study

9.4.2. Fringe Contrast and Sharpness

Expression (9.58) explains quantitatively the lack of sharpness of the Ronchi fringes, which is identical outside and inside the focus. It has been shown by Malacara (1990)

that for small defocusings, which are the ones we may find in practical arrangements, the bright fringes become sharper as the defocusing is increased. This means that the closer the fringes are, the sharper they appear. Thus, for the case of a parabolic mirror, the fringes are sharper around the edge of the pattern for inside of focus patterns and near the center for outside of focus patterns.

The fringes become very sharp when the phase shift between the first and zero orders in Eq. (9.58) is a multiple M of 2π , yielding

$$M \frac{2Rd^2}{\lambda r(R-r)} = 1 \quad (9.59)$$

This is the Talbot effect (1836) appearing in the Ronchi test, which has been reported by Malacara and Cornejo (1974b). The defocusing Δr given by this condition is extremely large even when M is 1. This condition is equivalent to having a fringe spacing equal to half the lateral shear S , which is as follows:

$$\alpha = \frac{\theta_s}{2M} \quad (9.60)$$

where θ_s is the angular lateral shear between any two consecutive orders and α is the angular separation between the fringes. This effect is illustrated in Figure 9.18.

We have another autoimaging condition, but with the contrast reversed, when $M = \frac{1}{2}$, that is, at half the Talbot defocusing. When $M = \frac{1}{4}$ and $M = \frac{3}{4}$, the fringe contrast disappears. This fringe blurring occurs when the fringe spacing is equal to twice the lateral shear.

Usually, the Ronchi test is interpreted using the geometrical approximation. This is valid as long as the lateral shear is small compared with the wavefront diameter. On the contrary, the number of fringes over Ronchi pattern cannot be very large. This means that the fringe spacing cannot become extremely small compared with the lateral shear.

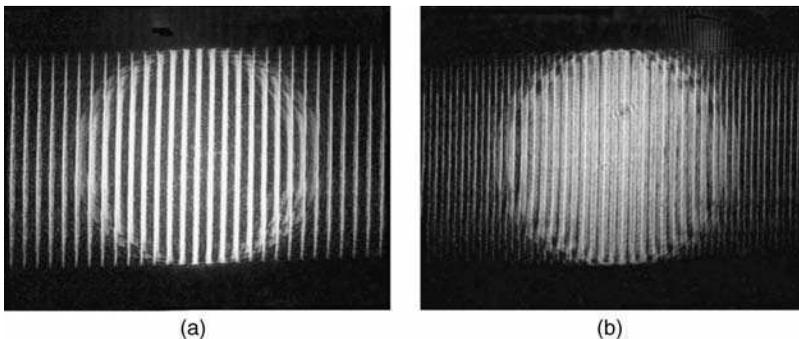


FIGURE 9.18. Talbot effect in the Ronchi test, (a) Ruling at the Talbot effect position, (b) Outside it. (From Malacara and Cornejo, 1974b.)

Lau (1948) showed that the autoimaging of gratings and rulings, or Talbot effect, also appears if an extended, incoherent, and periodical modulated light source with the proper period is used. This is done in practice by covering the extended light source with a ruling. In the Ronchi test, this type of illumination is frequently used as shown later in this chapter.

Numerous interferometers, based on both the Talbot and the Lau effects using gratings, have been invented. Lohmann and Silva (1971, 1972), Yokoseki and Susuki (1971a, 1971b), Silva (1972), Mallick and Roblin (1972), Hariharan et al. (1974), Patorski (1982), and Patorski (1984). These interferometers have been so successful because many applications have been found for them, for example, for measuring focal lengths (Nakano and Murata, 1985; Bernardo and Soares, 1988). Some interferometers have used a combination of the Talbot effect with Moiré patterns (Glatt and Kafri, 1987; Keren et al., 1988).

Experimental attempts have been made to enhance the sharpness of the bands by changing the relative width k_0 of the slits as shown in Figure 9.17. Some improvements have been obtained (Murty and Cornejo, 1973; Cornejo-Rodriguez et al., 1978). Studies about the contrast and sharpness of the observed fringes have also been carried out by Patorski and Cornejo-Rodriguez (1986a), Malacara (1990), and recently by Luna-Zayas et al. (2006).

The phase ruling is one in which, instead of dark and clear bands, there is a periodic change in the optical thickness. Such rulings can be made in many ways as described by Vogl (1964) and Ronchi (1965) or by any other of the conventional methods to make phase plates. This kind of grating is especially useful when the period is high enough such that no more than two wavefronts overlap (Fig. 9.19).

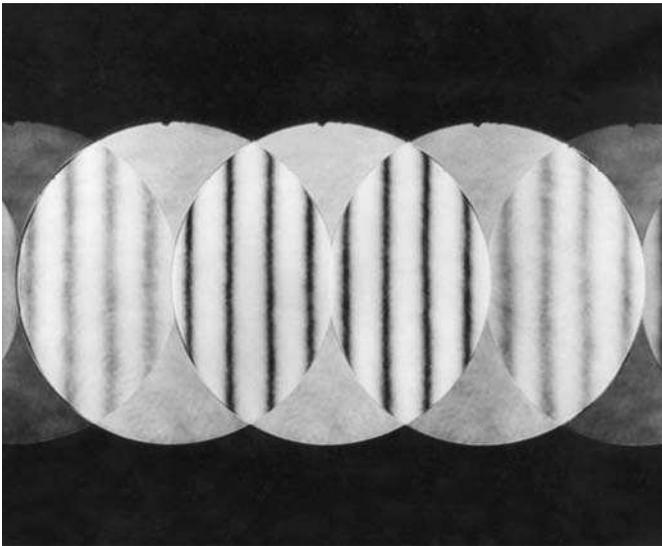


FIGURE 9.19. Ronchigram with a phase grating.

It is interesting that Pallotino (1941) shows experimentally that, with phase gratings, patterns placed symmetrically with the zero order are complementary; in other words, a dark fringe in one pattern corresponds to clear fringe in the other as in Figure 9.19. Later, Toraldo di Francia (1941a, 1941b) theoretically proved this effect. The explanation uses the fact that all diffracted orders in a phase ruling have a phase shift of magnitude $m\lambda/4$ with respect to the zero-order beam.

The sensitivity of the Ronchi test has been a subject of research almost since its invention by Ronchi (1930, 1940) himself and Bruscaioni (1933, 1939). More recently, Cornejo and Malacara (1970) studied the subject.

9.4.3. Physical versus Geometrical Theory

If the irradiance distribution over the exit pupil is a constant, we can write the pupil function $F_0(x_0, y_0)$ as

$$F_0(x_0, y_0) = \exp \left[i \frac{2\pi}{\lambda} W(x_0, y_0) \right] \quad (9.61)$$

where $W(x_0, y_0)$ for the primary aberrations is given by Eq. (9.4).

Let us now assume that the ruling has a period d small enough to produce interference patterns of no more than two overlapping beams as shown in Figure 9.19. Then, from Eq. (9.48), the interference pattern between two consecutive orders n_1 and n_2 can be shown to be given by

$$\begin{aligned} |G(x_1, y_1)|^2 = & B_{n_1}^2 + B_{n_2}^2 + 2B_{n_1}B_{n_2} \\ & \times \cos \left\{ \frac{2\pi}{\lambda} \left[W \left(x_1 + \frac{\lambda r n_1}{d}, y_1 \right) - W \left(x_1 + \frac{\lambda r n_2}{d}, y_1 \right) \right] \right\} \end{aligned} \quad (9.62)$$

It can be seen here that a bright fringe appears whenever

$$W \left(x_1 + \frac{\lambda r n_2}{d}, y_1 \right) - W \left(x_1 + \frac{\lambda r n_1}{d}, y_1 \right) = -m\lambda \quad (9.63)$$

where m is any positive or negative integer, including zero. If we now take the center between the two shears of the wavefront as a new origin, we find

$$W \left[x_1 + \frac{\lambda r (n_2 - n_1)}{2d}, y_1 \right] - W \left[x_1 - \frac{\lambda r (n_2 - n_1)}{2d}, y_1 \right] = -m\lambda \quad (9.64)$$

Expanding now in a Taylor series about the origin and taking $n_2 - n_1 = 1$, we have

$$\frac{\partial W(x_1, y_1)}{\partial x_1} + \frac{1}{6} \left(\frac{\lambda r}{2d} \right)^2 \frac{\partial^3 W(x_1, y_1)}{\partial x_1^3} = -\frac{m\lambda}{r} \quad (9.65)$$

When the third derivative term disappears, this expression becomes identical to Eq. (9.3) (with $\varphi = 0^\circ$), which was the basis of the geometrical theory of the Ronchi test. Therefore, in such cases the physical and geometrical patterns coincide. This result was derived by Toraldo di Francia (1947, 1954) who pointed out that the third derivative cancels out when the focus shifts and third-order astigmatism and coma cancel out when no power of x higher than the second appears [see Eq. (9.4)].

With regard to spherical aberration, we find substituting the expression for spherical aberration with defocusing from Eq. (9.4) in Eq. (9.65),

$$4A(x_1^2 + y_1^2)x_1 + 2\left(A\frac{\lambda^2 r^2}{2D^2} + D\right)x_1 = -\frac{md}{r} \quad (9.66)$$

which is identical to Eq. (9.10) for the geometrical pattern, except for a difference in the focus shift coefficient. The fusiform of uniform focus is obtained when the apparent focus shift is zero, giving

$$D = -A\frac{\lambda^2 r^2}{2d^2} \quad (9.67)$$

but $\lambda r/2d$ is the distance ρ_c from the center of one of the wavefronts to the center of the pattern. Therefore $D = -2A\rho_c^2$, which is the condition necessary to place the focus of the zone with radius ρ_c at the ruling. This effect was studied by Di Jorio (1939d).

9.5. PRACTICAL ASPECTS OF THE RONCHI TEST

The experimental apparatus used to perform the Ronchi test is of many types, as described, for example, by Kirkham (1953). The basic arrangement is shown in Figure 9.20(a). The light source is a white-light tungsten lamp illuminating a pinhole or a slit parallel to the ruling lines. However, De Vany (1974) used a ball bearing as a source, and Patorski and Cornejo-Rodriguez (1986b) used sunlight. Recently, a light emitting diode (LED) has been used by Nuñez et al. (2001). The advantage of a slit is that more light is used, producing a brighter pattern. The justification for using a slit light source is that any point source on the slit produces the same pattern because the corresponding images fall on the ruling displaced only along the grating lines, with no lateral shifts. In the beginning a slit source was used for the test, but a little later Anderson and Porter (1929) suggested allowing the grating to extend over the lamp, as shown in Figure 9.20(b), instead of employing a slit source. A significant advantage, in addition to the greater luminosity of the pattern, is that there is no need to worry about the parallelism between the ruling lines and the slit. The light source is then a multiple slit and the images of the slits are separated by a distance equal to the ruling period. This property justifies the use of the ruling over the light source, which may also be justified, however, in a more formal way as follows.

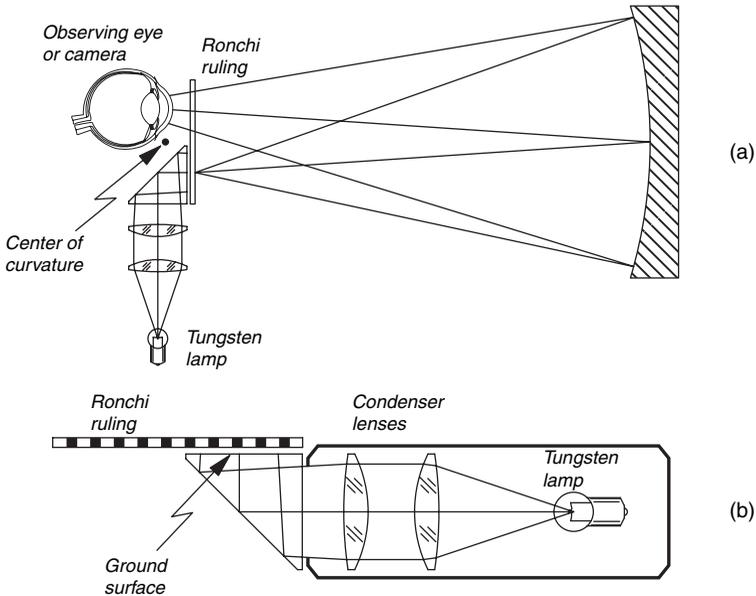


FIGURE 9.20. Diagram of an instrument to observe Ronchi patterns.

The well-known Van Cittert–Zernike theorem establishes that the coherence function of a source is given by the Fourier transform of the intensity distribution at the light source. Since the light source in this case is a ruling, the Fourier transform is as shown in Figure 9.21. Then, if two points over the wavefront are separated by the distance Δx between any two peaks in the wavefront, they are capable of producing interference fringes with good contrast. If the ruling on the light source and the examining Ronchi ruling are in the same plane, the distance between the peaks on the coherence function is just the lateral shear length between the diffracted wavefronts. This result completely justifies the use of the Ronchi ruling covering the extended light source.

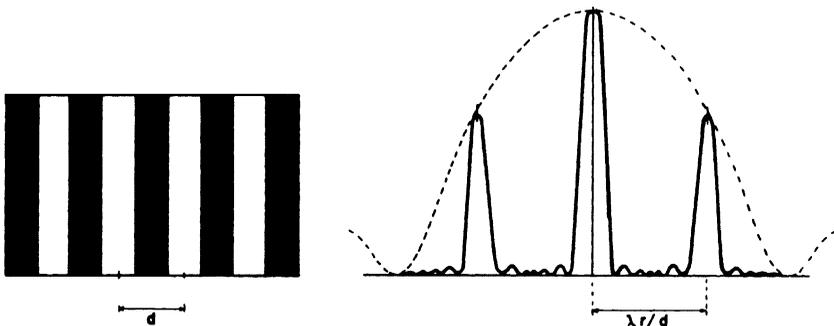


FIGURE 9.21. Ronchi ruling and its coherence function over the mirror being examined.

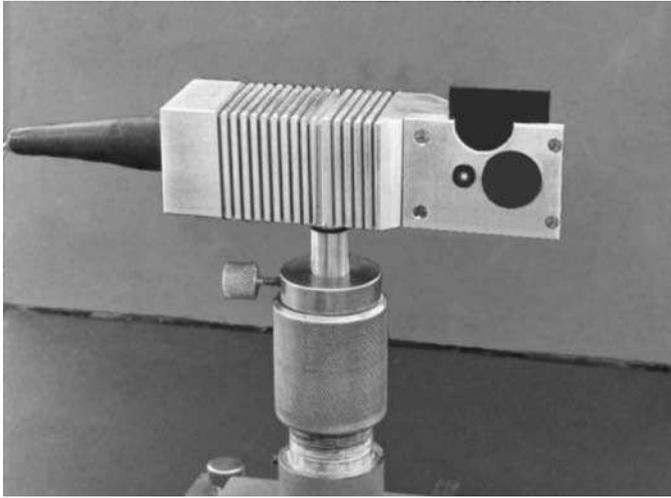


FIGURE 9.22. Instrument to observe Ronchi patterns.

It is interesting that, when the dark and clear bands on the ruling have the same width ($k_0 = \frac{1}{2}$), there is no coherence between points on the wavefront separated by a distance Δx equal to $2\lambda r/d$. Therefore, both first-order wavefronts interfere with the zero order but do not interfere between themselves. However, when using a point source or slit, all overlapping wavefronts interfere with each other (Braat and Janssen 1999).

The null Ronchi ruling described in Section 9.2.3 does not produce the same lateral shear for all points of the wavefront; therefore, the illuminating wavefront must have coherence between any two points at any distance between them. This is possible only if a point source is used. It must be kept in mind that a gas laser is the optimum light source because of its very high coherence.

Some Ronchigrams taken with the instrument in Figure 9.22 are shown in Figure 9.23. Although obvious from all past discussions, it should be pointed out

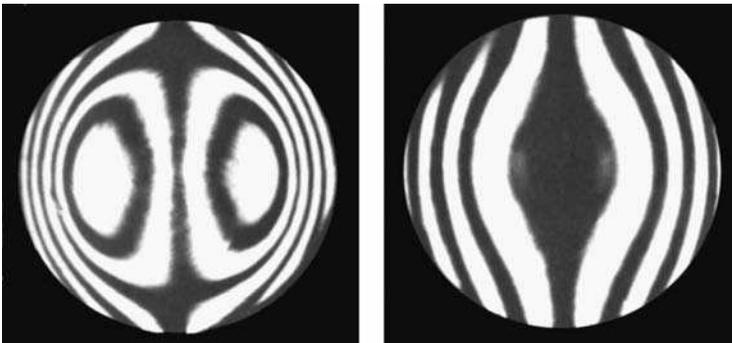


FIGURE 9.23. Ronchigrams taken with the instrument in Fig. 9.22

once more that the shape of the Ronchi fringes is not directly related to the shape of the surface but rather is related to the transverse aberration function. The photographs were taken with a normal photographic camera focused on the surface plane because the fringes we want to study are precisely in that plane. Nowadays a CCD detector with a lens system can be used, and the Ronchigram can be stored in a computer, for further fast processing, sometimes combined with the use of a liquid crystal device (Luna-Zayas et al. 2006).

When testing an aspheric mirror, care should be taken that the transverse dimensions of the caustic (see Appendix I) are small enough so that all the light goes through the observing eye or camera. Many a times the waist of the caustic is so big (i.e., in large telescope mirrors) that it cannot pass through the pupil of the eye, although it can pass through the lens of a camera.

Another point of great practical interest is that the light passing through each slit on the Ronchi ruling interferes with that passing through the other slits, producing an effect of many laterally sheared pupils. This effect reduces the accuracy of the test since the Ronchi pattern cannot be precisely located over the pupil, simply because there are many interfering images of it (Malacara, et al. 1990). An alternative method for improving the accuracy is to use a wire instead of the Ronchi ruling. The wire position is laterally displaced from the optical axis, and such displacement is analogous to the ruling period. The distance of the wire from the optical axis should be accurately measured.

If the Ronchi test is done with care and the patterns are correctly interpreted, the test is a very powerful tool to measure quantitatively or qualitatively the degree of corrections needed in an optical surface or lens.

New instruments to perform the Ronchi test and grating continuously developed. For example, an instrument based on Ronchi test was reported by Kamalov et al. (1980). The manufacturing of Ronchi gratings has been treated by Kuindzhi et al. (1980), Patorski (1980), Thompson (1987) and Steig (1987). Other interesting applications are those developed by Royo et al. (2000), Schreiber and Schrider (1997), Sirohi et al. (1995), and Blakley (1994).

9.6. SOME RELATED TESTS

9.6.1. Concentric Circular Grid

Instead of straight lines on the ruling, it is also possible to use a grid with circular lines. This modification was studied originally by Scandone (1931c, 1932) and more recently by Murty and Shoemaker (1966), whose treatment is presented here. Patterns with this kind of ruling differ from the normal Ronchi patterns. In this case, the equation of the fringes can be obtained in the same manner as in Eq. (9.3), giving

$$\left(R \frac{\partial W}{\partial x} - \bar{x}_0\right)^2 + \left(R \frac{\partial W}{\partial y} - \bar{y}_0\right)^2 = M\rho_1^2 \quad (9.68)$$

where (\bar{x}_0, \bar{y}_0) is the center of the set of concentric circles, ρ_1 is the radius of the innermost circle, and M is an integral number defined by

$$M = \begin{cases} n^2 & \text{for equally spaced circles} \\ n & \text{for a Fresnel zone plate} \end{cases}$$

n being a positive integer. If we use the wavefront W for primary aberrations as in Eq. (9.4), we find

$$\begin{aligned} & \{R[4A(x^2 + y^2)x + 2Bx + 2Cx + 2Dx + F] - \bar{x}_0\}^2 \\ & + \{R[4A(x^2 + y^2)y + B(3y^2 + x^2) \\ & + 6Cy + 2Dy + E] - \bar{y}_0\}^2 = M\rho_1^2. \end{aligned} \quad (9.69)$$

As pointed out by Murty and Shoemaker, this expression becomes extremely complicated, except for a few trivial cases. Some typical patterns are shown in the paper by the same authors.

It is important to point out that this test must be performed with a point source and that a slit cannot be used.

9.6.2. Phase Shifting Ronchi Test

This class of test was proposed by Thompson (1973), under the name AC grating interferometer to test laser wavefronts. It is basically a dynamic Ronchi test in which a rapid treatment of the data can be made by means of a minicomputer.

The fringe pattern is not stationary because the ruling is uniformly translated over its own plane, in the direction perpendicular to the ruling slits. Then, instead of photographing a stationary pattern, the moving fringes are recorded by a photoelectric detector such as a silicon diode array of a vidicon tube.

Each individual detector of the array gives an electrical periodical output, all with the same frequency, depending on the velocity of the moving Ronchi ruling. The relative phase of the signals from the detectors depends on the position of the detector over the wavefront and also on the aberrations W . A precise measurement of these phases is enough to reconstruct the fringe pattern and from it the wavefront.

To have a moving Ronchi ruling, Thompson (1973) suggested using a drum ruling with a circumference of about 40 cm. The slight curvature of the ruling is not very important.

Phase shifting Ronchi tests have been described by Koliopoulos (1980), Yatagai (1984), Omura and Yatagai (1988), and Wan and Lin (1990).

9.6.3. Sideband Ronchi Test

This is essentially a holographic test, as devised by Malacara and Cornejo (1976a). The method requires a special ruling (or hologram) that is fabricated as illustrated in Figure 9.24, where two concave mirrors reflect the light of a point source S over a common area on a photographic plate P . In front of mirror M_2 a very coarse ruling

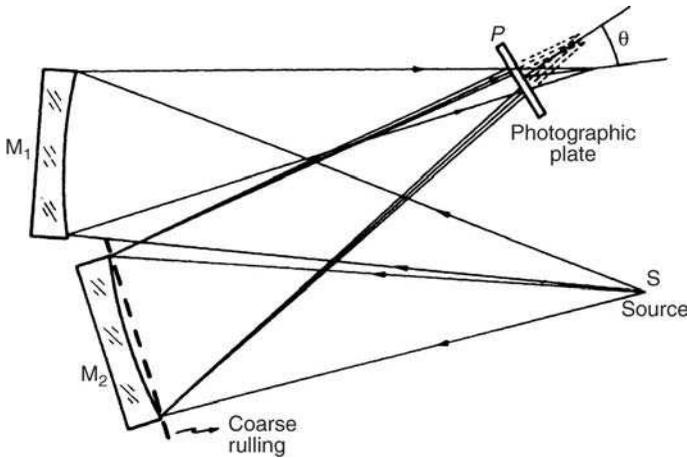


FIGURE 9.24. Setup to make grating for side band Ronchi test.

with straight bands is placed. The aim is to make a hologram in order to reconstruct the image of M_2 with the coarse ruling in front of it using the light from M_1 as a reference beam.

Once the hologram is developed, it is replaced in the beam from M_1 , but this time without mirror M_2 . If the mirror M_1 used during the reconstruction is identical to the mirror M_1 employed in forming the hologram, a set of straight fringes will be seen where mirror M_2 was placed. If the mirror is different, the observed fringes will not be straight, indicating a wave aberration given by twice the difference between the two mirror surfaces M_1 .

As shown in Figure 9.25, the main advantage of this method is that the bands are very sharp and well defined, unlike those in the normal Ronchi test. Further studies and modifications on the sideband Ronchi test have been reported by Malacara and Josse (1978), Schwider (1981), and Patorski and Salbut (1984).

9.6.4. Lower Test

This test was devised by Lower (1937), who pointed out that it slightly resembles an inverted Ronchi test. The normal Ronchi test cannot be used to test a parabolic mirror with a small f ratio, and for obvious reasons it would be even more difficult to test a fast Schmidt camera. Originally, Lower used his test to measure the degree of correction of a fast Schmidt camera. The test consists in placing a slit at the focus of the system under test and viewing image (at infinity) from a relatively large distance. This image may be considered as being localized on the aperture of the system, and it is a straight line when there is good correction. The straightness of the line remains as the eye traverses the aperture.

This test is very rapid and simple, but for many practical reasons, such as the finite observing distance, the accuracy is very limited. The factors that limit the sensitivity and accuracy of the Lower test have been analyzed by Rank et al. (1949) and Yoder

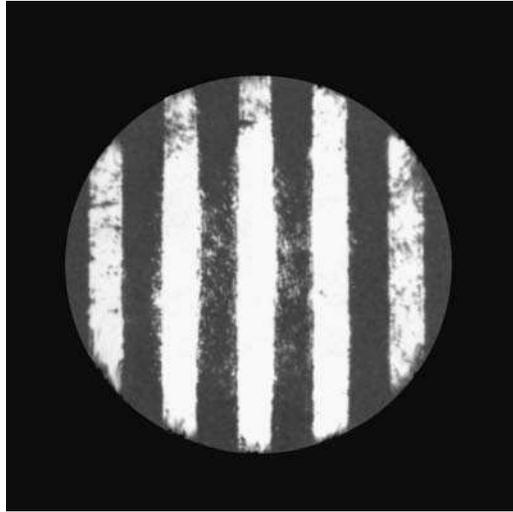


FIGURE 9.25. Fringes in sideband Ronchi test. (From Malacara and Cornejo, 1976a.)

(1959). A related test for paraboloidal reflectors, using a similar principle but a screen line, was developed by Hamsher (1946).

9.6.5. Ronchi–Hartmann and Null Hartmann Tests

Even though the Ronchi and Hartmann tests have been studied separately, Cordero-Davila et al. (1990) have shown the existence of a close relation between both the tests. In order to prove that relation, it was shown that for the two tests the screen or filter plane can be exchanged with the plane where the pattern is observed or recorded; thus one can go from one test to the other. The physical basis for this result rests in the geometrical optics concept of propagation of light between both

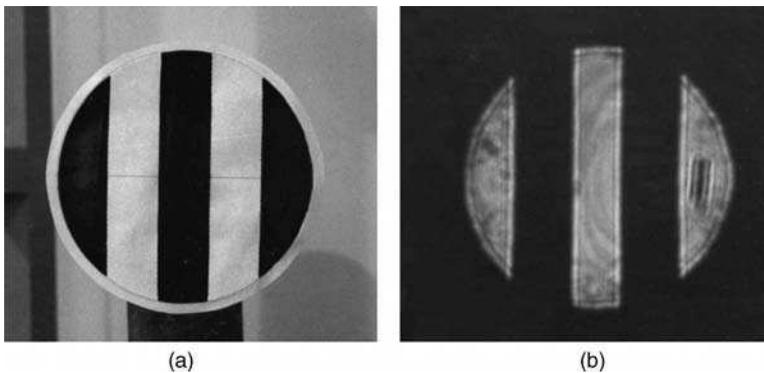


FIGURE 9.26. Ronchi Hartmann null test, (a) Ronchi-Hartmann screen in front of the surface under test, (b) Observed null fringe pattern. (From Cordero-Davila, et al., 1990.)

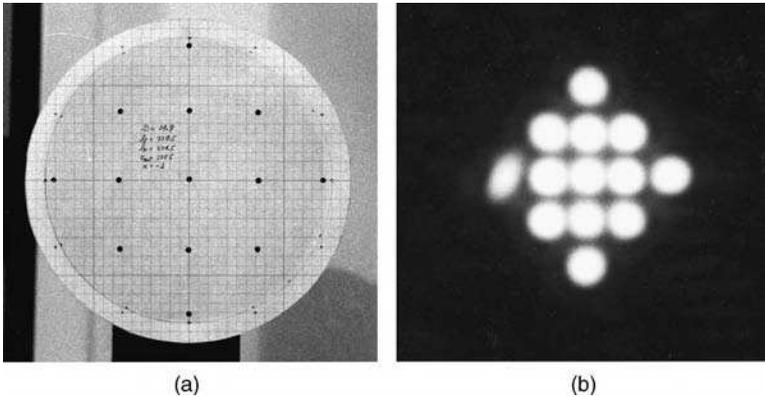


FIGURE 9.27. Hartmann null test. (a) Hartmann screen with holes. (b) Observed pattern, with spots along straight lines (null pattern). (from Cordero-Davila et al., 1990.)

planes. Using the mathematical development of Section 9.2.2, an ideal Ronchigram [Fig. 9.26(a)] was calculated and used as a one-dimensional Hartmann screen in front of the surface. The observed pattern is a null one because the new Ronchi–Hartmann screen compensates the asphericity of the surface under test. In Figure 9.26(b) the observed null pattern is shown.

Having the development of a null Hartmann test as a main aim, and after establishing the relation between the Ronchi and Hartmann test, the new Hartmann screen for a null Hartmann test was obtained by locating the holes on the screen, as the crossing points of two perpendicular ideal Ronchigrams. In Figures 9.27 (a,b) the screen and observed Hartmanngram are shown for the so-called Null Hartmann test. Recently, a common mathematical theory for the Ronchi and Hartmann test was developed by Cordero-Devila et al. (1992), which consider the results described in this section.

ACKNOWLEDGMENTS

I am indebted to Dr. Daniel Malacara for many stimulating and enlightening conversations, and to late Dr. Robert H. Noble for his valuable criticism. I am grateful to the Instituto Nacional de Astrofísica, Óptica, y Electrónica (INAOE), México, for the support given during the preparation of the three editions of the book, and in particular to its first General Director, late Dr. Guillermo Haro. Besides there, I am grateful to all the staff members of INAOE and the collegians of several Mexican institutions mentioned in the previous editions. In this edition I also want to thank F. Granados-Agustín, Gustavo Rodríguez-Zurita, Yaoltzin Luna, Noé Toto, and Ana Ma. Zarate for their help in several aspects.

REFERENCES

Adachi I., “Quantitative Measurement of Aberration by Ronchi Test,” *Atti. Fond. Giorgio Ronchi Contrib. Ist. Naz. Ottica*, **15**, 461–483 (1960a).

- Adachi I., "Quantitative Measurement of Aberration by Ronchi Test," *Atti. Fond. Giorgio Contrib. Ist. Naz. Ottica*, **15**, 550–585 (1960b).
- Adachi I., "The Recent History of the Grating Interferometer and Its Applications," *Atti. Fond. Giorgio Ronchi Contrib. Ist. Naz. Ottica*, **17**, 252–259 (1962a).
- Adachi I., "Measurement of Transfer Function of the Ronchi Test," *Atti. Fond. Giorgio Ronchi Contrib. Ist. Naz. Ottica*, **17**, 523–534 (1962b).
- Adachi I., "The Diffraction Theory of the Ronchi Test," *Atti. Fond. Giorgio Ronchi Contrib. Ist. Naz. Ottica*, **18**, 344–349 (1963).
- Agureev V. V., "Refractive-Index Distribution in Gradient-Index Fibers, Studied by Means of Shearing Interferometry," *J. Opt. Tech.*, **62**, 375–378 (1995).
- Agureev V. V., V. N. Morozov, G. O. Karapetyan and S. N. Ivashevskii, "Refractive-Index Distribution in Gradient-Index Fibers, Studied by Means of Shearing Interferometry," *J. Opt. Tech.*, **62**, 375–378 (1995).
- Anderson J. A. and R. W. Porter, "Ronchi's Method of Optical Testing," *Astrophys. J.*, **70**, 175–181 (1929).
- Assa A., A. A. Betseer, and J. Politch, "Recording Slope and Curvature Contours of Flexed Plates Using a Grating Shearing Interferometer," *Appl. Opt.*, **16**, 2504–2513 (1977).
- Barakat R., "General Diffraction Theory of Optical Aberration Tests from the Point of View of Spatial Filtering," *J. Opt. Soc. Am.*, **59**, 1432–1439 (1969).
- Bernardo L. M. and O. D. Soares, "Evaluation of the Focal Distance of a Lens by Talbot Interferometry," *Appl. Opt.*, **27**, 296–301 (1988).
- Blakley R., "Testing Concave Surfaces with a Rotating Ronchi Grating," *Opt. Eng.*, **33**, 3472–3475 (1994).
- Bocchino G., "L'Aberrazione Sferica Zonale Esaminata con i Reticoli a Bassa Frequenza," *Ottica*, **5**, 286 (1940).
- Bocchino G., "Un Metodo per la Determinazione Rapida e Precise dell' Aberrazione Sferica Semplice, Mediante la Frange d'Ombra," *Ottica*, **8**, 310 (1943).
- Braat J. and A. J. E. M. Janssen, "Improved Ronchi Test with Extended Source," *J. Opt. Soc. A*, **16**, 131–140 (1999).
- Briers J. D., "Interferometric Testing of Optical Systems and Components: A Review," *Opt. Laser Technol.*, **4**, 28–41 (1972).
- Briers J. D., "Ronchi Test Formulae 1: Theory," *Opt. Laser Technol.*, **11**, 189–196 (1979).
- Briers J. D. and M. J. Cochrane, "Ronchi Test Formulae 2: Practical Formulae and Experimental Verification," *Opt. Laser Technol.*, **11**, 245–257 (1979).
- Brookman E. C, L. D. Dickson, and R. S. Fortenberry, "Generalization of the Ronchi Ruling Method for Measuring Gaussian Beam Diameter," *Opt. Eng.*, **22**, 643–647 (1983).
- Bruscaglioni R., "Sulla Forma delle Fringe d'Interferenza Ottenute da Onde Affette da Astigmatismo Puro con Reticoli ad Orientamento Qualunque," *Rend. Accad. Naz. Lincei.*, **15**, 70 (1932a).
- Bruscaglioni R., "Sulla Misura dell' Astigmatismo e del Coma Mediante le Frange d'Ombra," *Boll. Assoc. Atti. Fond. Giorgio Ronchi Contrib. Ist. Naz. Ottica Ital.*, **6**, 46 (1932b).
- Bruscaglioni R., "Sulla Sensibilita della Rivelazione e Sulla Misura dell' Astigmatismo con Metodi Interferenziali," *Boll. Assoc. Atti. Fond. Giorgio Ronchi Contrib. Ist. Naz. Ottica Ital.*, **7**, 78 (1933).

- Bruscaglioni R., "Controllo della Afocalità di una Parte e Controllo di un Piano Campione con l'Interferometro Ronchi a Reticolo," *Ottica*, **4**, 203 (1939).
- Calamai G., "Su di una Formula per la Misura dell'Astigmatismo Mediante i Reticoli," *Ottica*, **3**, 41 (1938).
- Chen L. C., J. Chen, R. Zhu, Q. Wang, H. Liu, Y. Zhang, "Ronchi Test Using Overlapping Averaging 4-Frame Algorithm," *Proc. SPIE*, **2003**, 398–404 (1993).
- Cordero-Davila A., A. Cornejo Rodriguez, and O. Cardona-Nuñez, "Null Hartmann and Ronchi Hartmann Tests," *Appl. Opt.*, **29**, 4618 (1990).
- Cordero-Davila A., A. Cornejo-Rodriguez, O. Cardona-Nuñez, "Common Mathematical Theory for the Ronchi and Hartmann Test," *Appl. Opt.*, **31**, 2370–2376 (1992).
- Cordero-Dávila A., J. Díaz-Anzures, V. Cabrera-Peláez, "Algorithm for the Simulation of Ronchigrams of Arbitrary Optical Systems and Ronchi Grids in Generalized Coordinates," *Appl. Opt.*, **41**, 3866–3873 (2002).
- Cornejo-Rodriguez A., "The Ronchi Test for Aspherical Optical Surfaces," *Kogaku*, **12**, 278 (1983).
- Cornejo A., and D. Malacara, "Ronchi Test of Aspherical Surfaces: Analysis and Accuracy," *Appl. Opt.*, **9**, 1897–1901 (1970).
- Cornejo-Rodriguez A. and D. Malacara, "Wavefront Determination Using Ronchi and Hartmann Tests," *Bol. Int. Tonantzintla*, **2**, 127–129 (1976).
- Cornejo-Rodriguez A., H. Altamirano, and M. V. R. K. Murty, "Experimental Results in the Sharpening of the Fringes in the Ronchi Test," *Bol. Inst. Tonantzintla*, **2**, 313–315 (1978).
- Crino B., "Sulla Misura dell'Aberrazione Sferica, Coma e Astigmatismo Mediante la Frange d'Ombra Estrassiale Ottenute con Reticoli Rettilinei," *Boll. Assoc. Ottica Ital.*, **7**, 113 (1933).
- Crino B., "Nuovi Risultati Nello Studio Analitico delle Frange d'Ombra Ottenute per Interferenza di Onde Aberranti," d'Ombra Ottenute per Interferenza di Onde Aberranti," *Ottica*, **4**, 114 (1939).
- De Vany A. S., "Some Aspects of Interferometric Testing and Optical Figuring," *Appl. Opt.*, **4**, 831–833 (1965).
- De Vany A. S., "Quasi-Ronchigrams as Mirror Transitive Images of Interferograms," *Appl. Opt.*, **9**, 1844–1945 (1970).
- De Vany A. S., "Laser Illuminated Divergent Ball Bearing Sources," *Appl. Opt.*, **13**, 457–459 (1974).
- De Vany A. S., "Profiling pitch polishers" *Appl. Opt.*, **17**, 3022–3024 (1978).
- De Vany A. S., "Interpreting Wave-Front and Glass-Error Slopes in an Interferogram," *Appl. Opt.*, **19**, 173 (1980).
- De Vany A. S., "Patterns of Correlation of Interferograms and Ronchigrams," *Appl. Opt.*, **20**, A40–A41 (1981).
- Di Jorio M., "Ulteriore Approssimazione dello Studio delle Aberrazioni con l'Interferometro Ronchi a Reticolo," *Ottica*, **4**, 31 (1939a).
- Di Jorio M., "Una Formula Più Precisa delle Frange d'Ombra dell'Interferometro Ronchi a Reticolo," *Ottica*, **4**, 83 (1939b).
- Di Jorio M., "L'Aberrazione Sferica Esaminata con i Reticoli di Alta Frequenza," *Ottica*, **4**, 184 (1939c).
- Di Jorio M., "Estensione del Concetto del Fuoco Uniforme: il Fuoco Uniforme Zonale," *Ottica*, **4**, 254 (1939d).
- Di Jorio M., "Similitudine Degli Interferogrammi dell'Interferometro Ronchi al Variare x_p e Verifica Sperimentale della Costanza del numero b ," *Ottica*, **7**, 243 (1942).

- Di Jorio M., "Equazione dell'Interferometro Ronchi per le Onde Sferiche Aperte Fino al Quarto Ordine, e sua Discussione," *Ottica*, **8**, 288 (1943).
- Díaz-Uribe R., "Medium Precision Null Screen Testing of Off-Axis Parabolic Mirrors," *Appl. Opt.*, **39**, 2790–2804 (2000).
- Díaz-Uribe R., and M. Campos, "Null-Screen Testing of Fast Convex Aspheric Surfaces," *Appl. Opt.*, **39**, 2670–2677 (2000).
- Elster C., "Exact Two-Dimensional Wavefront Reconstruction from Lateral Shearing Interferograms with Large Shears," *Appl. Opt.*, **39**, 5353–5359 (2000).
- Elster C., and I. Weingartner, "Solution to the Shearing Problem," *Appl. Opt.*, **38**, 5024–5031 (1999).
- Erdős, P., "Ronchi Test of Fifth Order Aberrations," *J. Opt. Soc. Am.*, **49**, 865–868 (1959).
- Fischer D. J., H. P. Stahl, "A Vector Formulation for Ronchi Shear Surface Fitting," *Proc. SPIE*, **1755**, 228–238 (1994).
- Glatt I. and O. Kafri, "Determination of the Focal Length of Nonparaxial Lenses by Moiré Deflectometry," *Appl. Opt.*, **26**, 2507–2508 (1987).
- Hamsher, D. H., "Screen Line Tests of Paraboloidal Reflectors," *J. Opt. Soc. Am.*, **36**, 291–295 (1946).
- Harbers G., P. J. Kunst, and G. W. Leibbrandt, "Analysis of Lateral Shearing Interferograms by Use of Zernike Polynomials," *Appl. Opt.*, **35**, 6162–6172 (1996).
- Hariharan P., W. H. Steel, and J. C. Wyant, "Doubling Grating Interferometer with Variable Lateral Shear," *Opt. Comm.*, **11**, 317 (1974).
- Hegeman P., X. Christmann, M. Visser, J. Braat, "Experimental Study of a Shearing Interferometer Concept for At-Length Wave Characterization of Extreme-Ultraviolet Optics," *Appl. Opt.*, **40**, 4526–4533 (2001).
- Hopkins G. H. and R. H. Shagan, "Null Ronchi Gratings from Spot Diagram," *Appl. Opt.*, **16**, 2602–2603 (1977).
- Jentzsch F., "Die Rastermethode: Ein Verfahren zur Demonstration und Messung der Sphärischen Aberration," *Phys. Z.*, **24**, 66 (1928).
- Kamalov J. A., V. A. Komissaruk, and N. P. Mende, "Preparation of the IAB-451 Instrument for Operation as a Diffraction Interferometer," *Sov. J. Opt. Technol.*, **41**, 249–250 (1980).
- Kasana R. S., S. Boseck, and K. J. Rosenbrich, "Non-Destructive Collimation Technique for Measuring Glass Constants Using a Ronchi Grating Shearing Interferometer," *Opt. Laser Technol.*, **16**, 101–105 (1984).
- Keren E., K. M. Kreske, and O. Kafri, "Universal Method for Determining the Focal Length of Optical Systems by Moiré Deflectometry," *Appl. Opt.*, **27**, 1383–1385 (1988).
- Kessler D. and R. V. Shack, "Dynamic Optical Tests of a High-Speed Polygon," *Appl. Opt.*, **20**, 1015–1019 (1981).
- King J. H., "Quantitative Optical Test for Telescope Mirrors and Lenses," *J. Opt. Soc. Am.*, **24**, 250 (1934). Reprinted in: *Amateur Telescope Making*, vol. 2, A. G. Ingalls, Ed., Scientific American, New York, 1953, p. 104.
- Kirkham A. R., "The Ronchi Test for Mirrors," in: *Amateur Telescope Making*, vol. 1, A. G. Ingalls, Ed., Scientific American, New York, 1953, p. 264.
- Koliopoulos C. L., "Radial Grating Lateral Shear Heterodyne Interferometer," *Appl. Opt.*, **19**, 1523–1528 (1980).

- Kosłowska A. and M. Kujawinska, "Grating Interferometry with a Semiconductor Source," *Appl. Opt.*, **36**, 8116–8120 (1997).
- Kuindzhi V. V., S. A. Strezhnev and M. T. Popov, "Gratings for Diffraction Interferometer," *Sov. J. Opt. Technol.*, **47**, 240–241 (1980).
- Lau E., "Bengungserscheinungen an Deppeltrastern," *Ann. Phys.*, **6**, 417 (1948).
- Lenouvel M. L., "Methode de Determination et de Mesure des Aberrations des Systemes Optiques," *Rev. Opt.*, **3**, 211 (1924a).
- Lenouvel M. L., "Methode de Determination et de Mesure des Aberrations des Systemes Optiques," *Rev. Opt.*, **3**, 315 (1924b).
- Lenouvel M. L., "Etude des Objectifs de Reproduction," *Rev. Opt.*, **4**, 294 (1925a).
- Lenouvel M. L., "Essai d'Objectifs par le Coin d'Air" *Rev. Opt.*, **4**, 299 (1925b).
- Lin J. A., J. Hsu, and S. A. Shire, "Quantitative Three Beam Ronchi Test," *Appl. Opt.*, **29**, 1912 (1990).
- Lohmann A. W. and D. E. Silva, "An Interferometer Based on the Talbot Effect," *Opt. Comm.*, **2**, 413 (1971).
- Lohmann A. W. and D. E. Silva, "A Talbot Interferometer with Circular Gratings," *Opt. Comm.*, **4**, 926 (1972).
- Lower H. A., "Notes on the Construction of an F/1 Schmidt Camera," in *Amateur Telescope Making*, Vol. 2, A. G. Ingalls, Ed., Scientific American, New York, 1954, p. 410; formerly published by Munn and Co., 1937.
- Lumley E., "A Method of Making a Ronchi Test on an Aspheric Mirror," in *Amateur Astronomers*, Sidney, 1959; reprinted in *Atti. Fond. Giorgio Ronchi Contrib. Ist. Naz. Ottica*, **15**, 457 (1960).
- Lumley E., "Figuring a Paraboloid with the Ronchi Test," *Sky Telesc.*, **22** 298 (1961).
- Luna-Zayas Y., F. Granados-Agustín, and A. Cornejo-Rodríguez, "Ronchi Test with Sub-Structured Gratings," *Proc. SPIE*, **6034**, 480–486 (2006).
- Malacara D., Testing of Optical Surfaces, Ph.D. Thesis, University of Rochester, New York, University Microfilms, Ann Arbor, Mich., Order No. 65–12,013, 1965a.
- Malacara D., "Geometrical Ronchi Test of Aspherical Mirrors," *Appl. Opt.*, **4**, 1371–1374 (1965b).
- Malacara D., "Ronchi Test and Transversal Spherical Aberration," *Bot Obs. Tonanzintla Tacubaya.*, **4**, 73 (1965c).
- Malacara D., "Analysis of the Interferometric Ronchi Test," *Appl. Opt.*, **29**, 3633–3637 (1990).
- Malacara D. and A. Cornejo, "Modified Ronchi Test to Measure the Axial Chromatic Aberration in Lenses," *Appl. Opt.*, **10**, 679–680 (1971).
- Malacara D. and A. Cornejo, "Relating the Ronchi and Lateral Shearing Interferometer Tests," *Opt. Spectra.*, **8**, 54–55 (1974a).
- Malacara D. and A. Cornejo, "The Talbot Effect in the Ronchi Test," *Bol. Inst. Tonanzintla*, **1**, 193–196 (1974b).
- Malacara D. and A. Cornejo, "Null Ronchi Test for Aspherical Surfaces," *Appl. Opt.*, **13**, 1778–1780 (1974c).
- Malacara D. and A. Cornejo, "Shape Measurement of Optical Systems with Rotational Symmetry Using Ronchigrams," *Bol. Inst. Tonanzintla*, **1**, 277–283 (1975).
- Malacara D. and A. Cornejo, "Side Band Ronchi Test," *Appl. Opt.*, **15**, 2220–2222 (1976a).
- Malacara D. and A. Cornejo, "Third Order Computation of Null Ronchi Rulings," *Bol. Inst. Tonanzintla*, **2**, 91–92, (1976b).

- Malacara D. and M. Josse, "Testing of Aspherical Lenses Using Side Band Ronchi Test," *Appl. Opt.*, **17**, 17–18 (1978).
- Mallick S. and M. L. Roblin, "Coherent Imaging in Presence of Defect of Focus," *Trans. Opt. Soc. Am.*, **64**, 1944 (1972).
- Martin L. C., *Technical Optics*, vol. 2, Pitman, London, 1960, Chap. VII.
- Mobsby E., "Testing Parabolic Mirrors with Inverse Parabolic Grating," *Astronomy: J. Wessex Astron. Soc.*, **1**, 13 (1973).
- Mobsby E., "A Ronchi Null Test for Paraboloids," *Sky Telesc.*, **48**, 325–330 (1974).
- Morais C., "Riassunto delle Applicazioni dei Reticoli allo Studio delle Aberrazione dei Sistemi Ottici," *Atti. Fond. Giorgio Ronchi Contrib. Ist. Naz. Ottica*, **13**, 546–602 (1958).
- Murty M. V. R. K., "Interferometry Applied to Testing of Optics," *Bull. Opt. Soc. India*, **1**, 29–37 (1967).
- Murty M. V. R. K., "A Simple Method of Introducing Tilt in the Ronchi and Cube Type of Shearing Interferometers," *Bull. Opt. Soc. India*, **5**, 1–5 (1971).
- Murty M. V. R. K. and A. Cornejo, "Sharpening the Fringes in the Ronchi Test," *Appl. Opt.*, **12**, 2230–2231 (1973).
- Murty M. V. R. K. and A. H. Shoemaker, "Theory of Concentric Circular Grid," *Appl. Opt.*, **5**, 323–326 (1966).
- Nakano Y. and K. Murata, "Talbot Interferometry for Measuring the Focal Length of a Lens," *Appl. Opt.*, **24**, 3162–3166 (1985).
- Núñez A. J. M., A. Cordero Dávila, S. Vergara-Limón, and J. Cuautle-Cortés., "Improved Ronchi Tester," *Appl. Opt.*, **40**, 501–505 (2001).
- Omura K. and T. Yatagai, "Phase Measuring Ronchi Test," *Appl. Opt.*, **27**, 523–528 (1988).
- Pacella G. B., "Sulla Ricerca della Forma delle Onde Luminose dall'Esame delle Frange d'Ombra," *Rend. Accad. Naz. Lincei.*, **5**, 752 (1927).
- Páez G., M. Strojnik, and G. Ga, "Vectorial Shearing Interferometer," *Appl. Opt.*, **39**, 5172–5178 (2000).
- Pallotino P., "Sulla Dissimmetria delle Frange Dell'Interferometro Ronchi a Reticolo," *Ottica*, **6**, 26 (1941).
- Pastor J., "Hologram Interferometry and Optical Technology," *Appl. Opt.*, **8**, 525–531 (1969).
- Patorski K., "Measurements of the Wavefront Curvature of Small Diameter Laser Beams Using the Fourier Imaging Phenomenon," *Opt. Laser Technol.*, **11**, 91 (1979).
- Pastorski K., "Production of Binary Amplitude Gratings with Arbitrary Opening," *Opt. Laser Technol.*, **12**, 267–270 (1980).
- Patorski K., "Periodic Source Ronchi-Talbot Shearing Interferometer," *Optik*, **62**, 207–210 (1982).
- Patorski K., "Heuristic Explanation of Grating Shearing Interferometry Using Incoherent Illumination," *Opt. Acta*, **31**, 33–38 (1984).
- Patorski K., "Grating Shearing Interferometer with Variable Shear and Fringe Orientation," *Appl. Opt.*, **25**, 4192–4198 (1986).
- Patorski K. and A. Cornejo-Rodriguez, "Fringe Contrast Interpretation for an Extended Source Ronchi Test," *Appl. Opt.*, **25**, 2790–2795 (1986a).
- Patorski K. and A. Cornejo-Rodriguez, "Ronchi Test with Daylight Illumination," *Appl. Opt.*, **25**, 2031–2032 (1986b).

- Patorski, K. and L. Salbut, "Reversed Path Ronchi Test," *Optica Appl.*, **15**, 261 (1984).
- Phillips F. W., "Aspherizing and Other Problems in Making Maksutov Telescopes," *Sky Telesc.*, **25**, 110 (1963).
- Popov G. M., "Methods of Calculation and Testing of Ritzley-Chrétien Systems," *Izv. Krym. Astrofiz. Obs.*, **45**, 188 (1972).
- Porter R. W., "Notes on the Ronchi Band Patterns," in: *Amateur Telescope Making*, vol. 1, A. G. Ingalls, Ed., Scientific American, New York 1953, p. 268.
- Rank D. H., P. R. Yoder Jr., and J. Vrabel, "Sensitivity of a Rapid Test for High Speed Parabolic Mirrors," *J. Opt. Soc. Am.*, **39**, 36–38 (1949).
- Rayces J. L., "Exact Relation Between Wave Aberration and Ray Abberation," *Opt. Acta*, **11**, 85–88 (1964).
- Ricci E., "Nuovi Criteri per la Misura dell'Aberrazione Sferica di un Sistema Ottico; col Metodo delle Frange d'Ombra," *Ottica*, **4**, 104 (1939).
- Ronchi V., "Le Frange di Combinazioni Nello Studio delle Superficie e dei Sistemi Ottici," *Riv. Ottica Mecc. Precis.*, **2**, 9 (1923a).
- Ronchi V., "Due Nuovi Metodi per lo Studio delle Superficie e dei Sistemi Ottici," *Ann. Sc. Norm. Super Pisa.*, **15**, (1923b).
- Ronchi V., "Sopra la Caratteristiche dei Cannocchiali di Galileo e la Loro Autenticita," *Rend. Accad. Naz. Lincei.*, **2**, 162 (1923c).
- Ronchi V., "Sopra i Cannocchiali di Galileo e Sopra una Lente di Evangelista Torricelli," *L'Universo*, **4**, 10 (1923d).
- Ronchi V., "Sullo Studio dei Sistemi Ottici col Biprisma e Gli Specchi del," *Ren. Accad. Naz. Lincei.*, **3**, 314 (1924a).
- Ronchi V., "Ancora sull' Impiego dei Reticoli Nelo Studio dei Sistemi Ottici," *Nuovo Cimento*, **1**, 209 (1924b).
- Ronchi V., *La Prova dei Sistemi ottici*, Zanichelli, Bologna, 1925.
- Ronchi V., "Sur la Nature Interferentielle des Franges d'Ombre dans l'Essai des Systemes Optiques," *Rev. Opt.*, **5**, 441 (1926a).
- Ronchi V., "Uber die Schattenstreifen zum Studium der Lichtwellen," *Zt. Instrumentenknd*, **46**, 553 (1926b).
- Ronchi V., "Sul Comportamento e l'Impiengo delle Frange d'Ombra Nella Prova dei Systemes Optiques," *Nuovo Cimento*, **4**, 297 (1927).
- Ronchi V., "Sul Comportamento e l'Impiengo delle Frange d'Ombra nella Prova dei Systemes Optiques," *Rev. Opt.*, **7**, 49 (1928).
- Ronchi V., "Le Frange d'Ombra Nello Studio delle Aberrazioni Sferiche," *Rend. Accad. Naz. Lincei.*, **11**, 998 (1930).
- Ronchi V., "Sulla Sensibilita delle Frange d'Ombra all Aberrazione Sferica," *Ottica*, **5**, 275 (1940).
- Ronchi V., *Corso di Ottica Tecnica*, 2nd ed. Associazione Ottica Italiana, Firenze, 1954.
- Ronchi V., "An Elementary Introduction to the Use of the Grating Interferometer," *Atti. Fond. Giorgio Ronchi Contrib. Ist. Naz. Ottica*, **13**, 368–403 (1958).
- Ronchi V., "Forty Years of Gratings," *Atti. Fond. Giorgio Ronchi Contrib. Ist. Naz. Ottica*, **17**, 93–143 (1962a).
- Ronchi V., "Forty Years of Gratings," *Atti. Fond. Giorgio Ronchi Contrib. Ist. Naz. Ottica*, **17**, 240–251 (1962b).

- Ronchi V., "Forty Years of History of a Grating Interferometer," *Appl. Opt.*, **3**, 437–451 (1964).
- Ronchi V., "On the Phase Grating Interferometer," *Appl. Opt.*, **4**, 1041–1042 (1965).
- Rosenbruch K. J., "Testing of Optical Components and Systems," *J. Optics (India)*, **14**, 23–43 (1985).
- Royo S., J. Arasa, and C. Pizarro, "Profilometry of Toroidal Surfaces with an Improved Ronchi Test," *Appl. Opt.*, **39**, 5721–5731 (2000).
- Salzmann H., "A Simple Interferometer Based on the Ronchi Test," *Appl. Opt.*, **9**, 1943–1944 (1970).
- Scandone F., "Sulla Forma delle Frange d'Ombra Dovute ad Onie Luminose Affette da Aberrazione Zonale," *Nuovo Cimento*, **7**, 289 (1930).
- Scandone F., "Sulla Forma delle Frange d'Ombra Dovute ad Onlde Luminose Affette de Aberrazione Estrassiale," *Nuovo Cimento*, **8**, 157 (1931a).
- Scandone F., "Sulla Forma delle Frange d'Ombra Estrassiale Ottenute con un Reticolo a Tratti Inclinati sul Piano de Simmetria del Sistema Ottico," *Nuovo Cimento*, **8**, 310 (1931b).
- Scandone F., "Sulla Forma delle Frange d'Ombra Ottenute con Reticolo Circolari a Frequenza Costante," *Nuovo Cimento*, **8**, 378 (1931c).
- Scandone F., "Sulla Forma delle Frange d'Ombra Ottenute con Reticolo Circolari a Frequenza Costante non Centrato sull'Asse Ottico," *Boll. Assoc. Ottica Ital.*, **6**, 35 (1932).
- Scandone F., "Sulla Frange d'Ombra Estrassiali Ottenute con Reticoli in Presenza di Aberrazione Sferica sull'Asse," *Boll. Assoc. Ottica Ital.*, **7**, 100 (1933).
- Schreiber H. and J. Schwider, "Lateral Shearing Interferometer Based in Two Ronchi Phase Gratings in Series," *Appl. Opt.*, **36**, 5321–5324 (1997).
- Schulz L. G., "Über die Prufung Optischer Systeme mit Rastein," *Ann. Phys.*, **35**, 189 (1928).
- Schulz L. G., "Preparation of Aspherical Refracting Optical Surfaces by an Evaporation Technique," *J. Opt. Soc. Am.*, **38**, 432–441 (1948).
- Schwider J., "Single Sideband Ronchi Test," *Appl. Opt.*, **20**, 2635–2642 (1981).
- Seiichi O. and T. Nomura, "High-Precision Analysis of a Lateral Shearing Interferogram by Use of the Integration Method and Polynomials," *Appl. Opt.*, **39**, 5179–5186 (2000).
- Servin M., D. Malacara, and F. J. Cuevas, "Direct-Phase Detection of Modulated Ronchi Rulings Using a Phase-Locked Loop," *Opt. Eng.*, **33**, 1193–1199 (1994).
- Sherwood A. A., "A Quantitative Analysis of the Ronchi' Test in Term of Ray Optics," *J. Br. Astron. Assoc.*, **68**, 180 (1958).
- Sherwood A. A., "Ronchi Test Charts for Parabolic Mirrors," *J. Proc. R. Soc. New South Wales*, **43**, 19 (1959); reprinted in *Atti. Fond. Giorgio Ronchi Contrib. Ist. Naz. Ottica*, **15**, 340–346 (1960).
- Silva D. E., "Talbot Interferometer for Radial and Lateral Derivatives," *Appl. Opt.*, **11**, 2613–2624 (1972).
- Sirohi R. S., T. Eiju, and K. Matsuda, "Multiple Beam Lateral Shear Interferometer for Optical Testing," *Appl. Opt.*, **34**, 2864–2870 (1995).
- Steig H., "More on Ronchi Gratings," in: *Advanced Telescope Techniques, vol. 1: Optics*, A. Mackintosh, Ed., Willmann-Bell, Inc., Richmond VA, 1987, p. 46.
- Stoltzmann D. E., "Application of Ronchi Interferometry to Tssting Large Aperture Flat Mirrors," *Proc. SPIE*, **153**, 68–75 (1978).
- Strong J., *Procedure in Experimental Physics*, Prentice-Hall, Englewood Cliffs, NJ, 1983, Chap. 11.

- Talbot F., "Facts Relating to Optical Science No. IV," *Philos Mag.*, **9**, 401 (1836).
- Thompson B. J., Studies in Optics, Technical Report, U.S. AFAL-TR-73-112, U.S. Government, Washington, D.C. 1973.
- Thompson E. T., "Making Ronchi Gratings," in *Advanced Telescope Techniques, vol. 1: Optics*, A. Mackintosh, Ed., Willmann-Bell, Inc., Richmond, VA, 1987, p. 45.
- Toraldo di Francia G., "Sulla Frange d'Interferenza delle Onde Aberranti," *Ottica*, **6**, 151 (1941a).
- Toraldo di Francia G., "Saggio di una Teoria Generate dei Reticoli," *Ottica*, **6**, 258 (1941b).
- Toraldo di Francia G., "Limiti di Validita dell'Ipotesi della Rotazione Rigida per il Reticolo Rettilineo," *Ottica*, **7**, 282 (1942a).
- Toraldo di Francia G., "La Prova dell'Aberrazione Cromatica con l'Interferometro a Reticolo," *Ottica*, **7**, 302 (1942b).
- Toraldo di Francia G., "Ancora su le Aberrazioni delle Onde Diffratte dial Reticolo Rettilineo," *Ottica*, **8**, 1 (1943a).
- Toraldo di Francia G., "La Formula Esatta per le Frange del' Interferometro Ronchi," *Ottica*, **8**, 225 (1943b).
- Toraldo di Francia G., "Introduzione alla Teoria Geometrica e Interferenziale delle Onde Aberranti," *Atti. Fond. Giorgio Ronchi Contrib. Ist. Naz. Ottica*, **1**, 122-138 (1946).
- Toraldo di Francia G., "Introduzione alla Teoria Geometrica e Interferenziale delle Onde Aberranti," *Atti. Fond. Giorgio Ronchi Contrib. Ist. Naz. Ottica*, **2**, 25-42 (1947).
- Toraldo di Francia G., "Geometrical and Interferential Aspects of the Ronchi Test," in: *Optical Image Evaluation*, Nat. Bur. Stand. U.S. Circ. No. 256, U.S. Government Printing Office, Washington, D.C., 1954, Chap. 11, p. 161.
- Van Brog. H., "Zernike Polynomials as a Basis for Wavefront Fitting in Lateral Shearing Interferometry," *Appl. Opt.*, **36**, 2778-2790 (1987).
- Van Brug H., "Zernike Polynomials as a Basic for Wavefront Fitting in Lateral Shearing Interferometry," *Appl. Opt.*, **36**, 2788-2790 (1997).
- Villani F., "Sulla Misurazione dell' Astigmatismo e del Coma Mediarue le Frange d'Ombra," *Nuovo Cimento*, **7**, 248 (1930).
- Villani F. and R. Bruscazioni, "Sulla Forma delle Frange d'Ombra ottenute de Onde Affette da Astigmatismo e Coma," *Nuovo Cimento*, **9**, 1 (1932).
- Vogl G., "A Phase Grating Interferometer," *Appl. Opt.*, **3**, 1089 (1964).
- Waland R. L., "Note on Figuring Schmidt Correcting Lenses," *J. Sci. Instrum.*, **15**, 339 (1938).
- Wan D. S. and D. T. Lin, "Ronchi Test and a New Phase Reduction Algorithm," *Appl. Opt.*, **29**, 3255-3265 (1990).
- Wan D. S. and D. T. Lin, "Ronchi Test and a New Phase Reduction Algorithm," *Appl. Opt.*, **29**, 3555-3565 (1990).
- Wan D.-S. and M.-W. Chang, "Effect of Grating Spacing on the Ronchi Test," *Opt. Eng.*, **32**, 1084-1090 (1993).
- Wehn R., "Die Methode der Ronchi-Gitter in der Praxis," *Atti. Fond. Giorgio Ronchi Contrib. Ist. Naz. Ottica*, **17**, 39-96 (1962).
- Xu Y., O. Sasaki, and T. Susuki, "Double Grating Interferometer for Measurement of Cylinder Diameters," *Appl. Opt.*, **43**, 537-541 (2004).

- Yatagai T., "Fringe Scanning Ronchi Test for Aspherical Surfaces," *Appl. Opt.*, **23**, 3676–3679 (1984).
- Yokoseki S. and T. Susuki, "Shearing Interferometer Using the Grating as the Beam Splitter," *Appl. Opt.*, **10**, 1575–1580 (1971a).
- Yokoseki S. and T. Susuki, "Shearing Interferometer Using the Grating as the Beam Splitter. Part 2," *Appl. Opt.*, **10**, 1690–1693 (1971b).
- Yoder P. R., Jr., "Further Analysis of the 'Lower' Test for High-Speed Parabolic Mirrors," *J. Opt. Soc. Am.*, **49**, 439–440 (1959).

10

Hartmann, Hartmann–Shack, and Other Screen Tests

D. Malacara-Doblado and I. Ghozeil

10.1. INTRODUCTION

This chapter has been rewritten and updated by the first author; it includes some of the materials taken from the chapter in previous versions of the book. The Hartmann test, as the Ronchi test and lateral shearing interferometry, measures the wavefront slope instead of the optical path difference (OPD), like the Twyman–Green and Fizeau interferometers. A very large wavefront deformation may produce small slope changes if the extension of this deformation is large. In an analogous manner, small wavefront deformations may produce large slope changes if their extension is small. However, since most of the times the final aim is the retrieval of the wavefront shape, this has to be obtained from these slope measurements. These slope are measured from the transverse aberrations at some observation plane near the focal plane using a geometric optics approach.

This chapter is concerned mainly with the methods of sampling a wavefront or mirror surface through the use of such screens and similar methods. The sections on applications of these methods and comparison of the various screen tests among themselves are included in this chapter. Although the methods described are applicable to most lens systems, the presentation in the rest of this chapter will be made for large concave mirrors, which are the most commonly tested.

The Hartmann test, invented by Hartmann (1900, 1904a,b and c) to test the Great Refractor at Postdam and illustrated in Figure 10.1, has its antecedent in the measurements of eye refractive defects using a screen with two holes in front of the eye as described by Tscherning (1894). It uses a screen with an array of holes placed close to the entrance or exit pupil of the system under test. The most frequent screen has a rectangular array of holes with one at the center as shown in Figure 10.2. As illustrated in Figure 10.3, the wavefront deformations $W(x, y)$ are

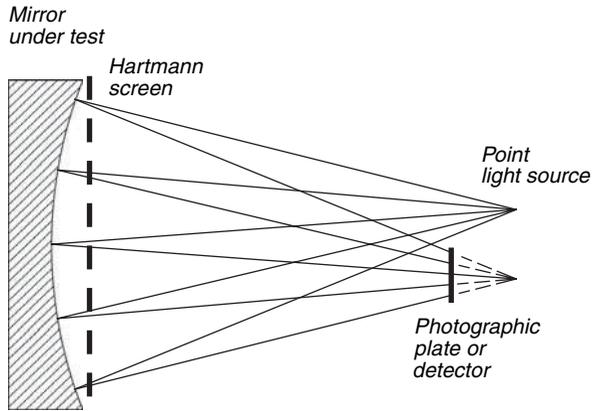


FIGURE 10.1. Hartmann test schematics of a concave mirror.

calculated from the measurements of the transverse aberrations TA_x and TA_y related by the expressions:

$$\frac{\partial W(x, y)}{\partial x} = -\frac{TA_x(x, y)}{r} \quad (10.1)$$

and

$$\frac{\partial W(x, y)}{\partial y} = -\frac{TA_y(x, y)}{r} \quad (10.2)$$

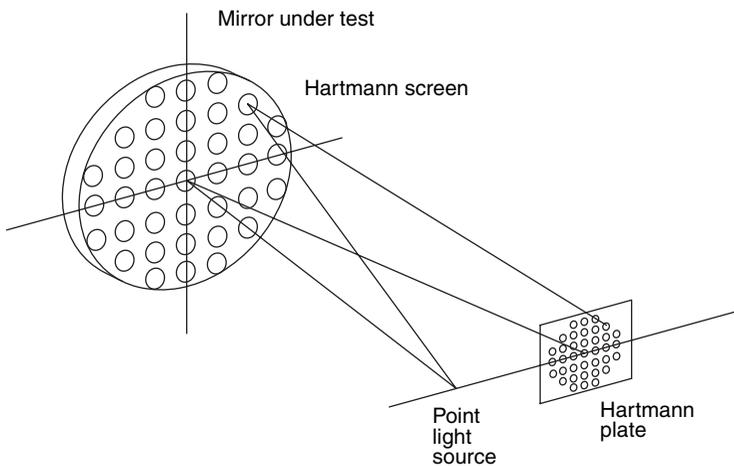


FIGURE 10.2. Hartmann test perspective schematics showing the Hartmann screen over a mirror to be tested.

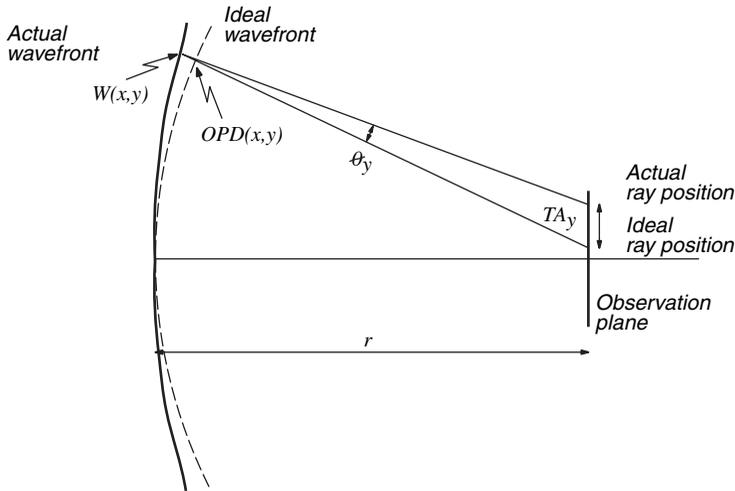


FIGURE 10.3. Relation between the transverse aberrations and the wavefront deformations.

where r is the distance from the pupil of the wavefront, which is to be evaluated, to the Hartmann plate. If the wavefront is convergent, as shown in Figure 10.1, and the Hartmann plate is close to the point of convergence, r is the radius of curvature of the wavefront.

Hartmann patterns are nothing else than the spot diagrams obtained in lens design evaluation, as pointed out by Stavroudis and Sutton (1965). When measuring a wavefront, the errors can be reduced if the test procedure used affords ways of obtaining a given result by several independent means. A test pattern that allows for several independent ways of obtaining a surface deviation at a sample point is likely to give a closer approximation to the actual value of that deviation than a pattern that lacks such a feature.

10.2. SOME PRACTICAL ASPECTS

The screen must be centered accurately on the aperture of the mirror under test. This is particularly true in the presence of spherical aberration, which is the case for aspheric mirrors tested at their center of curvature. A decentering of the screen in such a situation leads to an apparent presence of coma as described by Landgrave and Moya (1986). They present results for parabolic mirrors of various F numbers.

After the screen is placed and centered on the mirror, but before any plates are taken, the point light source used to illuminate the mirror through the screen must be centered properly relative to the mirror to prevent the introduction of off-axis aberrations. This can be done conveniently if the housing of the point source is of such a size that it is narrowly missed by the beams returned by the region near the center of the mirror. Without elaborate equipment, one can then center the source by judging and equalizing visually the gaps between the housing and the light beams.

It is important that a well thought-out support system be used with a mirror under test: Only deformations that will be present in actual use of the mirror are allowable. Some of the many effects of support systems have been described by Malvick (1972).

For convenience, the point source is placed at a position closer to the mirror than the center of curvature. This position places the conjugate image beyond the center of curvature and permits the interception of the converging light inside and outside the focus modes without the obstruction of the illuminating beam; the so-called focus referred to is the conjugate image location of the point source. In general, the image is highly aberrated for mirrors other than a sphere, and for classical telescope primaries, spherical aberration should be the only one present. To facilitate the identification of the dark spots on the Hartmann plate, it is necessary that the photographic plate or detector be placed outside the caustic limits. The source can be any light that is dim, or dimmed, enough to permit long exposure times, which do not result in overexposure of the photographic plate or detector. Exposure times in the range 90–180 s, depending on source brightness, should be sufficient to average random air turbulence.

The photographic plate or detector should be very close to being perpendicular to the optical axis; otherwise, a fictitious astigmatism will be obtained from the data reduction. There are several ways of orienting the plate holder correctly. A straightforward approach is to load the holder with a clear glass plate that has a cross hair scribed at its center. Then, with an alignment telescope, a zone on the mirror is viewed through the center of the plate. The alignment scope is first brought into alignment with the mirror, and then the fainter return from the glass plate is centered by tilting its holder.

With the screen and source properly located, one or more photographic plates are recorded, usually at a position between the source and its conjugate image. The location of the dots obtained on the photographic plate can then be determined to a high degree of accuracy with a micro-densitometer having an x-y traveling stage. Typically, a position of uncertainty of less than 0.003 mm can be expected. This error represents a surface tilt uncertainty of 0.015 s of arc (1.41×10^{-7} rad) for a mirror having a radius of curvature of 21.3 m as in the case of the 4-m primaries. The measuring stage is one of the few critical aspects of the test. When the photographic plate is placed in the micro-densitometer, this instrument is made to have its two perpendicular cross hairs present on a viewing screen through which the plate can also be seen. These cross hairs are aligned with the two rows of dots that correspond to the two perpendicular rows of holes in the square array, and which intersect at the center of the screen. The rows of dots do not form straight lines unless a very good mirror is under test. The plate and cross hairs are aligned through a succession of small plate rotations and displacements until some minimal departure is obtained. The task is easy with a mirror nearing completion because the dots line up nicely.

In general, if the pattern is physically large, say of the order of 60 mm, alignment can be made with a possible rotation error of less than 30 s of arc for the plate relative to the cross hairs. This error has little effect on the test results, as can be shown through repeated measurements of the same plate with slightly differing alignment

criteria. If a gross error of the order $0.5^\circ-1'$ is introduced, the resultant data reduction gives a somewhat astigmatic looking, saddle-shaped aspect for the surface figure.

An important assumption in screen testing is that between samples the wavefront changes are gradual rather than abrupt. This is a safe assumption because abrupt changes can be readily detected by other means, such as a Foucault knife-edge test, and in principle their presence can be taken into account or allowance can be made for them. As a consequence, the screen test gives results that are more representative of the actual surface features as the mirror surface becomes smoother. The need for this assumption is due to the fact that each sample point is taken to give the average-tilt for a certain area on the mirror surface. For the average-tilt approach to hold, one must assume surface smoothness. This assumption must also be made if a polynomial fit is used.

Also, it is assumed that the air turbulence is random in magnitude and direction between the mirror and the recording plane. The main turbulence condition that influences the results is the case of a laminar sheet of air rising or falling. Such type of turbulence will uniformly displace the light from a number of holes and systematically alter the results. Frequently, this condition is present on a small scale, and it is difficult to detect.

A common source for this sort of turbulence is a thermal imbalance across the width of the test beam. The typical situation arises when one is testing with the optical axis horizontal in the presence of a thermal gradient across the optical path, as can be found between the floor and the ceiling of a test area when their temperature is not controlled adequately. If the air reaches a stable condition, then a layering results, with the coolest and densest layer at the bottom and the warmest and lightest layer at the top. This air density gradient is also an index of refraction gradient, with the higher index at the bottom and the lowest at the top. Although this is not a turbulence in the usual sense, nevertheless, it results in an astigmatic wavefront when the system under test may not have such an aberration. This astigmatism is aligned with the vertical axis and can cause further confusion when astigmatism due to a drooping of the optics is expected because of self-weight deflection of those optics under the local gravity gradient.

The other typical thermal gradient situation arises when the test is performed with the optical axis vertical near the outside wall of a building. In such a case, if the wall is warm, a chimney effect could take place and obviously this would be undesirable. Similarly, if the wall is cold, a cascade of cold air would also cause problems for the test. Proper insulation and care in thermal control are the necessary design criteria in setting up the test under such circumstances.

The screen must be centered accurately on the aperture of the mirror under test. This is particularly true in the presence of spherical aberration, which is the case for aspheric mirrors tested at their center of curvature. A decentering of the screen in such a situation leads to an apparent presence of coma as described by Landgrave and Moya (1986). They present results for parabolic mirrors of various F numbers.

After the screen is placed and centered on the mirror, but before any plates are taken, the point light source used to illuminate the mirror through the screen must be centered properly relative to the mirror to prevent the introduction of off-axis

aberrations. This can be done conveniently if the housing of the point source is of such a size that it is narrowly missed by the beams returned by the region near the center of the mirror. Without elaborate equipment one can then center the source by judging and equalizing visually the gaps between the housing and the light beams.

10.3. HARTMANN TEST USING A RECTANGULAR SCREEN

The distance from the wavefront at the exit pupil to the Hartmann plate is r and the linear transverse aberrations are TA_x and TA_y . The wavefront can thus be calculated by integration of the transverse aberrations measured at the Hartmann plate.

For an adequate identification of the spots in the Hartmann plate, this plate has to be located displaced from the focus, frequently inside the focus, so that the light beams do not cross each other. In other words, if the wavefront has spherical aberration, the Hartmann plate must be placed outside the caustic region.

As pointed out by Gozheil in a previous version of this chapter, the use of the square screen was suggested by R. V. Shack in private communications, and it was first employed to test a number of large telescope primary mirrors, like the first 4-m telescope, made at Kitt Peak National Observatory (Simmons and Ghozeil, 1971; Ghozeil and Simmons, 1974); the 3.8-m Anglo-Australian telescope, which was tested at Grubb Parsons; the second 4-m telescope, made at Kitt Peak for the Cerro Tololo Interamerican Observatory (Ghozeil 1974); the 2.6-m telescope for the Irénée Dupont Observatory, made at the Optical Sciences Center of the University of Arizona; and the 3.6-m telescope for the Canada-France-Hawaii Observatory, made at the Dominion Astrophysical Observatory. The square array not only gives uniform sampling of the surface but also does so in a manner that has no circular symmetry. Thus, there is no need to make assumptions concerning the disposition and symmetry of the errors to be detected, and the possibility of an artificial circular error buildup has been removed. Furthermore, a much higher surface-sampling frequency can be obtained with this method than is possible with radial or helical screens. The tests performed in this way are independent of random air turbulence and can be done with or without intervening optics, such as null lenses. The second option of having no intervening matter other than air is preferred, so that the introduction of confounding alignment and surface errors caused by intervening materials can be prevented.

As the mirror nears completion, there is a need for a more detailed and reliable representation of the surface under test. This can be achieved through the superposition and averaging of several tests obtained with the screen rotated at some known amount about its center, relative to the mirror (Ghozeil, 1974). The superposition can be done numerically with an electronic computer or photographically by expressing different surface deviations by corresponding densities on a photographic plate. The deviations obtained from separate tests for each region are used to obtain new average values for all the regions. This composite method is capable of giving surface deviations with an uncertainty of less than 0.05 wavelength. Results obtained in this way have shown faint small features observable with a knife-edge test, while

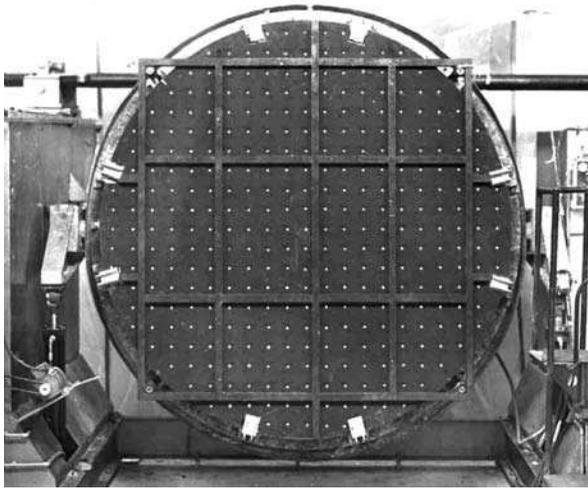


FIGURE 10.4. Hartmann screen over the 2.1 m primary mirror of the Guillermo Haro Observatory telescope made at the Instituto Nacional de Astrofísica, Óptica y Electrónica, en Tonatzintla Pue. Mexico.

simultaneously revealing overall slowly varying asymmetrical features. The square array method of testing requires the use of a precision centroid finding method.

With this test method, as with other screens, it is not possible to detect small scale surface changes taking place between holes in the screen. However, as shown by Gozheil (1974), this test, and potentially the helical screen method, can be made to

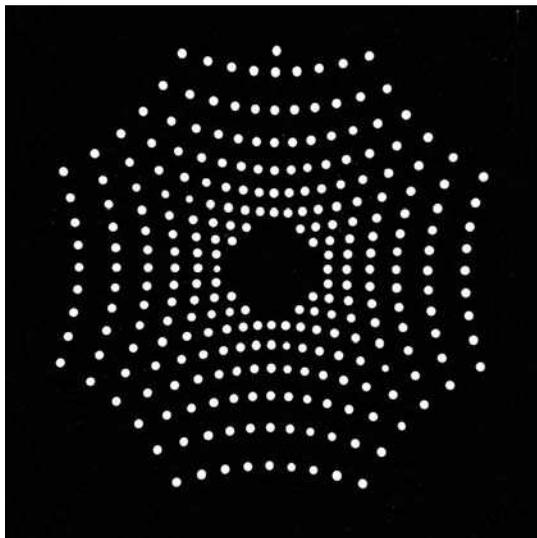


FIGURE 10.5. Hartmann spots of the 2.1 m primary mirror made at the Instituto Nacional de Astrofísica, Óptica y Electrónica, en Tonatzintla Pue. Mexico.

show zones that are small in extent through a composite of several tests of the same mirror under the same conditions, but with the screen rotated at a known amount relative to the surface.

The screen is placed near the mirror as shown in Figure 10.4, which shows the screen used with the two 2.1-m primary mirror made at Instituto Nacional de Astrofísica, Óptica y Electrónica (INAOE), in Tonatzintla, Mexico, and shows the 341 holes in the screen and the ribbed structure incorporated in the screen to make it rigid. The Hartmann pattern for this hyperboloid mirror is shown in Figure 10.5.

10.4. WAVEFRONT RETRIEVAL

Several methods of integration can be used with somewhat more reliable results. Some of these methods are as follows.

- (a) The well known trapezoidal integration. In most cases this trapezoidal rule will suffice.
- (b) The Southwell integration algorithm.
- (c) The polynomial fitting of the wavefront.
- (d) Other methods.

10.4.1. Tilt and Defocus Removal

Any possible lateral displacement of the whole Hartmann pattern, which would introduce a tilt in the retrieved wavefront, must be subtracted from the data. To do so, the centroid of the measured dot positions need to be found and established as the origin of the coordinate system to be used in the reduction. This is done by averaging all the x and y values of the measured dot locations and then subtracting these average values from each measurement. The reason for this step is that the intersection of the plate with the optical axis of the mirror is unknown and the centroid is the best method for establishing this intersection. Surface deformations that occur randomly or symmetrically will result in a set of spots whose centroid corresponds with the location of the intersection of the optical axis and the plate.

The necessary large defocusing in the Hartmann test introduces a large linear component on the transverse aberrations, which is equivalent to a large quadratic component in the wavefront. To greatly reduce the effect of this defocusing, the transverse aberrations must be obtained with respect to the spots in an ideal Hartmann pattern with the same defocusing. Ideally, a least-squares fitting of the two measured transverse aberrations to a pair of linear expressions with the same linear coefficient must be obtained. The ideal transverse aberrations Δx and Δy with tilt and defocusing, for a given spot, can be written as

$$\begin{aligned}\Delta x_m &= A_x + Bx_m \\ \Delta y_m &= A_y + By_m\end{aligned}\tag{10.3}$$

where A_x , A_y , and B are constants and x and y are measured on the corresponding hole on the Hartmann screen.

This procedure can be separated in two steps, using the fact that the Hartmann screen is symmetrical. Then, for every x coordinate of a hole there is a negative counterpart. As a consequence, summations over x reduce to zero. Thus, the centroid can be set by

$$A_x = \frac{\sum_{m=1}^N \Delta x_m}{N} \quad (10.4)$$

$$A_y = \frac{\sum_{m=1}^N \Delta y_m}{N}$$

where N is the number of holes over which the summations are to take place.

After the centroid has been established as the origin of the coordinate system to be used, the ideal positions for the defocused pattern must be calculated. A defocus situation results in a radial linear departure of the ray intercepts. Thus, the residual deviations can be brought to the best focus by subtracting the linear term, which is calculated by means of a two-dimensional linear least-square fit. Both, the centroid (or tilt) and the defocusing removal can be performed with this least squares as follows:

$$B = \frac{\sum_{m=1}^N x_m (\Delta x_m) + \sum_{m=1}^N y_m (\Delta x_m)}{\sum_{m=1}^N x_m^2 + \sum_{m=1}^N y_m^2} \quad (10.5)$$

Once the constants have been found, the values of the residuals at the best focus, illustrated in Figure 10.6(b), can be determined by subtracting the value of Δx_m from the measured values of TA_x and TA_y . This gives

$$TA_{x\ m} = TA_{x\ m\ measured} - \Delta x_m \quad (10.6)$$

where $TA_{x\ m}$ is the new value.

The new transverse aberration values can be used to obtain an estimate of the energy distribution of the light returned by the mirror under test. This is done by counting the number of points having a distance $r = \sqrt{TA_{x\ m}^2 + TA_{y\ m}^2}$ from the origin of the coordinate system. If the test is performed at the center of curvature and the energy distribution at the focus is desired, the residuals can be divided by a factor of 2. This division is a good approximation that introduces a small error, equivalent to a surface deformation of less than 0.05 wavelength, at the edge of an $f/2.7$ mirror. It is worth remembering that this energy distribution is the result of geometrical optics considerations and that the real image obtained with the mirror

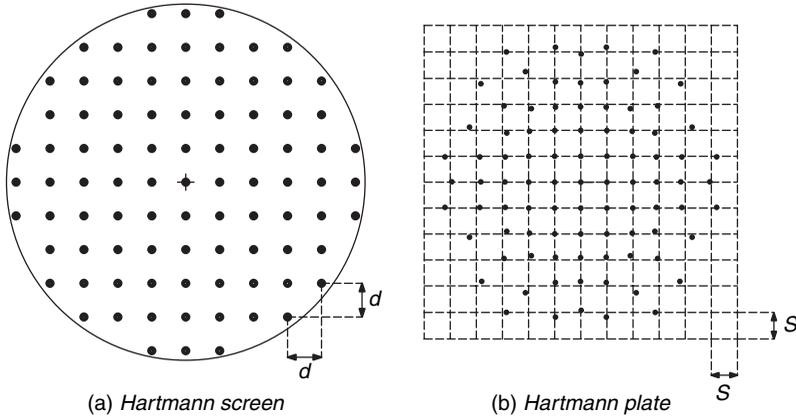


FIGURE 10.6. Schematics of a rectangular array of holes and the positions of the registered spots at the Hartmann plate. The rectangular array with dotted lines represent the positions of the spots for an ideal aspheric mirror.

will be affected by diffraction. In general, that effect will tend to give a higher energy concentration in the core of the image.

If the wavefront is highly aspherical, it is convenient also to subtract the aspheric component before retrieving the wavefront. This can be done by one of the two different processes, that is by a least-squares fit to a cubic transverse aberration function, representing primary spherical aberration or by taking the difference between the newly established x and y coordinates of each dot, and the corresponding coordinates expected from a perfect mirror is found. The expected coordinates of the spot can be obtained from geometrical ray tracing or from closed form equations such as Eq. (9.22) in this book (Malacara, 1965).

10.4.2. Trapezoidal Integration

This method of wavefront retrieval, sometimes call a *zonal method*, is illustrated in Figure 10.7, with the linear approximation in the TA measurements; the wavefront shape is continuous, synthesized by second-order segments with constant curvature in each segment. Thus, if the wavefront is quadratic, which happens when only tilts, defocus, and astigmatism are present, the trapezoidal integration becomes exact.

Integrating expressions (10.1) and (10.2), we obtain (the minus sign is ignored for simplicity)

$$W(x, y) = \frac{1}{r} \int_0^x TA_x(x, y) dx \quad (10.7)$$

and

$$W(x, y) = \frac{1}{r} \int_0^y TA_y(x, y) dy \quad (10.8)$$

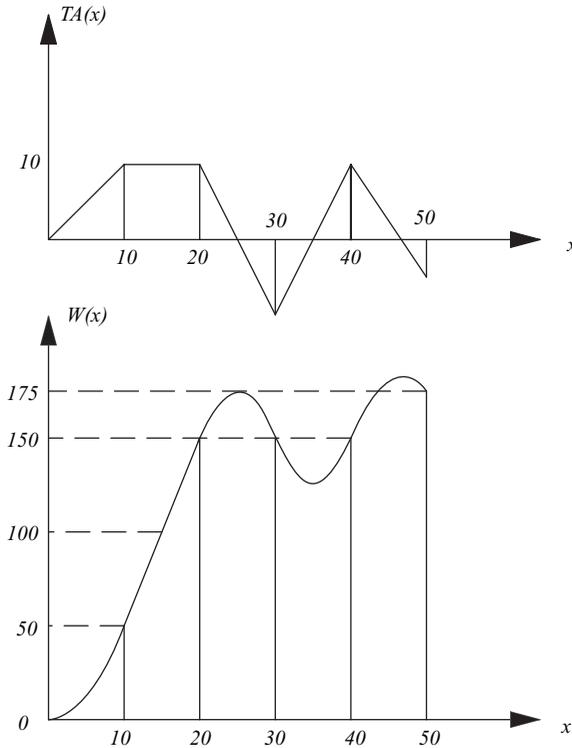


FIGURE 10.7. Illustration of a trapezoidal integration with straight lines joining the measured values of the transverse aberration. The wavefront is reconstructed with quadratic segments. The wavefront is continuous at the measured points, but not the wavefront slopes.

Since the measurements are made only at points in a rectangular array, in the classical Hartmann test the wavefront shape is found by a first-order trapezoidal integration of Eqs. (10.1) and (10.2) by writing

$$W_{n,m} = W_{n-1,m} + \frac{d}{2r} [TA_x(n-1,m) + TA_x(n,m)] \tag{10.9}$$

along the x axis and

$$W_{n,m} = W_{n,m-1} + \frac{d}{2r} [TA_y(n,m-1) + TA_y(n,m)] \tag{10.10}$$

along the y axis, and

$$W_{n,m} = W_{n-1,m-1} + \frac{d}{2r} [(TA_x(n-1,m-1) + TA_x(n,m-1)) + (TA_y(n,m-1) + TA_y(n,m))] \tag{10.11}$$

along a diagonal line. Here, d is the separation between two consecutive holes in the Hartmann screen. This expression is calculated for successive spots coming from holes along a line on the Hartmann screen, beginning with $W(0, m)$, $W(n, 0)$, $TA(0, m)$, and $TA(n, 0)$ equal to zero. Then, a new line is scanned, until the whole pattern is covered from several directions. These expressions give the surface deviation at any point (n, m) , taking the first point as a reference.

The holes on the Hartmann screen should be placed very accurately to avoid errors in computing the surface deviations. The hole diameter should be small but not so small that their diffraction images on the photographic plate overlap each other and should be of sufficiently small size so as to permit an accurate measurement of the wavefront aberrations present (Golden, 1975; Vitrichenko et al., 1975; Morales and Malacara, 1983).

Because numerical integration methods tend to have an accumulation of errors inherent in them, some care must be taken to reduce that error. This is best done through the use of paths of integration that intersect at only one point, which means that a surface height at any location is obtained by independent means. It is also worthwhile to integrate along any path in one direction and in its reverse direction, and then average the results obtained.

The scheme used in the integration follows a philosophy that is shown schematically in Figure 10.8. First, the summations originate on the x and y axes and have as

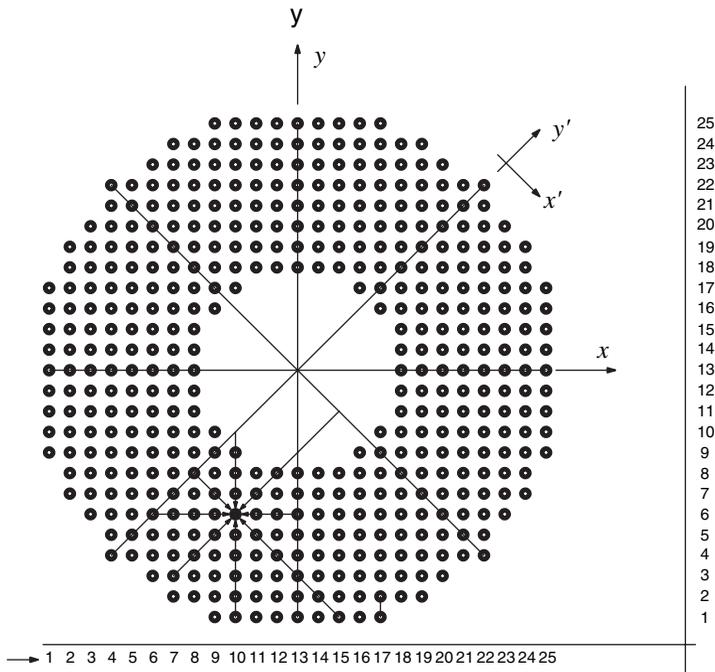


FIGURE 10.8. Diagram showing the holes on the Hartmann screen. Different integration paths are followed along vertical lines, horizontal lines, and also lines at 45° and -45° .

starting value of each integral the value obtained from the other coordinates, integral going through the point in question. Since the surface departure values obtained through the x and y integrations should be the same at each point, the results of the two integrations are averaged at each point. Then, the reverse path summation is done and averaged with its corresponding integral.

For greater accuracy, the next step can be to rotate the coordinate axes by 45° about the origin and to repeat the entire integration process, using a different hole interval and different integration paths. Since these integrations should give the same results as those obtained with the first scheme, the values obtained by both the processes are averaged at each point. This means that each surface departure value is obtained in at least four ways, with a majority obtained in eight ways. This replication tends to reduce the systematic accumulation of errors, as well as the introduction of spurious ones.

To further reduce spurious errors, two photographic plates may be measured and reduced in this fashion, and the results of the two tests may be averaged together. This reduces errors due to imperfect photographic emulsions and nonrandom air turbulence between mirror and plate, as well as measuring errors.

10.4.3. Southwell Algorithm

This method, sometimes called the Southwell (1980) reconstructor is also a zonal method. When the linear integration is made from one point to another following two different paths, the two results may be slightly different due to several factors like measurement errors in the spot positions, local wavefront deviations from sphericity, and numerical errors. The idea behind this method is to take into account the wavefront deformation for some vertical and horizontal adjacent spots in the calculation of the wavefront deformation for all spots.

Southwell proposed an iterative solution where for any point (n, m) the wavefront is calculated with Eqs. (10.10) and (10.11) integrating from four neighboring points, one above, one below, one at the left, and one to the right of the point being considered as illustrated in Figure 10.9. Then, the final value $W_{n,m}$ is just the average of the four values. For notational simplicity, we define $S_{n-1,m}^x$ as

$$S_{n-1,m}^x = \frac{d}{2r} [TA_x(n-1, m) + TA_x(n, m)] \quad (10.12)$$

Thus, with Eqs. (10.12) and (10.13) for the four neighboring points in Figure 10.9 we can write

$$\begin{aligned} W_{n,m} &= W_{n-1,m} + S_{n-1,m}^x \\ W_{n,m} &= W_{n+1,m} - S_{n+1,m}^x \\ W_{n,m} &= W_{n,m-1} + S_{n,m-1}^y \\ W_{n,m} &= W_{n,m+1} - S_{n,m+1}^y \end{aligned} \quad (10.13)$$

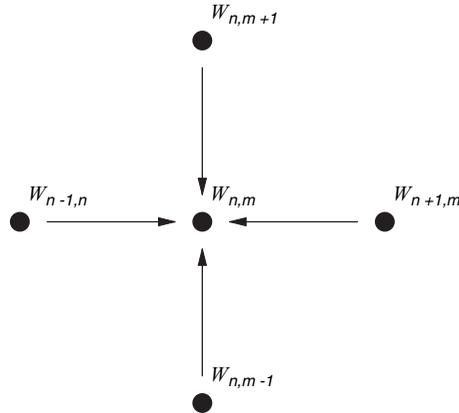


FIGURE 10.9. Wavefront calculation for point n, m going to this point from four neighboring points in cross.

Thus, the weighted average of these four measurements is

$$W_{n,m} = \frac{(\sigma_{n-1,m} W_{n-1,m} + \sigma_{n+1,m} W_{n+1,m} + \sigma_{n,m-1} W_{n,m-1} + \sigma_{n,m+1} W_{n,m+1})}{(\sigma_{n-1,m} + \sigma_{n+1,m} + \sigma_{n,m-1} + \sigma_{n,m+1})} + \frac{(\sigma_{n-1,m} S_{n-1,m}^x - \sigma_{n+1,m} S_{n+1,m}^x + \sigma_{n,m-1} S_{n,m-1}^y - \sigma_{n,m+1} S_{n,m+1}^y)}{(\sigma_{n-1,m} + \sigma_{n+1,m} + \sigma_{n,m-1} + \sigma_{n,m+1})} \quad (10.14)$$

where the weights σ_{nm} are defined as one for all spots. This expression is calculated for all points (n, m) in a square array where the whole Hartmann screen is included. All spots that are nonexistent or located outside of the Hartmann screen should have the corresponding weight factor σ equal to zero. After all the spots had been covered, a second iteration is made, until about a number of iterations equal to the total number of spots are completed.

10.4.4. Polynomial Fitting

We may also use a polynomial fit to express the changes in slope deviations present in the mirror. This procedure is frequently called a *modal method*, because each polynomial component can be considered as a mode. The fit is made to agree with some criterion governing how closely the fit and the sample points should agree. Thus, rather than using tilted flat planes to perform the integration to obtain the surface departures, one integrates over a polynomial (Cornejo and Malacara, 1976). The wavefront can then be written in terms of Zernike polynomials. Theoretically, this is ideal because errors due to numerical integration method are removed. The difficulty lies in obtaining a good polynomial fit. It is known that such fitting can introduce errors by smoothing out relevant sharp features or by introducing oscillations when none exists. The first effect usually is due to

low-order polynomial fits and the second effect is due to high-order fits. Needless to say, this leads to a reduction in the reliability of the test procedure, especially near the edge of the mirror.

A polynomial fit can be made through a least-squares method as described next. Let us assume that the wavefront can be represented by a linear combination of some polynomials, for example, Zernike polynomials, as

$$W(x, y) = \sum_{n=0}^N A_n P_n \quad (10.15)$$

where N is the number of polynomials being used (it is not the maximum power) and A_n are the unknown coefficients of the linear combination. Then, from Eqs. (10.1) and (10.2), we can write

$$TA_x(x, y) = r \sum_{n=0}^N A_n \frac{\partial P_n}{\partial x} \quad (10.16)$$

and

$$TA_y(x, y) = r \sum_{n=0}^N A_n \frac{\partial P_n}{\partial y} \quad (10.17)$$

Then, a least-squares fit of these analytical functions to the measured corresponding values of the transverse aberrations can be made. From a polynomial fit, one can then extract classical aberration terms such as coma and astigmatism.

10.4.5. Other Methods

Another approach to retrieve the wavefront with a two-dimensional fast Fourier transform (FFT) algorithm has been reported (Freischlad and Koliopoulos, 1985). Its main advantage is that it is less sensitive to noise in the data, which has been used to obtain two-dimensional wavefront errors.

Also, some of the methods used to demodulate interferograms can be used with Hartmann patterns as shown by Servín et al. (1996 and 1999).

If a non-linear approximation is desired when a strongly aspheric wavefront is to be measured, another procedure would be to use two-dimensional cubic spline functions (Ahlberg et al., 1967). If properly used, this method of fitting can give a better approximation between sampling points than with the trapezoidal rule, which assumes quadratic curves instead of straight lines in the transverse aberration functions between two consecutive points. For this method to improve results, it is necessary that the transverse aberrations to be integrated are measured with respect to the ideal aspheric wavefront and not with respect to a close sphere.

10.5. HARTMANN TEST USING A SCREEN WITH FOUR HOLES

Sometimes, the detailed wavefront topography is not desired and only a measurement of some primary aberrations are needed. In this case, the Hartmann test can be performed with only four holes near the edge of the pupil of the system. If we have only four holes, we take two measurements, that is, TA_x and TA_y , for the light from each hole. Thus, we have a total of eight measurements. We can determine a maximum of eight aberration terms, where the constant or piston term cannot be included. These aberrations can be

$$W(x, y) = Bx + Cy + D(x^2 + y^2) + E(x^2 - y^2) + Fxy + G(x^2 + y^2)y + H(x^2 + y^2)x \quad (10.18)$$

where the coefficients B , C , D , E , F , G , and H represent the following:

- B = Local cell tilt about the y axis.
- C = Local cell tilt about the x axis.
- D = Defocusing (local cell spherical deformation).
- E = Astigmatism with axis at 0° or 90° .
- F = Astigmatism with axis at $\pm 45^\circ$.
- G = Coma along the y -axis.
- H = Coma along the x -axis.

The next aberrations are the two triangular astigmatism terms, with bases along each of the two axes, x and y , but we cannot find both with only eight measurements. So, they are not to be calculated.

If the tilt and coma terms are not present the wavefront has a toroidal shape whose radial curvature can be written (Salas-Peimbert et al., 2005) as

$$c_r = [2D + (4E^2 + F^2)^{1/2}] - 2(4E^2 + F^2)^{1/2} \sin^2(\theta - \alpha) \quad (10.19)$$

where the first term is the mean spherical contribution and the second term is the astigmatic contribution with an orientation (astigmatic axis) α given by

$$\tan 2\alpha = \frac{F}{2E} \quad (10.20)$$

Let us now consider the four Hartmann spots produced by the light beams passing through the four holes at the Hartmann screen. Then, the x components of the transverse aberrations are given by

$$\frac{\partial W(x, y)}{\partial x} = \frac{TA_x}{r} = B + 2Dx + 2Ex + Fy + 2Gxy + H(3x^2 + y^2) \quad (10.21)$$

and the y components of these aberrations are given by

$$\frac{\partial W(x, y)}{\partial y} = \frac{TA_y}{r} = C + 2Dy - 2Ey + Fx + G(x^2 + 3y^2) + 2Hxy \quad (10.22)$$

We have not yet considered any particular configuration for the four holes. The four holes can be located at $\pm 45^\circ$ (in X) or at $0^\circ, 90^\circ, 180^\circ$, and 270° , (in cross) as will be considered next.

10.5.1. Four Holes in Cross

Now we will consider the following coordinates for the four holes, as described by Malacara and Malacara (1992) and illustrated in Figure 10.10(a): $\alpha(-d/2, 0)$, $\beta(0, -d/2)$, $\gamma(-d/2, 0)$, and $\delta(0, -d/2)$.

We can apply Eqs. (10.21) and (10.22) to the four holes to find the aberration coefficients and by taking averages where redundant information is available, as in the case of the G and H coefficients; we finally find all the aberration coefficients for the square cell to be given by

$$\begin{aligned} B &= -\frac{(TA_{xz} + TA_{x\gamma}) - 3(TA_{x\beta} + TA_{x\delta})}{4r} \\ C &= -\frac{3(TA_{yz} + TA_{x\gamma}) - (TA_{y\beta} + TA_{y\delta})}{2r} \\ D &= -\frac{(TA_{xz} - TA_{x\gamma}) + (TA_{y\beta} - TA_{y\delta})}{4rd} \\ E &= -\frac{(TA_{xz} - TA_{x\gamma}) - (TA_{y\beta} - TA_{y\delta})}{4rd} \\ F &= -\frac{(TA_{yz} - TA_{y\gamma})}{2rd} \\ G &= -\frac{(TA_{yz} - TA_{y\gamma} - TA_{y\beta} - TA_{y\delta})}{2rd^2} \\ H &= \frac{(TA_{xz} + TA_{x\gamma} - TA_{x\beta} - TA_{x\delta})}{rd^2} \end{aligned} \quad (10.23)$$

The typical patterns for some primary aberration are given in Figure 10.11(a).

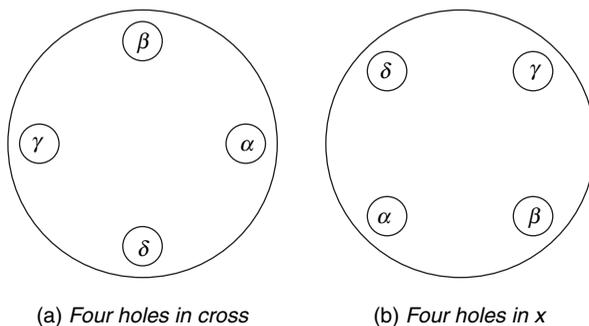


FIGURE 10.10. Two Hartmann screens with four holes:(a) in cross and (b) in X.

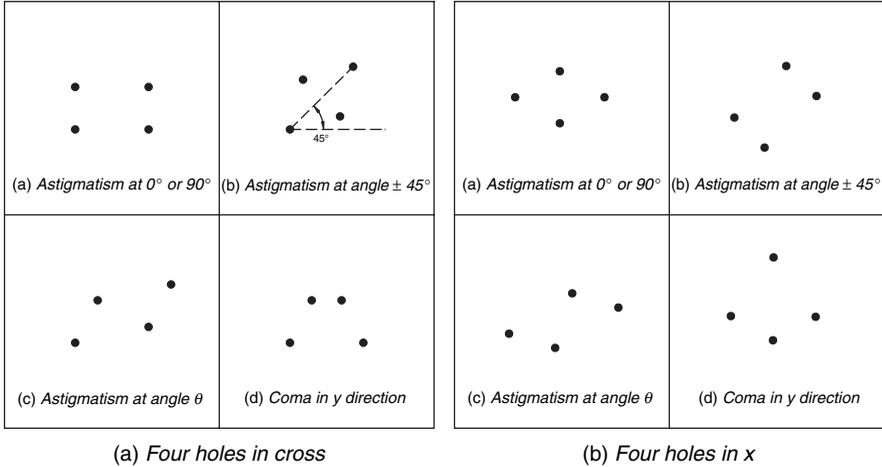


FIGURE 10.11. Typical hartmann patterns for a Hartmann screen with four holes: (a) in cross and (b) in X.

10.5.2. Four Holes in X

Now we will consider the following coordinates for the four holes, as described by Salas Peimbert et al. (2005) and illustrated in Figure 10.10(b): $\alpha(-d/2, -d/2)$, $\beta(d/2, -d/2)$, $\gamma(-d/2, d/2)$, and $(d/2, d/2)$. In a manner similar to the case of the four points in cross, we can find

$$\begin{aligned}
 B &= \frac{(TA_{x\alpha} + TA_{x\beta} + TA_{x\gamma} + TA_{x\delta}) - 2(TA_{y\alpha} + TA_{y\beta} + TA_{y\gamma} - TA_{y\delta})}{4r} \\
 C &= \frac{(TA_{y\alpha} + TA_{y\gamma} + TA_{y\beta} + TA_{y\delta}) - 2(TA_{x\alpha} - TA_{x\beta} + TA_{x\gamma} - TA_{x\delta})}{4r} \\
 D &= -\frac{(TA_{x\alpha} - TA_{x\beta} - TA_{x\gamma} + TA_{x\delta}) + (TA_{y\alpha} + TA_{y\beta} - TA_{y\gamma} - TA_{y\delta})}{8rd} \\
 E &= -\frac{(TA_{x\alpha} - TA_{x\beta} - TA_{x\gamma} + TA_{x\delta}) - (TA_{y\alpha} - TA_{y\beta} - TA_{y\gamma} - TA_{y\delta})}{8rd} \quad (10.24) \\
 F &= -\frac{(TA_{x\alpha} + TA_{x\beta} - TA_{x\gamma} - TA_{x\delta})}{2rd} \\
 G &= \frac{(TA_{x\alpha} - TA_{x\beta} + TA_{x\gamma} - TA_{x\delta})}{2rd^2} \\
 H &= \frac{(TA_{y\alpha} - TA_{y\beta} + TA_{y\gamma} + TA_{y\delta})}{2rd^2}
 \end{aligned}$$

The typical patterns for some primary aberration are given in Figure 10.11(b).

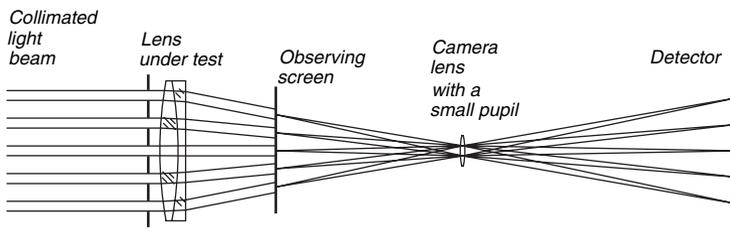


FIGURE 10.12. Hartmann pattern projected on a transparent diffusing observing screen. This pattern is then observed with a camera lens with a small aperture.

10.6. HARTMANN TEST OF OPHTHALMIC LENSES

Salas-Peimbert et al. (2004) have applied the Hartmann test to the measurement and evaluation of ophthalmic lenses. These lenses can have positive as well as negative focal lengths. Also the wavefront produced can be spherical, toric, or even highly aspherical as in the progressive power lenses. In conclusion, if they are illuminated with a collimated beam of light, the refracted wavefront can be convergent or divergent and spherical or aspherical.

If the refracted wavefront is projected on a translucent diffusing screen, even larger than the lens, as illustrated in Figure 10.12, the Hartmann pattern can be observed with a digital television camera to be digitally processed. With this method, even the refractive power of the lens can be determined. It is only necessary to measure the distance from the vertex of the ophthalmic lens to the observing screen.

The presence of a large distortion aberration on the camera lens can affect the results, but this distortion can be numerically compensated and eliminated when processing the Hartmann pattern. So, it is desirable to previously calibrate the camera lens by observing a rectangular array placed at the desired observing distance.

10.7. HARTMANN TEST USING NONRECTANGULAR SCREENS

Over the years, sampling screens with a number of different hole patterns have been used in testing mirror surfaces. The various types reported in the literature along with their advantages and disadvantages are discussed in this section. Although recently there has been an increase in the use of lenticular screens, we will continue to use “hole” as a short term instead of the more descriptive term “sampling aperture” in the following discussion. It will be made obvious to the reader, in later parts of the text, when a lenticular screen is being discussed and that the term “hole” is not applicable.

For convenience in use, it is worthwhile that a few holes be added that do not fall on the hole pattern of interest. These extra holes serve to aid in orienting the screen and subsequently the recorded pattern of dots. Classically, photographic plates were used to record the location of dots obtained from the screen tests and for subsequent measuring of dot locations recorded on these plates.

10.7.1. Radial Screen

By far, the most common screen in use, until recently, was the radial pattern shown in Figure 10.13. This type of screen was used first by Hartmann (1900, 1904) and subsequently, without substantial change in the original concept, by a number of other experimenters. A bibliography of the work published on the radial screen test is given at the end of this chapter.

The basic concept of the radial pattern test is that a sample of points on the surface of a concave mirror that has a circular aperture can be taken using an opaque cover in which a number of small holes have been made. The holes are spaced evenly along a number of diameters of the circular mirror aperture as shown in Figure 10.13. The choice of such a pattern has several advantages. First, a circular aperture is easily analyzed in a polar coordinate system. Second, the grid pattern will detect the most common flaws of mirror surfaces of any size zonal errors and concentric “hills” and “valleys” fractions of a micrometer in magnitude that result from classical grinding and polishing techniques. Thus, in principle, the radial screen could afford the luxury of an easily analyzed test that detects the major noticeable surface defects.

In fact, the radial screen test has many shortcomings. A major one is that if one considers each hole in the screen to be sampling an area on an annulus having as mean radius the radial location of the hole, it follows that the area sampled is considerably larger for holes far from the mirror center than for those near that center. This means that the area of the surface contributing to the major portion of the light gathering is known with least certainty, with the result that asymmetric defects not detectable by other means are likely to go undetected by this method as well. Another major shortcoming is due to the circular symmetry of some of the defects, which go undetected if the hole spacing is not sufficiently small. Intertwined with these procedural disadvantages is the assumption, often made in the data reduction to

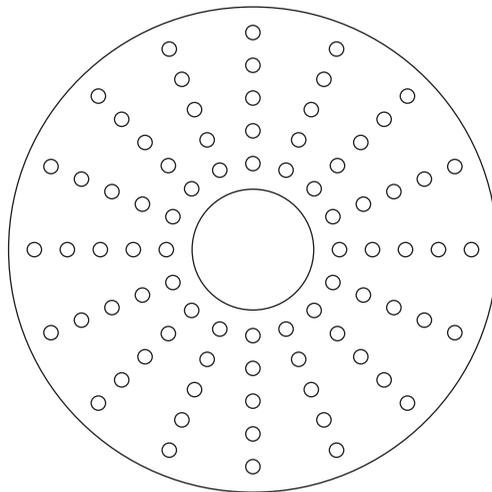


FIGURE 10.13. Classical Hartmann radial pattern.

deduce the surface, that the slope deviations detected are part of a circularly symmetric system. The foregoing objections can be reduced somewhat by using a radial screen with holes spaced moderately close together, with many diameters sampled and consequently with some of the diametral arms not having holes near the mirror center, to prevent overcrowding in that region. The assumption on symmetry could also be abandoned. Incorporating these changes into the classical test, however, would lower its ease of reduction.

Also present in this form of testing is a buildup of integration errors along the circular paths of integration, as discussed in the next section on application. This error accumulation could be eliminated by a method similar to the replication developed for the square array screen (Ghozeil and Simmons, 1974), but applied to circular integration.

The radial screen can be used for testing a concave mirror in collimated light, or with a point source placed at or near the paraxial center of curvature. The first test is suitable for systems such as telescopes. It requires that the screen be placed either flush with the surface or sufficiently far from the concave mirror so that the beams, after reflection, do not strike the screen before they come to focus. When the mirror is under test with light from a point source, the screen suffers from the same location constraint already described and should be placed as close to the mirror as possible.

In the radial screen method, it has been customary to record, on two photographic plates, the light reflected from the mirror, with the mirror illuminated through the screen. One plate can be placed inside and the other outside the theoretical point to which the light should be converging, or both plates can be on the same side of that location. Either of the two-plate techniques allows one to compensate for plate tilt and for errors in hole spacings in the screen. The mirror for recording plane distance need not be known accurately, but the plate separation is important in the computation unless a least-squares refocusing is to be done as described in Section 10.3.3.

In the classical test reduction, where the photographic plates are placed with one inside and the other outside of focus, a mean weighted focus position, F_0 , can be obtained as follows:

$$F_0 = \frac{\sum R_i F_i}{\sum R_i} \quad (10.25)$$

where F_i is the focus obtained with the light from a pair of screen holes diametrically located to either side of the vertex, and R_i is the radial distance of the “zone” that contains that pair of points. These values of R_i act as weight factors. To obtain an indication of the size of a point imaged by the lens, the Hartmann T criterion is used, as expressed by

$$T = \frac{200,000}{F_0^2} \left(\frac{\sum R_i^2 |F_i - F_0|}{\sum R_i} \right) \quad (10.26)$$

This gives the point image size in seconds of arc.

The astigmatism of the mirror can be estimated by comparing the average focus obtained from another arm of the radial screen with the average focus obtained from another arm perpendicular to the first one. However, this utilizes a relatively small number of sampling points and, consequently, does not give a very reliable estimate of that astigmatism.

Only one plate at a time needs be reduced if the screen is well known, as was the case for the 3-m (120-in.) Lick Observatory primary mirror (Mayall and Vasilevskis, 1960). This modification is a great time saver if there are many points to measure on the photographic plates.

The measured spot locations are reduced to give surface deviations by the general method described in Section 10.2. In the classical form of the test, only radial integration profiles are obtained. More recently, radial and tangential integrations have been used (Mayall and Vasilevskis, 1960; Schulte, 1968; Vitrichenko et al., 1975) to obtain more reliable results than the radial profiles alone. The two integration modes should differ only by a constant and should be made to agree. In fact, circular error buildup due to the tangential integration does not readily permit a simple shift to bring about the agreement. This method would be improved if some ways were established to introduce replication in the determination of surface deviations.

10.7.2. Helical Screen

In an attempt to overcome the disadvantages of the radial screen, a helically distributed set of holes in a screen was used to test the Lick Observatory 3-m (120-in.) primary mirror (Mayall and Vasilevskis, 1960). A portion of this pattern is shown schematically in Figure 10.14. Although basically a radial screen pattern, it differs from the classical pattern insofar, as holes along any radius are shifted radially relative to adjacent radii so as to form a helix as shown in Figure 10.15.

This choice of pattern affords an opportunity for the pattern to intercept zonal errors that might go undetected with a classical screen having the same hole spacing.

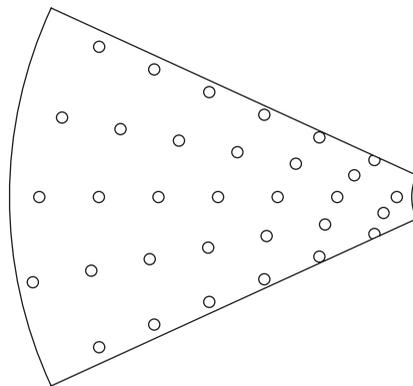


FIGURE 10.14. A portion of a helical screen showing the staggering of holes in the radial arms of the screen.

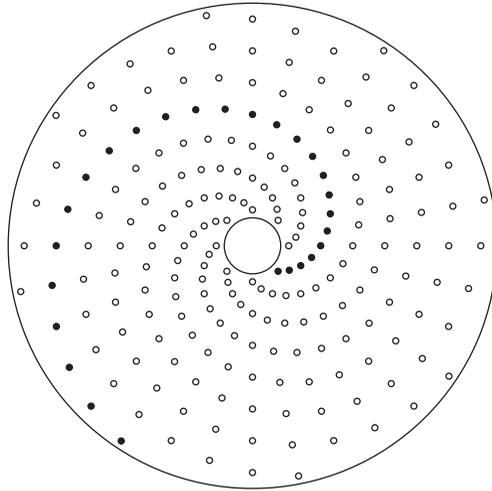


FIGURE 10.15. Complete helical screen showing the spiral distribution of holes.

Thus, it overcomes one of the disadvantages of the classical screen. However, it still suffers from unequal area sampling so that the most poorly sampled area of the mirror is the one that will gather a great percentage of the light that will be used in forming an image with the instrument in actual use.

This screen test makes extensive use of both radial and tangential integration paths in its application. This takes into account the reality that mirror deformations result, not merely in radial tilts, but in tilts having two components (Kingslake, 1927–1928) that are due to asymmetric surface departures from the ideal mirror surface. Difficulties arise in relating the slopes by sampling holes that are radially displaced relative to neighboring radial arms of the screen. The methods needed to reduce the data are thoroughly reported in the only known occurrence of the use of this screen (Mayall and Vasilevskis, 1960).

10.8. HARTMANN–SHACK TEST

Instead of using a screen with an array of holes, Platt and Shack (1971) proposed using a lenticular screen as illustrated in Figure 10.16. There are several small but important differences between this method and the classic Hartmann test, for example:

- (a) In the classic test, the Hartmann pattern is obtained in a convergent light beam close to the focus, whereas in the Hartmann–Shack sensor the test is made in a nearly collimated beam of light. If the wavefront under test is convergent, a collimating lens has to be used before the sensor. Figure 10.17 shows three possible optical arrangements.

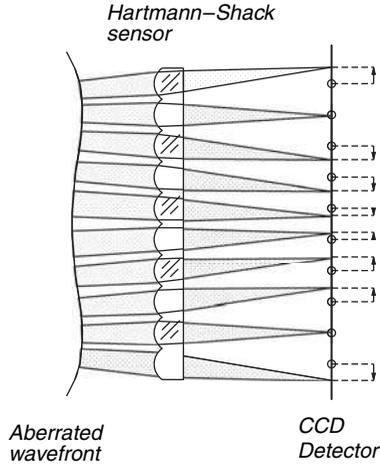


FIGURE 10.16. Hartmann–Shack test using a plate with a rectangular array of lenslets.

- (b) An advantage in the Hartmann–Shack method is that any positive or negative power can be easily detected and measured.
- (c) Each of the spots is individually focused on the detector, making the light energy density of the spot much higher than in the classic Hartmann test.

The Hartmann–Shack lenticular screen was originally made with two identical layers of cylindrical lenses perpendicular to each other. Now, they are made as lenslet

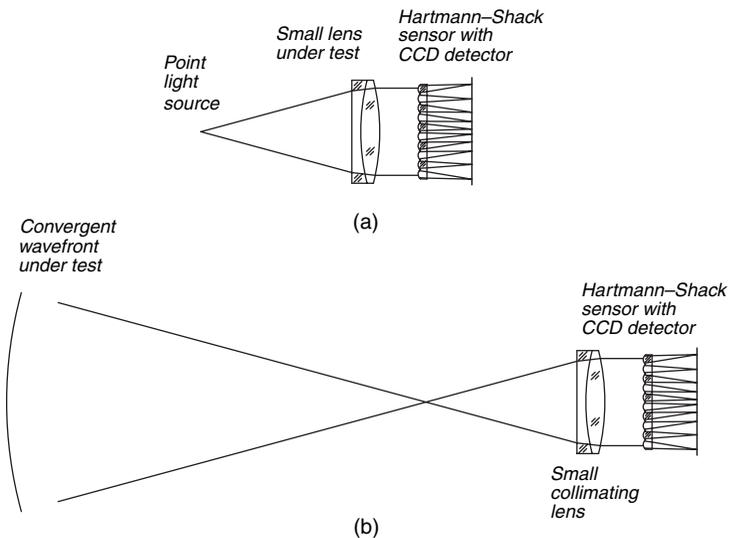


FIGURE 10.17. Wavefront testing with a Hartmann–Shack sensor: (a) A small lens about the same size of the sensor, (b) a convergent wavefront.

TABLE 10.1. Main characteristics of some lenslet arrays used in Hartmann–Shack sensors.

Array size	Full size in mm	Lenslet diameter in mm	Lenslet focal length in mm	Focal ratio $f/\#$
100×100	50×50	0.5	8.31	16.6
60×60	61×61	1.0	40.0	40.0
55×55	62×62	1.1	105.0	95.5
30×30	70×70	2.2	209.0	95.0

arrays in molded plastic, glass, or fused silica. The Hartmann–Shack plates are arrays between 10×10 to more than 100×100 lenslets. Each lenslet has a diameter in the approximate range 0.2–2 mm. A typical focal ratio $f/\#$, which is nearly the size of the spot, in μm is large, between $f/10$ and $f/200$. Table 10.1 shows the main characteristics of some lenslet arrays.

If the wavefront is flat, the light beam passing through each lens is focused near its optical axis. Since the lens array is not perfect, the lenticular array must be previously calibrated with a reference flat wavefront.

The diffraction spot size semidiameter ρ is given by

$$\rho = \frac{f\lambda}{d} \quad (10.27)$$

where d and f are the diameter and the focal length of the lenslets, respectively, and λ is the wavelength of the light. The angular dynamic range of this sensor is defined as the maximum possible angular aberration that can be measured without any crossing or overlapping of the spots. Thus, the maximum allowed spot deviation is about $d/2 - \rho$, so that the angular dynamic range becomes

$$\theta_{\max} = \frac{d/2 - \rho}{f} \quad (10.28)$$

The size σ of each pixel in the detector determines the angular sensitivity, which is defined by the inverse of the minimum measurable angular slope, given by

$$\theta_{\min} = \frac{\sigma}{f} \quad (10.29)$$

Since the spot displacement on the detector is equal to the wavefront slope multiplied by the focal length of the lenslet, a shorter focal length will give a greater dynamic range but a reduced angular sensitivity. The most adequate focal length will depend on the intended application. The several possible sources of error in the Hartmann–Shack test had been analyzed by some authors, for example, by Neal et al. (2002).

Because of its compactness, ease of calibration, high surface-sampling frequency, and other advantages, the method of reimaging the aperture onto a small screen is finding increasing applications (Loibl, 1980). The Hartmann–Shack wavefront sensor

has been widely used for adaptive optics in astronomy (Ftaclas, 2001), for eye measurements (Liang et al., 1994), and for aspheric lens testing (Greivenkamp et al. (2001).

10.9. CROSSED CYLINDER TEST

It has been pointed out that in the classic Hartmann test the observation plane has to be out of focus so that the hole from which each of the detected light spots can be properly identified. Once the Hartmann plate is taken, the amount of applied defocus can be measured if the radius of curvature of the wavefront is well known. However, this is not always the case and sometimes we would like to measure this radius of curvature from the Hartmann plate or be able to identify and measure any defocusing. Two possible examples are the human eye and a photographic camera. Also, in these two cases we cannot introduce a large defocusing, unless a convergent or divergent lens is placed in front of the system.

Howland (1968) has proposed to use a pair of crossed cylinders in front of the system, one of them convergent and the other divergent, to produce the desired defocusing, instead of using a single convergent lens. This arrangement has several

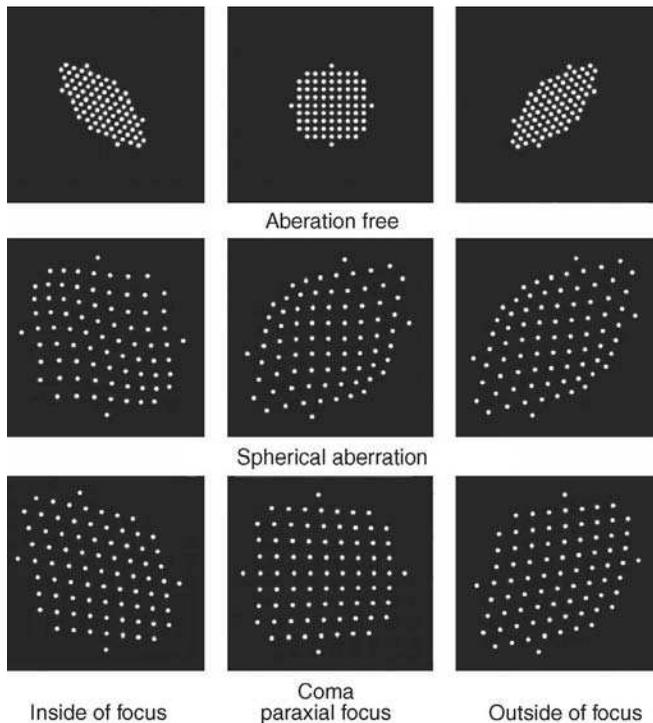


FIGURE 10.18. Hartmann patterns with the two crossed cylinder technique in the presence of primary aberrations.

advantages, since any defocusing can be measured, including its sign, from the Hartmann plate. The two crossed cylinders are placed in front of the system under test, a positive cylinder with its axis at -45° with respect to the x axis and a negative cylinder at $+45^\circ$ with respect to the x axis. Instead of a pair of crossed cylinders, a single lens with two opposite curvatures in orthogonal direction could be used, but it is more complicated and expensive to produce this lens than two cylinders.

Figure 10.18 shows the Hartmann patterns using this method for an aberration free system and for a system with spherical aberration and coma.

10.10. TESTING WITH AN ARRAY OF LIGHT SOURCES OR PRINTED SCREENS

Instead of using a screen with an array of holes in front of the pupil of the system under test and a single point light source, it is possible to use an array of independent light sources, each one illuminating a small zone on the pupil of the system. The working principle is the same, which is based on geometrical optics. There are several possible different configurations, depending on the type of optical element or system being tested.

Let us consider a lens with spherical aberration as in Figure 10.19. If the lens is illuminated by a wavefront with a distortion that just matches the spherical aberration, the refracted wavefront will be spherical and thus focused at single focus. Now, instead of illuminating with a single wavefront, an array of point light sources is placed on a plane in front of the lens. Then, the refracted rays originating from the rays traveling perpendicularly to the aspherical wavefront before entering the lens are selected by means of a pinhole at the focus of the lens. If the light sources in the array are equally spaced, ray intersections at any virtual or real observation plane will not be equally spaced, as in an aberration free lens.

The distance from the ray intersection to the ideal ray intersection in the observation plane can be interpreted as the transverse aberration. The radius of curvature of the wavefront is the distance from the observation plane to the pinhole. Although the observation plane can be located anywhere, ideally, the observation plane is on the

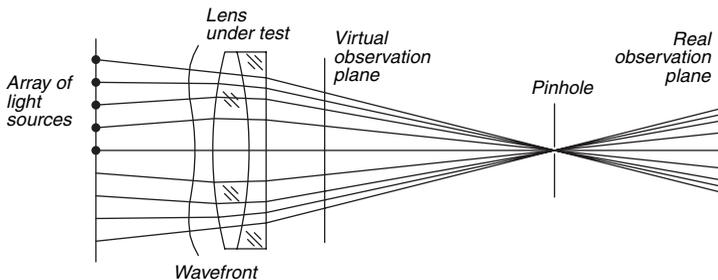


FIGURE 10.19. A lens with spherical aberration, illuminated by an aberrated wavefront that just matches the spherical aberration of the lens. The refracted wavefront is spherical.

pupil of the system under test. Then the distance between the real and the ideal intersections is the transverse aberration $TA(x, y)$, and the location of the intersection of the real ray has the coordinates x, y , with origin at the center of the pupil. It must be noticed that since neither the pinhole nor the light sources are infinitely small, and the light spots being observed are not small either. Similar to the classic Hartmann test, the centroid of the spot can be taken. This measurement principle can be used to test lenses or optical surfaces as will now be described.

A ruled screen, that is, a screen with two rectangular arrays of straight lines drawn on it can be used instead of using arrays of point light sources. Otherwise, the test is nearly the same and its pattern interpretation is identical. These tests will be briefly described in the next paragraphs.

10.10.1. Testing Convergent Lenses

A convergent lens can be tested as illustrated in Figure 10.20. An array of point sources is placed in front of the lens. If the pupil of the observing system is at the focus, the object is assumed to be at infinity. The distance from the lens to the light source is relatively unimportant, but has to be large enough so that the size of each of the light sources and any lateral displacement from its correct position do not affect the results much.

If the lens is not designed for an object at infinity, but at a finite distance, the imaging lens has to be placed at the correct image position. This imaging lens has a small aperture and forms an image of any desired virtual observing plane at the observing screen. Ideally, this lens has to form an image of the pupil of the system under test at the observation plane. The imaging lens with the observation plane can be replaced if desired by the human eye.

A similar arrangement is shown in Figure 10.21. The difference is that a back telecentric system formed by the collimator and a small stop, with the focusing lens, are used in front of the array of light sources (frequently solid state lasers). With this telecentric system, an image of the array of point light sources is formed at the entrance pupil of the system under test. The stop after the lens under test is not used because the small aperture stop at the collimator helps to produce narrow parallel beams of light entering the lens under test. Then, the Hartmann spots are formed at the position sensing detector as in the classic Hartmann test.

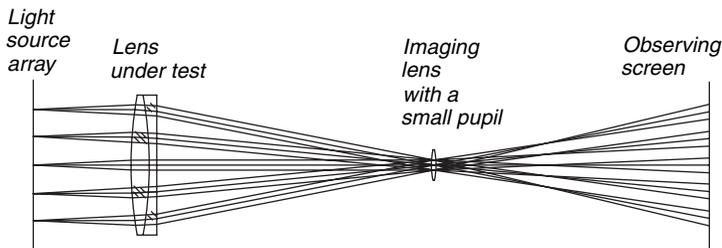


FIGURE 10.20. Testing a convergent lens with an array of light sources in front of the lens. A small stop at the imaging lens selects only the rays that pass through a common focus.

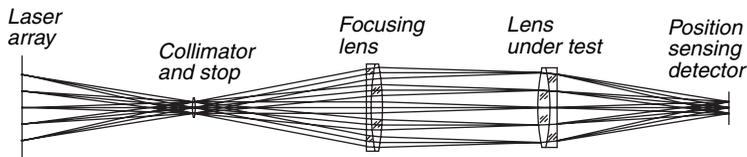


FIGURE 10.21. Testing a convergent lens with an array of point light sources. A bundle of narrow and parallel beams of light illuminate the lens under test.

An advantage of having independent light sources is that they can be switched on and off in sequence. This permits an easy identification of the stops at the observation plane even if no defocus is applied, since only one light source is on at the time of measurement. Another advantage is that the absence of defocus allows the detection and measurement of the convergence or divergence power.

As described by López-Ramírez et al. (2000), instead of using an array of point light sources, a screen with two orthogonal arrays of straight and parallel lines can be placed in front of the lens, as shown in Figure 10.22. The imaging lens and the observation plane can be a human eye. Aspherical lenses can be tested with this method in a simple manner, observing the patterns as shown in Figure 10.23.

10.10.2. Testing Concave and Convex Surfaces

A convex surface or divergent lens can also be tested with this test using an array of light sources. Of special importance is the testing of strongly convex surfaces. The two problems that appear when observing the virtual image of a point light (real or virtual) source in front of the convex surface are (a) that the observed image has a strong astigmatism, thus appearing two focal surfaces, the tangential and the sagittal surfaces and (b) that the two focal surfaces are quite strongly curved. If the object is a virtual plane behind the surface, the shape of the two focal surfaces is close to an ellipsoid, but not exactly, as illustrated in Figure 10.24.

To test a strongly curved surface, as for example the cornea of the human eye, the arrangement as shown in Figure 10.25 can be used. The point light sources can be located at an ellipsoidal surface between the sagittal and tangential surfaces. The virtual images of the point light sources are then formed at a plane behind the surface.

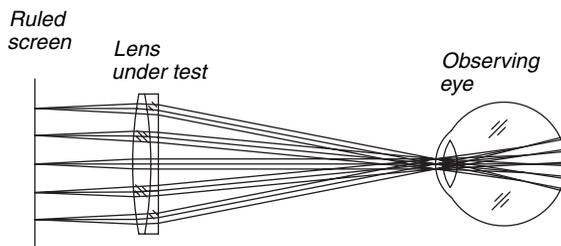


FIGURE 10.22. Testing a convergent lens by observing a ruled screen located at the front of the focal plane of the lens.

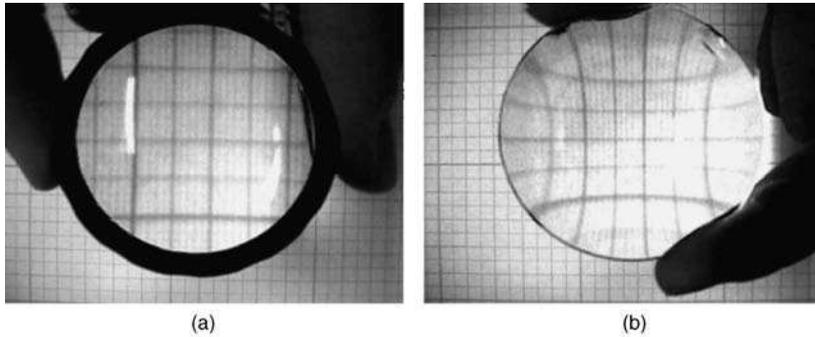


FIGURE 10.23. Observing the ruled screen in front of the lens under test. (a) A lens without spherical aberration (with an aspherical surface) and (b) a lens with spherical aberration.

At a hole close to the ellipsoid, a lens is placed to form the image of the surface under test at the observing screen. The lens with the observing screen can be a digital or a television camera. This ellipsoidal configuration of the surface with the array of light sources has been used by Mejía-Barbosa Y. and D. Malacara-Hernandez (2001) and Mejía-Barbosa (2006) to test convex mirrors.

The ellipsoidal shape is sometimes unpractical, and a simple configuration is desired. Diaz-Uribe and Campos-García (2000) proposed a cylindrical configuration with its axis aligned with the optical axis of the convex surface under test as illustrated in Figure 10.26. In this case, instead of an array of point light sources, a printed screen can be used. The lines to produce an image with a rectangular array are quite distorted as shown in Figure 10.27. Their correct shape is calculated by ray tracing. The image observed from the end of the tube as shown in Figure 10.28 is obtained with a strong convex aspherical surface as shown in Figure 10.29.

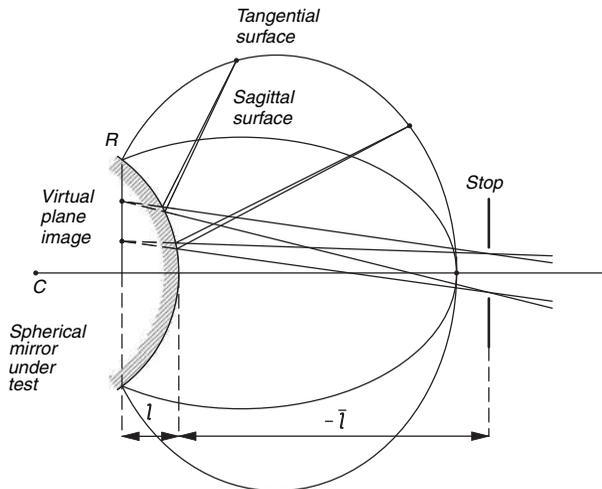


FIGURE 10.24. Tangential and sagittal surfaces produced with a plane virtual object behind a strong convex surface and a stop centered on the optical axis, close to the image surfaces.

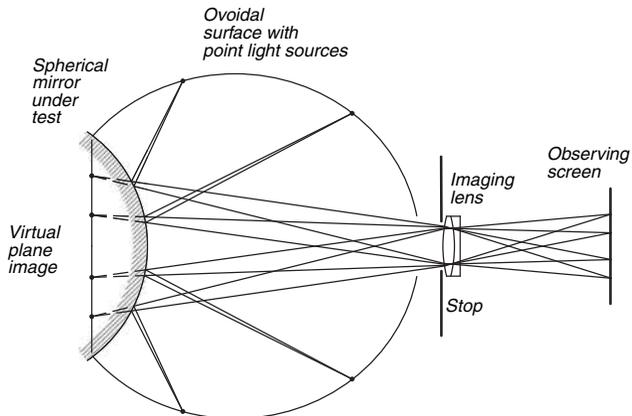


FIGURE 10.25. Testing a strong convex surface with an array of light sources on an ovoidal surface.

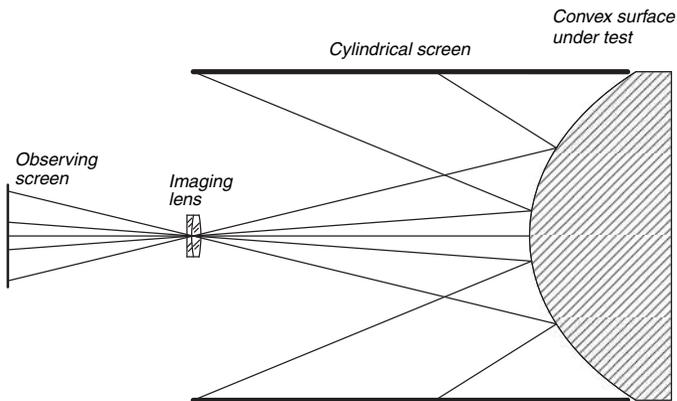


FIGURE 10.26. Null testing of a strong convex surface using a cylindrical ruled screen with its axis being coaxial with the optical axis of the convex surface.

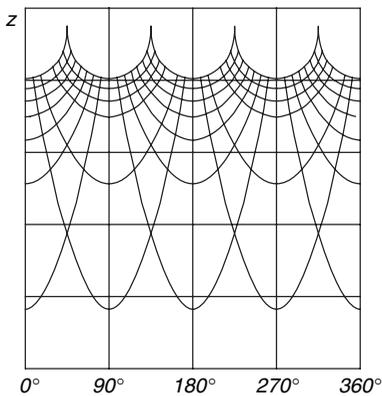


FIGURE 10.27. Ruled screen to be folded as a cylinder used to test a strong convex surface.

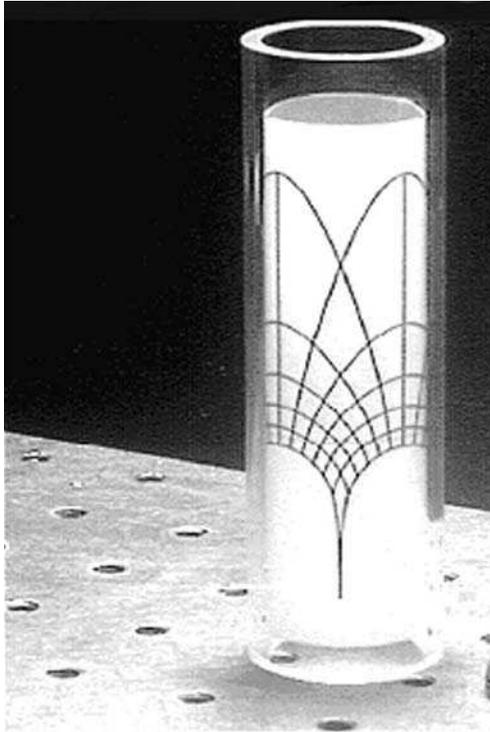


FIGURE 10.28. Plastic cylinder with the ruled screen inside it.

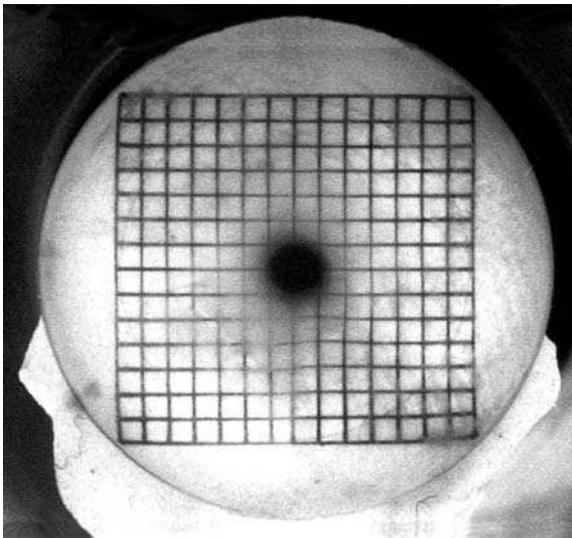


FIGURE 10.29. Pattern observed with the cylindrical ruled screen.

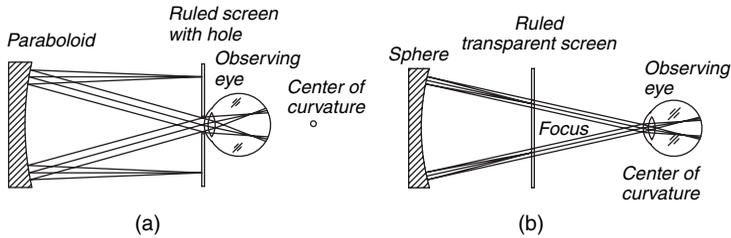


FIGURE 10.30. Testing with a flat ruled screen: (a) A concave paraboloid and (b) a concave sphere.

As shown by López-Ramírez et al. (2000), concave surfaces can also be tested with printed screens as illustrated in Figure 10.30 for a paraboloid and an spherical surface. Off-axis concave paraboloid can be tested in a similar manner with a tilted flat screen as described by Avendaño-Alejo and Díaz-Urbe (2006).

10.11. MICHELSON–GARDNER–BENNETT TESTS

A test was devised by Michelson (1918) to determine mirror surface deformations by the interference of light reflected from the surface illuminated through two holes in a screen. The Michelson setup is shown schematically in Figure 10.31. The light source is a slit illuminated by a monochromatic lamp. The image of the slit is observed by means of a microscope. A series of screens with two apertures or, equivalently, a double slit mechanism is placed in front of the surface under test. One aperture is fixed at the center, and the other changes its radial distance. Interference bands are observed in the microscope, with the central fringe exactly at the position

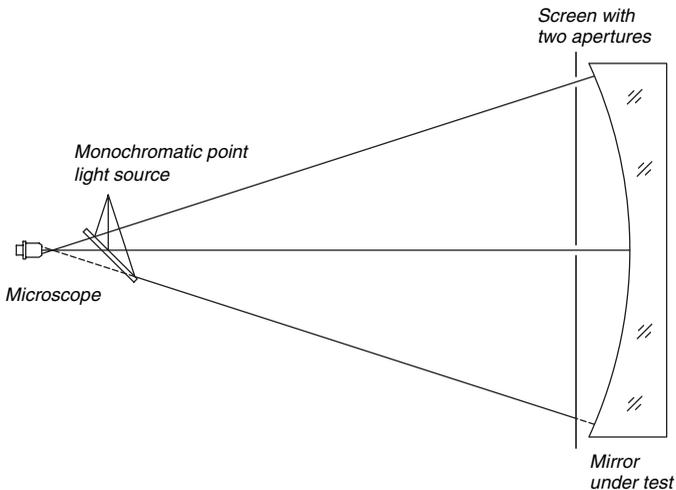


FIGURE 10.31. Michelson test setup. The central screen hole is not moved, whereas the other hole is moved in the survey of the entire surface.

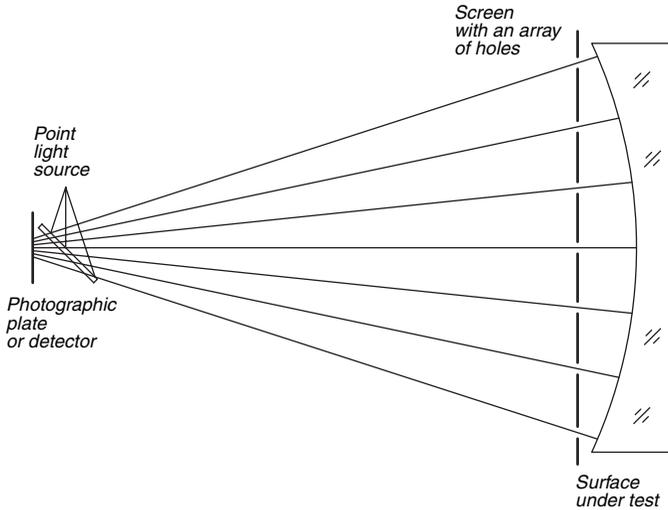


FIGURE 10.32. Gardner–Bennett test setup.

of the slit image if the wavefront is spherical. When the wavefront is not spherical, the distance between the central fringe and the image of the slit gives the error. The fringe displacement, in fractions of the separation between fringes, is the wavefront error. By placing the movable hole over many places on the aperture, a complete mapping of the wavefront errors can be obtained.

Michelson's method was later modified to a diametral screen with several holes (Merland, 1924; Gardner and Bennett, 1925; Bennett and Gardner, 1925) for testing refractive elements and is also applicable to mirrors. The Gardner–Bennett setup is illustrated in Figure 10.32, which shows that the interference fringes between adjacent holes on the screen are recorded, not at the focus, as in the Michelson test, but outside or inside the focus, as in the Hartmann test. The defocusing must be sufficiently small, however, so that the light from adjacent holes does interfere. The deviation of the central fringe from its ideal position gives an indication of the phase difference between the wavefront regions covered by two holes as in the Michelson test.

Although these forms of testing have not gained much popularity, they may be worth a thorough investigation for potential use in uncommon testing situations. A bibliography of publications on applications of various forms of this method is given at the end of this chapter.

10.12. OTHER DEVELOPMENTS

One important development was the modification of the classical test to incorporate an electro-optical detector array instead of the customary photographic plate (Hausler and Schneider, 1988; Pearson, 1990). This permits rapid data collection

and analysis of several spot patterns, and with the aid of a microcomputer these can be averaged. The averaging can be done either with the raw data consisting of spot centroids or with the wavefront maps that result from the analysis of the individual patterns. This approach also permits the dynamical sampling of the aperture (Golden, 1975; Hausler and Schneider, 1988).

An improvement on the accuracy in measuring the spot centroids has been obtained by the intentional overlapping of these spots, thus causing an interference effect to occur (Korhonen, 1983; Korhonen et al., 1986). Such an approach allows a closer packing of the spots and a possible higher surface-sampling frequency. This would make the use of detector arrays more attractive, especially since the detector could be biased to ignore the low-intensity noise present between adjacent spots.

With the use of the electro-optical detector arrays and fast microcomputers, a rapid turnaround in testing can be achieved. Also this combination makes it possible to test the same surface several times, without changing it, and thus makes it possible to reduce test errors by averaging the several test results. Finally, we have seen that continuous work is being done in the application and analysis of the Hartmann test, with very promising innovations being considered.

REFERENCES

- Ahlberg J. H., E. N. Nilson, and J. L. Brown, *The Theory of Splines and Their Applications*, Academic Press, New York, 1967.
- Avedaño-Alejo M. and R. Díaz-Uribe, "Testing a Fast Off-Axis Parabolic Mirror by Using Tilted Null Screens," *Appl. Opt.*, **45**, 2607–2614 (2006).
- Bennett A.H. and I. C. Gardner, "Interference Effects and the Hartmann Test," *J. Opt. Soc. Am.*, **11**, 126 (1925).
- Cornejo A. and D. Malacara, "Wavefront Determination Using Ronchi and Hartmann Tests," *Bol. Inst. Tonantzintla*, **2**, 127 (1976).
- Díaz-Uribe R. and M. Campos-García, "Null-Screen Testing of Fast Convex Aspheric Surfaces," *Appl. Opt.*, **39**, 2670–2677 (2000).
- Freischlad K. and C. Koliopoulos, "Wavefront Reconstruction from Noisy Slope or Difference Data Using the Discrete Fourier Transform," *Proc SPIE.*, **551**, 74 (1985).
- Ftaclas C. and A. Kostinski, "Curvature Sensors, Adaptive Optics, and Neumann Boundary Conditions," *Appl. Opt.*, **40**, 435–438 (2001).
- Gardner I. C. and A. H. Bennett, "A Modified Hartmann Test Based on Interference," *J. Opt. Soc. Am.*, **11**, 441–452 (1925).
- Ghozeil I., "Use of Screen Rotation in Testing Large Mirrors," *Proc. SPIE*, **44**, 247 (1974).
- Ghozeil I. and J. E. Simmons, "Screen Test for Large Mirrors," *Appl. Opt.*, **13**, 1773–1777 (1974).
- Golden L. J., "Dynamic Hartmann Test," *Appl. Opt.*, **14**, 2391–2399 (1975).
- Greivenkamp J. E., D. G. Smith, R. O. Gappinger, and G. A. Williby, "Optical Testing Using Shack-Hartmann Wavefront Sensors," *Proc. SPIE*, **4416**, 260–263 (2001).
- Hartmann J., "Bemerkungen über den Bau und die Justirung von Spektrographen," *Zt. Instrumentenk.*, **20**, 47 (1900).

- Hartmann J., “Objektivuntersuchungen,” *Zt. Instrumentenk.*, **24**, 1 (1904a).
- Hartmann J., “Objektivuntersuchungen,” *Zt. Instrumentenk.*, **24**, 33 (1904b).
- Hartmann J., “Objektivuntersuchungen,” *Zt. Instrumentenk.*, **24**, 97 (1904c).
- Hausler G and G. Schneider, “Testing Optics by Experimental Ray Tracing with a Lateral Effect Photodiode,” *Appl. Opt.*, **27**, 5160–5164 (1988).
- Howland B., “Use of Crossed Cylinder Lens in Photographic Lens Evaluation,” *Appl. Opt.*, **7**, 1587–1599 (1968).
- Kingslake R., “The Absolute Hartmann Test,” *Trans. Opt. Soc.*, **29**, 133 (1927–1928).
- Korhonen T. K., “Interferometric Method for Optical Testing and Wavefront Error Sensing,” *Proc. SPIE*, **444**, 249 (1983).
- Korhonen T. K., S. T. Haarala, Lo. Piironen, and A. K. Sillanpaa, “Interferometric Optical Test and Diffraction Based Image Analysis,” *Proc. SPIE*, **628**, 486–491 (1986).
- Landgrave J. E. A. and J. R. Moya, “Effect of a Small Centering Error of the Hartmann Screen on the Computed Wavefront Aberration,” *Appl. Opt.*, **25**, 533–536 (1986).
- Liang J., B. Grimm, S. Goelz, and J. F. Bille, “Objective Measurements of the Wave Aberrations of the Human Eye with the Use of the Hartmann–Shack Wave-Front Sensor,” *J. Opt. Soc. Am.*, **11**, 1949–1957 (1994).
- Loibl B., “Hartmann Tests on Large Telescopes Carried out with a Small Screen in a Pupil Image,” *Astron. Astrophys. (Germany)*, **91**, 265 (1980).
- López-Ramírez J. M., D. Malacara-Doblado, and D. Malacara-Hernández, “New Simple Geometrical Test for Aspheric Lenses and Mirrors,” *Opt. Eng.*, **39**, 2143 (2000).
- Malacara D., “Geometrical Ronchi Test of Aspherical Mirrors,” *Appl. Opt.*, **4**, 1371 (1965).
- Malacara D. and Z. Malacara, “Testing and Centering by Means of a Hartmann Test with four Holes,” *Opt. Eng.*, **31**, 1551–1555 (1992).
- Malvick A. J., “Theoretical Elastic Deformation of the Steward Observatory 230 cm and the Optical Sciences Center 154-cm Mirrors,” *Appl. Opt.*, **11**, 575 (1972).
- Mayall N. U. and S. Vasilevskis, “Quantitative Tests of the Lick Observatory 120-Inch Mirror,” *Astron. J.*, **65**, 304 (1960); reprinted in *Lick Obs. Bull.* No. 567, 1960
- Mejía-Barbosa Y. and D. Malacara-Hernández, “Object Surface for Applying a Modified Hartmann Test to Measure Corneal Topography,” *Appl. Opt.*, **40**, 5778–5786 (2001).
- Mejía-Barbosa Y., “Hartmann Test of Small F/# Convex Mirrors,” *Opt. Commun.* **263**, 17–24 (2006).
- Merland M. A., “Sur La Methode de M. M. Michelson et Cotton pour l’Etude des Systemes Optiques,” *Rev Opt.*, **3**, 401 (1924).
- Michelson A. A., “On the Correction of Optical Surfaces,” *Astrophys. J.*, **47**, 283 (1918).
- Morales A. and D. Malacara, “Geometrical Parameters in the Hartmann Test of Aspherical Mirrors,” *Appl. Opt.*, **22**, 3957–3959 (1983).
- Neal D. R., J. Copland, and D. Neal, “Shack-Hartmann Wavefront Sensor Precision and Accuracy,” *Proc. SPIE*, **4779**, 148–160 (2002).
- Pearson E. T., “Hartmann Test Data Reduction,” *Proc. SPIE*, **1236**, 628 (1990).
- Platt B. C. and R. V. Shack, “Lenticular Hartmann Screen,” *Opt. Sci. Newsl.*, **5**, 15–16 (1971).
- Salas-Peimbert D. P., G. Trujillo-Schiaffino, D. Malacara-Hernández, D. Malacara Doblado, S. Almazán-Cuellar, “Ophthalmic Lenses Measurement Using Hartmann Test,” *Proc. SPIE*, **5622**, 102–106 (2004).

- Salas-Peimbert D. P., D. Malacara Doblado, V. M. Durán-Ramírez, G. Trujillo-Schiaffino, and D. Malacara-Hernández, “Wave-Front Retrieval from Hartmann Test Data,” *Appl. Opt.*, **44**, 4228–4238 (2005).
- Schulte D. H., “A Hartmann Test Reduction Program,” *Appl. Opt.*, **7**, 119–124 (1968).
- Servín M., D. Malacara, J. L. Marroquín, and F. J. Cuevas, “New Technique for Ray Aberration Detection in Hartmanngrams Based on Regularized Bandpass Filters,” *Opt. Eng.*, **35**, 1677–1683 (1996).
- Servín M., F. J. Cuevas, D. Malacara, and J. L. Marroquín, “Direct Ray Estimation in Hartmanngrams by Use of a Regularized Phase-Tracking System,” *Appl. Opt.*, **38**, 2868–2869 (1999).
- Simmons J. E. and I. Ghozeil, “Double-Option Technique for Testing Large Astronomical Mirrors” (abstract only), *J. Opt. Soc. Am.*, **61**, 1586 (1971).
- Southwell W. H., “Wave-Front Estimation from Wave-Front Slope Measurements,” *J. Opt. Soc. Am.*, **70**, 998 (1980).
- Stavroudis O. N. and L. E. Sutton, *Spot Diagrams for the Prediction of Lens Performance from Design Data*, U.S. Department of Commerce, National Bureau of Standards Monograph No. 93, Washington, D.C., 1965.
- Tscherning M., Die monochrotischen Aberrationen des menschlichen Auges, *Z. Psychol. Physiol. Simm.* **6**, 456–471 (1894).
- Vitrichenko E. A., F. K. Katagarov, and B. G. Lipovetskaya, “Methods of Investigation of Astronomical Optics. II: Hartmann Method,” *Izv. Spetz. Astrofis. Obs.*, **7**, 167 (1975).

11

Star Tests

D. Malacara and W. T. Welford[†]

11.1. INTRODUCTION

This chapter has been rewritten and updated by the first author, largely based on the previous version by late Prof. W. T. Welford. The star test is conceptually, perhaps, the most basic and simplest of all methods of testing the image-forming optical systems: We examine the image of a point source formed by the system and judge the image quality according to the departure from the ideal image form. In principle, the test can be made quantitative by, for example, photoelectric measurement techniques, but in practice the star test in the workshop is almost always carried out visually and semiquantitatively. We shall mainly discuss visual techniques. We can divide star testing methods into two groups: (a) those in which very small aberrations, near or below the Strehl tolerance limit, are examined and (b) those in which relatively large aberrations are studied. Group (a) is typified by tests on microscope and telescope objectives and group (b) by tests on camera lenses in which, for example, the star test is used to plot the astigmatic field surfaces or to estimate transverse chromatic aberration.

The monochromatic image of a point source, or point spread function as it is often called, has a very complex structure, particularly in the presence of aberrations. The structure depends in a complicated way on the geometrical aberrations, but it would be out of place in a book on practical methods to go deeply into this aspect. Moreover, it is always possible in principle to calculate the point spread function from the aberrations, although the calculation may in practice be costly in computer time; but for the present purposes we should like to be able to estimate the aberrations from the form of the point spread function. This, however, is generally impossible in principle. If the aberration is axially symmetrical, this calculation could be done from very careful measurements of the light intensity in the star image. We have to make estimates based on experience and on the many examples of point spread functions that have been computed and photographed from known aberrations. Thus, the star test is semiquantitative, and considerable experience is needed to get the best results from it. Nevertheless, it is an important testing technique because it is rapid and

reliable in experienced hands and is very sensitive. For example, the test is used in the final adjustment of a critical air space in microscopic objectives of high numerical aperture, since in this process it is desirable to try a new spacing and to retest rapidly.

There are not many publications about the star test probably because it depends so much on experience. The basic reference is still the work of Taylor (1891). Useful descriptions of the method have also been given by Martin (1961) and by Twyman (1942); the latter reproduces much of Taylor's descriptive material.

In discussing the star test we shall mostly assume that the pupil of the system under test is filled with light of uniform intensity, and we shall also assume that the star image is formed at a reasonably large convergence angle, say, not less than 0.01. Modest variations of intensity across the pupil generally do not affect the star image much, but we make an exception in the case of Gaussian beams from lasers. (A true Gaussian beam has an indefinitely large aperture, while at the same time the theory by which the Gaussian beam profile is derived uses the paraxial approximation! This apparently blatant contradiction is resolved by the fact that, in practice, the intensity in the Gaussian profile is negligibly small at distances from the central maximum greater than, say, three times the $1/e^2$ radius.) Gaussian beams will be discussed in Section 11.2.6, and the effects of very small convergence angles will be mentioned in Section 11.2.7. At very large convergence angles, say, numerical aperture greater than 0.5, the shape of the point spread function is again different in detail from the one to be described here, and an excellent study of such effects was recently published by Stamnes (1986); for reasons of technical convenience, it is very unlikely that a practical star test would be carried out at such large numerical aperture (NA); the usual procedure is to transform the NA by a reliably aberration-free system such as a well-corrected microscope objective, which itself has been star tested. Thus, for our main descriptions of the aberration free point spread function (Figs. 11.1–11.9), we have used the usual scalar wave approximation as the basis of the calculations.

11.2. STAR TEST WITH SMALL ABERRATIONS

In star testing systems that are nearly diffraction limited (possibly apart from chromatic aberration), we need a background of information about the appearance of the point spread function with small aberrations. This is given in Sections 11.2.1–11.2.5, but without mathematical details.

Photographs and computations of point spread functions are to be found throughout the literature on applied optics, although often not in a form particularly applicable to practical star testing. One of the most useful sets of photographs is that which is used as a frontispiece by Taylor (1891) and subsequently reproduced in many other publications, for example, Martin (1961). Among other photographs, we may mention the work of Nienhuis (1948), which was reproduced by, for example, Born and Wolf (1975); however, we should note that Nienhuis took some of his very beautiful spread function photographs with a coherent background in order to enhance the secondary rings and fringes, and this rather falsifies them for the purposes of star testing. In fact all photographs are to be mistrusted for star testing: The nonlinearity of the photographic emulsion is compounded with that of the halftone process and the result cannot be

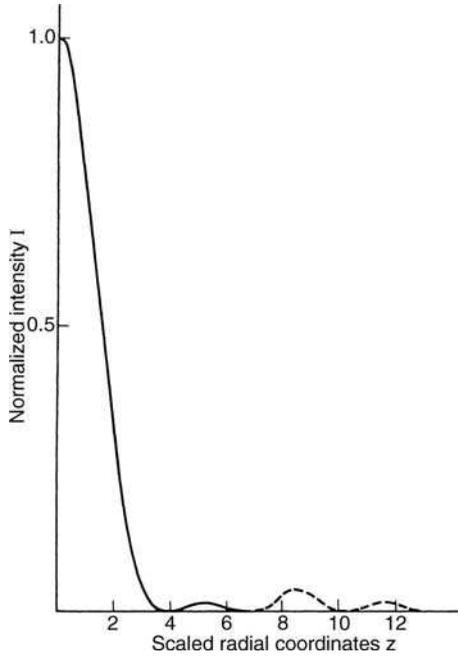


FIGURE 11.1. The Airy pattern; the aberration-free image of a bright point formed in monochromatic light by a system with a circular aperture. The function plotted is

$$I = \left[\frac{2J_1(z)}{z} \right]^2$$

where $z = (2\pi/\lambda) \sin \alpha \eta$, $\sin \alpha$ is the convergence angle of the image-forming beam, and η is the radial distance from the center of the aperture. The second and third rings are plotted at 10 times the actual ordinate.

relied on over any reasonable dynamic range. We therefore give only graphical presentations of spread functions in this chapter.

11.2.1. The Aberration Free Airy Pattern

Figure 11.1 shows the Airy pattern, the monochromatic aberration-free point spread function for a system with a circular aperture of uniform transmission, and Figure 11.2 shows the same with logarithmic vertical scale. What is plotted in these two figures is the light intensity (vertical) against the radial distance from the center of the image (horizontal). The light intensity is scaled to unity at the center of the pattern. The radial coordinate z , which lies in the image plane, is given by

$$z = \frac{2\pi}{\lambda} \sin \alpha \eta \quad (11.1)$$

where λ is the wavelength of the light, α is the convergence angle (cone semiangle) of the image-forming beam, and η is the actual radial distance. Thus the radial coordinate as shown is dimensionless and is often said to be in z units or diffraction

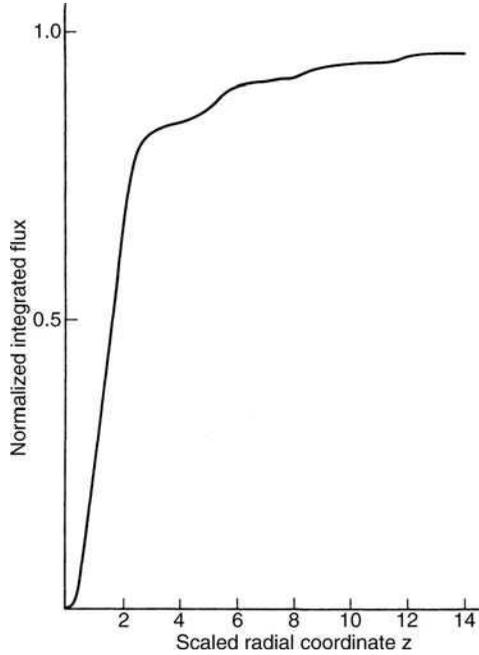


FIGURE 11.3. The “Encircled Energy,” that is, the proportion of total light flux inside a circle of radius z in the image plane. See Table 11.1.

encircled flux function, which gives the proportion of the total flux within a given radius from the center. The encircled flux (sometimes called the encircled energy) has occasionally been proposed for use in the development of the star test (see, e.g., Barakat and Newman, 1963), but probably the practical difficulties are too great for it to be generally adopted. However, the function is useful as a guide to the general properties of the Airy pattern. This function is plotted in Figure 11.3.

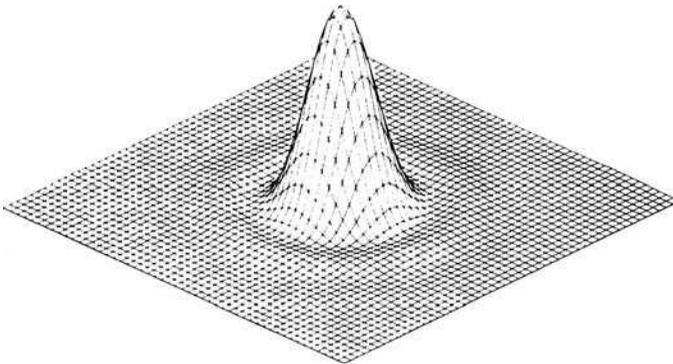


FIGURE 11.4. The diffraction solid for the Airy pattern; 2.5 squares in the horizontal plane correspond to a z unit. (Computed and plotted by M. W. L. Wheeler.)

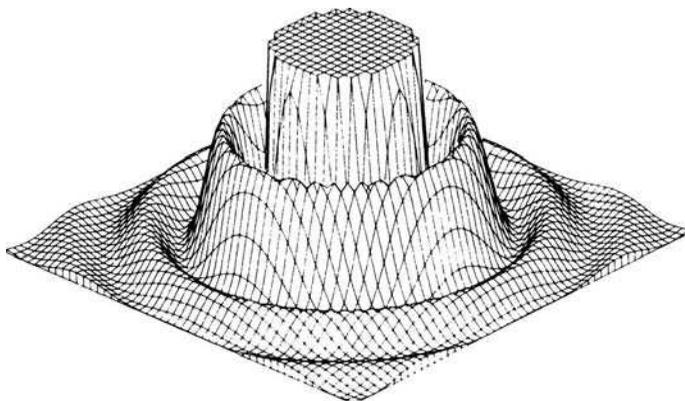


FIGURE 11.5. Similar to Figure 11.4, but with increased vertical scale truncated to relative intensity 0.03. (Computed and plotted by M. W. L. Wheeler.)

Figure 11.4 shows an alternative way of presenting the Airy pattern; the vertical height of the figure represents the relative intensity. Such a figure is sometimes called a diffraction solid. Figure 11.5 represents the same solid with increased vertical scale to show the details of the outer rings.

11.2.2. The Defocused Airy Pattern

One of the earliest findings of practical experience in star testing is that it is very useful to examine the defocused point spread function on both sides of the best focal plane. This may seem paradoxical, but it is found to be much easier to diagnose the aberrations from the defocused image than from the image at best focus. We therefore show in Figure 11.6 the contours of constant intensity (isophots) in a plane containing the optical axis. The axes correspond to a radial coordinate in the image plane (vertical axis) and to the defocus distance along the principal way. The image plane coordinate is the same dimensionless coordinate as was used in Figures 11.1 and 11.2 and the defocus coordinate, explained in the caption, is also dimensionless. The scales are such that the geometrical cone of rays would have a semiangle of 45° , but by using appropriate values of the convergence angle x and the wavelength λ , the diagram can represent any cone[†]

Several features of this distribution are worth noting. It is symmetrical about the true focal plane. There are zeros (“dark spots”) along the axis, equally spaced (apart from one missing at the focal plane) and alternating with maxima. The intensity along the axis is shown in Figure 11.7. In Figure 11.6, the out-of-focus patterns consist of bright

[†]It should be noted, however, that this applies only to the case in which the diameter of the first dark ring is much less than the diameter of the exit pupil of the optical system. This situation would, of course, almost invariably exist in most workshop testing situations, but occasionally it may happen that the convergence angle is very small and then Fig. 11.6 would not be a true picture. This happens more frequently in beams with Gaussian profiles, as produced by a single-mode TEM₀₀ laser, since these are often used nearly collimated.

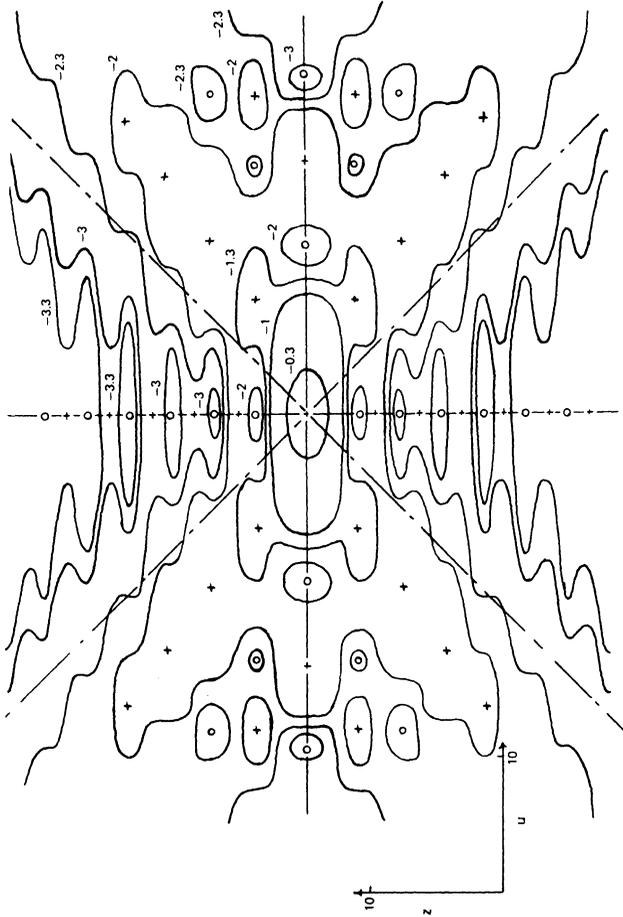


FIGURE 11.6. Isophots (contours of constant intensity) in the defocused Airy pattern. The line $z = u$ is the boundary of the ray cone. The circles denote minima (zeros on the axes); the crosses, maxima. The contours are labeled with the logarithm of the intensity. (From values computed by J. C. Dainty.)

and dark rings, spreading out with increasing defocus. This cannot be seen very clearly in Figure 11.6 but is more evident in Figure 11.8, which is again a set of isophots, but taken over a greater range of both the coordinates and showing only one quadrant of the symmetrical pattern. The isophots are labeled with the logarithm, to base 10, of the relative intensity, so as to give a greater range of intensities in the diagram. From Figure 11.8 it can be seen that a typical defocused pattern, easily recognizable by anyone who has looked at point spread functions, will show a series of roughly equally spaced rings, increasing in intensity toward the outside and having the outermost ring brighter, wider, and larger in diameter than would be expected from the general progression of the rings. In fact, the rings are not so regularly graded as they appear at first to be; this can be seen from a very careful inspection of the monochromatic point spread function or from a study of Figure 11.8. Nevertheless, this apparent regularity is easily upset by

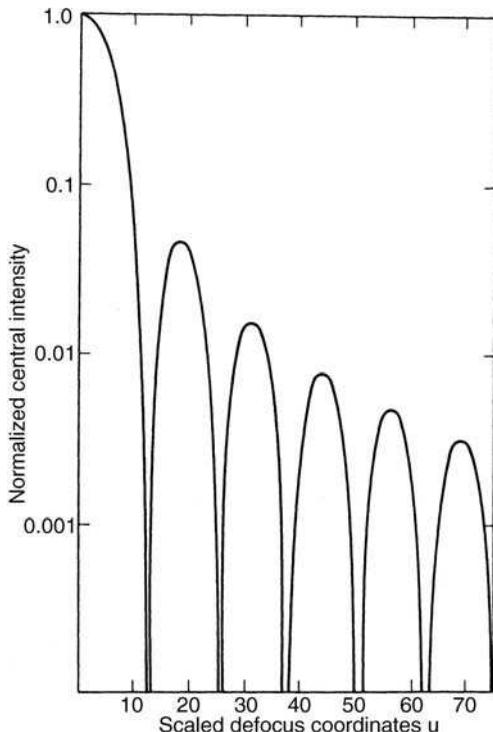


FIGURE 11.7. The intensity at the center of the defocused Airy pattern. The quantity plotted is $\sin \frac{1}{4}u / \frac{1}{4}u$ where $u = (2\pi/\lambda)\zeta \sin^2 \alpha$, α is the convergence angle of the beam, and ζ is the defocus distance.

small aberrations, and this is a very useful property of the star test. As the defocus is steadily increased, the appearance is as if new rings appear from the center and spread out, taking the rest of the pattern with them like the ripples in a pond.

Some of these effects and those described in Section 11.1.5 are illustrated by photographs taken by Cagnet et al. (1962). Beiser (1966) gave a perspective sketch of the intensity distribution through focus, and Taylor and Thompson (1958) reported careful measurements to verify the predicted distributions. Figure 11.9 shows the inside of focus, in focus and out of focus images of an aberration free system.

11.2.3. Polychromatic Light

The descriptions in Sections 11.2.1 and 11.2.2 apply, of course, to monochromatic light. All features of the point spread function (diameter of dark rings, distances between axial zeros, etc.) scale in direct proportion to the wavelength. Thus, even if an optical system has no chromatic aberration according to geometrical optics (i.e., a purely reflecting system), we would still expect to see color effects when star testing; however, in practice, these are not very noticeable. We can, therefore, use white light for testing most of the reflecting systems without danger of confusion from the weak chromatic effects.

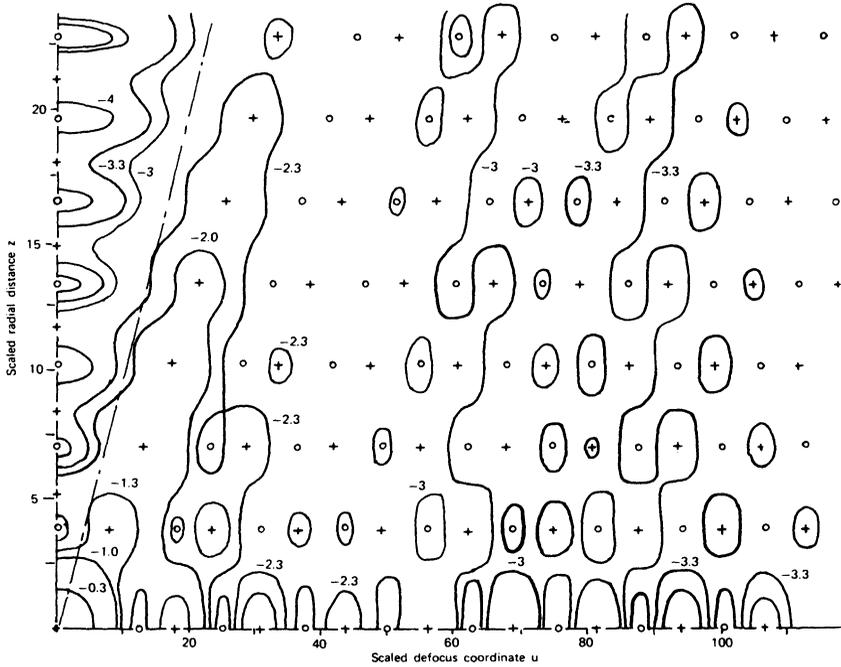


FIGURE 11.8. Isophots in the defocused Airy pattern; this diagram covers a greater of defocus than Figure 11.6. The chain line ($z = u$) denotes the boundary of the cone of rays forming the pattern. The circles denote minima (zeros on the axes); the crosses, maxima. The contours are labeled with the logarithm of the intensity, scaled to unity at the origin. (From values computed by J. C. Dainty.)

In refracting systems, however, the star test easily reveals “secondary spectrum,” that is, the residual uncorrectable chromatic aberration. We are not always concerned to actually test chromatic correction by star testing because achromatism depends on overall choices of glass types and powers of components and is therefore affected only by relatively gross errors in construction that would affect other aberrations more markedly. Thus, in testing refracting systems we should use either a fairly

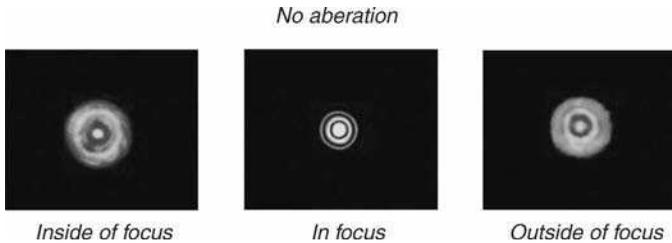


FIGURE 11.9. Inside of focus, in focus, and outside of focus images for an aberration free system.

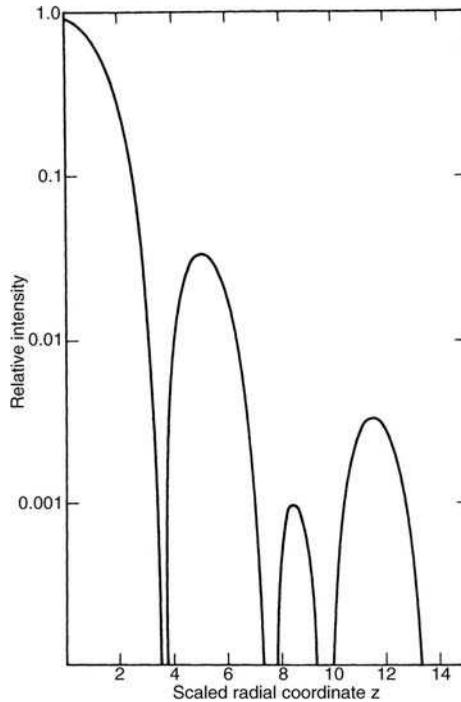


FIGURE 11.10. The in-focus point spread function for a circular aperture with a central obstruction which is one quarter of the full diameter. This should be compared with Figure 11.2 to see the effect on the central intensity and the outer rings.

narrow wavelength band, selected by means of a filter (Section 11.3.2.) or a laser. In this way the monochromatic aberrations can be more clearly seen; this is usually the object in workshop tests, since it is these aberrations that are susceptible to correction by refiguring or by tuning air spaces.

Linfort and Wolf (1952) calculated the total flux in polychromatic star images from refracting telescopes, weighted according to the visibility curve of the human eye. These results are of great interest for optical designers, particularly, since the calculations were done for existing telescope objectives, but they are probably not very useful as a background to the star test, since in actual operation the varying colors seen probably make at least as much impression as the total brightness. However, they do show how the interesting color changes that are observed arise from secondary spectrum effects.

11.2.4. Systems with Central Obstructions

Most large telescopes and many other optical systems have a central “hole” in the pupil, due to a secondary mirror. Some of the effects of this central obstruction on the point spread function are rather surprising. There are three main effects: (a) The

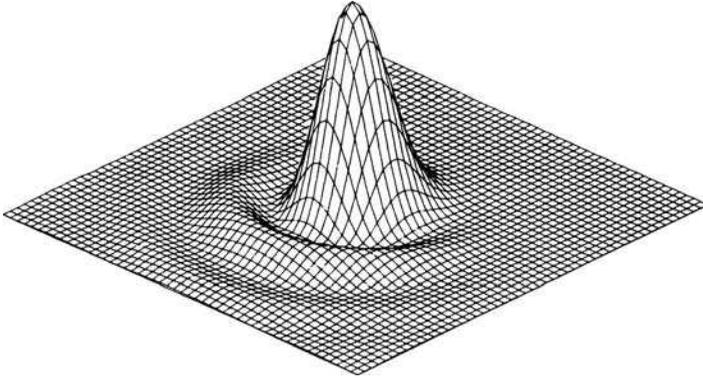


FIGURE 11.11. The diffraction solid for a 0.6 of a wavelength of coma (the Strehl tolerance limit). (Computed and plotted by M. W. L. Wheeler.)

central maximum becomes narrower; (b) the outer rings become irregular in brightness—some brighter, some dimmer; and (c) the spacing along the axis of the dark spots increases. A detailed theoretical study made by Linfoot and Wolf (1953) showed that, if the obstruction ratio (diameter, not area) is E , the axial spreading of the features is according to the factor $(1 - \epsilon^2)^{-1}$. Figure 11.10 shows the in-focus pattern for an obstruction ratio $\epsilon = 0.25$, which is typical for an astronomical telescope. Features (a) and (b) may be seen in comparison with those in Figure 11.2. Linfoot and Wolf (1953) give isophot diagrams showing the out-of-focus effects.

11.2.5. Effects of Small Aberrations

In an optical design context, “small aberration” usually means a distortion of the wavefront from the ideal spherical shape of the order of magnitude $\lambda/4$. This is so because the system of aberration tolerances based on the Strehl criterion leads to aberration tolerances of this order. It is generally accepted that, when the aberrations are at this tolerance limit, as described later, the effect on the formation of images of extended objects is negligible, and the optical system is said to be diffraction limited[†]

The Strehl aberration tolerance system is explained by, for example, Born and Wolf (1975). It depends on the concept that the initial effect of introducing a small amount of aberration of any kind into the wavefront forming the point spread function is to reduce the maximum intensity slightly: The width at half maximum of the central maximum stays unchanged, and the flux taken from the center is redistributed to the outer rings. A 20% drop in the maximum intensity is generally taken as the tolerance limit, and simple formulas are given (e.g., by Born and Wolf 1975) for the amounts of different aberrations that will produce this tolerance limit.

[†] Sometimes the term “Rayleigh limit” is used, since Lord Rayleigh proposed the quarter-wave-length criterion.

However, the eye is not very good at judging changes in the absolute intensity levels; so the Strehl criterion is not very useful in workshop practice, although it is invaluable in optical design. It is also found that, as mentioned earlier, one can see much less than in the Strehl limit in many cases. This can be understood by looking at the photographs of aberrated point spread functions in some of the references already cited; to choose an example at random. Born and Wolf (1975) reproduce in their Figure 9.8 a photograph of coma at about the Strehl tolerance limit, taken by Nienhuis (1948), and it is quite easy to see that this is an aberrated point spread function. The same point is made in Figure 11.11, which shows the diffraction solid for coma at the Strehl limit. There seems to be no doubt, on the basis of experience, that the semiquantitative methods beautifully described by Taylor (1891) are extremely sensitive; however, it is difficult to set a limit to the sensitivity since this depends on the kind of aberration, for example, whether the wavefront shape varies rapidly or slowly across the aperture. In a study of spherical aberration, Welford (1960) suggested that it is possible to detect $\lambda/20$ of slowly varying aberration and $\lambda/60$ of rapidly varying aberration.

11.2.6. Gaussian Beams

By convention, the equivalent of a focus for a Gaussian beam is called the beam waist; the beam profile is the same at all points along the beam, simply scaling the intensity and width depending on the distance from the beam waist. There are well-known formulas relating the width of the beam to the distance from the waist and giving the curvature of the phase front (e.g., Kogelnik and Li, 1966). If a system is to be used with Gaussian beams, then by implication the lenses and other components must have enough spare aperture to ensure negligible truncation of the Gaussian profile; the equivalent of the star test would then be to check that the final beam waist has a true Gaussian profile; but since the defocused waist does not change qualitatively in a Gaussian beam, this is not very easy to assess. In practice, it is probably better to expand the input beam and test the system with fully filled aperture (so-called hard-edge beam) and with more or less uniform amplitude; this can be an overrigorous test, since the outer parts of the aperture have less amplitude with a Gaussian beam. However, unless the Gaussian beam very much underfills the aperture, the difference is negligible. If a Gaussian beam is used in the test, there is the danger that any slight truncation of the beam at some point in the system will introduce ringing or fringing in the phase front; it is then not easy to tell if this is due to truncation or aberrations in the optical system.

11.2.7. Very Small Convergence Angles (Low Fresnel Numbers)

When a hard-edged beam is almost collimated, so that the focus is formed with a beam of very small convergence angle, the diffraction pattern in the focal region is qualitatively different from that described in Section 11.2. This is discussed in some detail by Stamnes (1986) and by Li and Wolf (1984). From the point of view of star testing, the most marked effect is that for an aberration free beam, the intensity

distribution is no longer symmetrical on either side of the focus; this would considerably complicate the task of assessing the quality of the system, so that it is advisable to avoid testing in a nearly collimated space.

This raises the question, how nearly collimated may we be without upsetting the simpler descriptions of the star image in Section 11.2. This is best answered in terms of the Fresnel number of the beam: Let α be the radius of the exit pupil and let R be the radius of curvature of the emerging phase front, that is, R is the distance from the pupil to the focus; the Fresnel number N is defined as

$$N = \frac{\alpha^2}{\lambda R} \quad (11.2)$$

The significance of the Fresnel number is that $N/2$ is the number of whole wavelength depth of the curvature of the phase front at the edge of the pupil, and it is shown in the references quoted earlier that the asymmetric effects in the focal region occur when N is not very large, say, less than 10. Thus, star testing should always be done in a space where the Fresnel number is not too small according to this criterion.

11.3. PRACTICAL ASPECTS WITH SMALL ABERRATIONS

We now discuss in more detail the effects to be looked for in visual star testing for small aberrations. In what follows we assume that approximately monochromatic light is used.

11.3.1. Effects of Visual Star Testing

Asymmetry of the in-focus pattern denotes a coma-like aberration. This may be either a genuine coma due to working off-axis or a “manufacturer’s aberration” due to poor centering of components (often quite strong asymmetric color effects can be seen because of bad centering), to nonsymmetrical polishing of surfaces, or even to nonuniform distribution of the refractive index in a lens component. A rare source of coma-like effects is an asymmetrical phase change on reflection at a mirror surface. This can be caused by uneven coating of multilayer dielectrics; but if a dielectric-coated mirror is used to fold a beam at a large angle and if the beam has a large convergence angle, enough asymmetrical phase change may result in giving coma-like effects, due to varying angles of incidence across the beam.

In practice, asymmetries always show up more clearly in the defocused image. With coma, in particular, there is a very marked effect of the out-of-focus asymmetry, which is useful in finding the axis or center of the field of a system; this is what is meant by the term “squaring-on,” applied to telescope objectives for astronomical use.

Small amounts of astigmatism show up as a “Maltese cross” effect in focus, and again there is a considerable gain in defocusing. In particular, on moving steadily through focus from one side to the other, the effect of a switch in the direction of the astigmatic focal lines is easily detected for much less than the Strehl limit.

Of course, neither astigmatism nor coma should be present in the axial point spread function of a well-constructed system, and the most important thing to be able to judge is the state of spherical aberration correction. Here almost nothing can be obtained from studying only the in-focus image (but this is not so for large aberrations; see Section 11.4.1), and it is essential to use defocus.

In describing the effects we use the generally accepted terms *undercorrection* and *overcorrection*; undercorrection means that for primary spherical aberration the rays from the edge of the pupil focus nearer to the optical system than the paraxial rays, if a real image is formed, and vice versa for overcorrection. In terms of wavefront shapes, a zone of the aperture has undercorrection if the wavefront in that zone is in advance of the reference sphere (ideal or unaberrated wavefront), and again, vice versa for overcorrection.

With primary spherical aberration or with spherical aberration of any single higher order, we find that undercorrection produces sharper, well-defined rings on the side of focus nearer to the optical system (“inside of focus”) with a particularly bright outermost ring, whereas with “outside of focus,” all rings are blurred and show less contrast and the outer ring, in particular, fades to a vague smudge with poor definition. For overcorrection the same appearances are seen, but with “inside of focus” and “outside of focus” interchanged. Figure 11.12 shows diagrammatically the case of undercorrection in relation to the ray diagram; it can be seen that the much brighter outermost ring occurs on the side where the ray caustic is formed. This is a convenient way to remember how to relate these effects, since we would naturally expect a great concentration of light flux near the caustic. In other words, we see more light intensity near a greater concentration of rays. This rule is generally true, provided that we are far enough from the main focus; it offers a simple way of interpreting zonal aberration effects. Thus suppose that we go a considerable distance

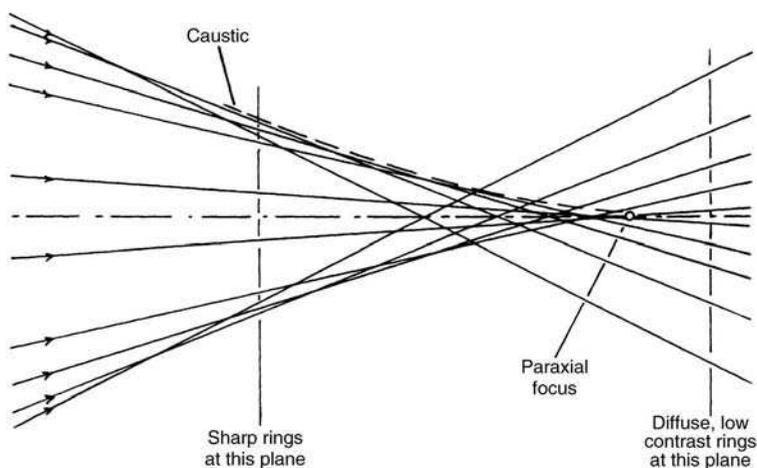


FIGURE 11.12. Zonal spherical aberration: this ray-theoretic interpretation shows approximately how effects due to the zone would appear.

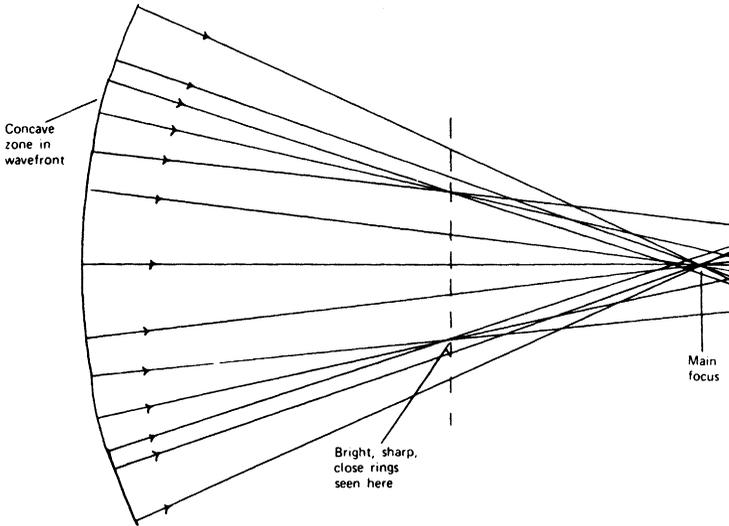


FIGURE 11.13. The effect of zonal wavefront error in bunching the rays.

inside focus and we see that at, say, two thirds of the diameter of the spread-out pattern, the rings are brighter and sharper; this indicates that the rays from the corresponding part of the aperture are bunching together here or, alternatively, that the corresponding zone of the wavefront is more concave than it should be. The opposite effect of fainter and more diffuse rings would be seen outside focus. Figure 11.13 shows the effect of a zonal wavefront error in bunching the rays.

11.3.2. The Light Source for Star Testing

In practice, we never use a real star as a test object for star testing because atmospheric turbulence causes the star image to vary in intensity, position, and aberrations to the extent that a critical appraisal is impossible.

The laboratory artificial star is a pinhole with a lamp of the right spectral composition. The pinhole must be small enough to be quite unresolved, that is, its angular subtense at the objective must be much less than λ/D , where D is the diameter of the objective aperture. In practice, it is easy to check that the pinhole is of the correct size by noticing whether any trace of the pinhole edges can be seen in the star image, since most pinholes are irregular enough in outline to show an image; if only diffraction structure can be seen, the pinhole is small enough.

Very costly pinholes are available in special mounts for use as spatial filters in laser beams. For workshop use, it is cheaper and simpler to employ electron microscope aperture stops. These come as thin, electroformed copper disks about 3 mm in diameter with holes from about 5 to 50 μm in diameter, as ordered; the holes are not always very accurately circular in shape, but this does not matter for testing purposes.

Before electron microscope apertures became generally available, several other devices were used, as described in the references cited at the beginning of this

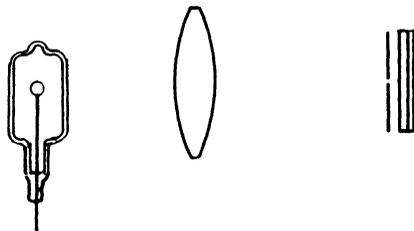


FIGURE 11.14. Lamp and pinhole assembly. The drawing shows a tungsten-halogen lamp, which should be on a centering and focusing mount, followed by a condenser, pinhole, and filter. The condenser should focus the filament onto the pinhole.

chapter. Among these was the reflection of a distant light source in a small, clean globule of mercury, obtained by condensing mercury vapor onto a microscope slide. (In particular, this has been suggested for star-testing microscope objectives.) It is worth mentioning in connection with workshop practice that mercury is a cumulative poison with appreciable amounts that can be absorbed from mercury vapor at room temperature and pressure. This method should therefore never be used.

An intense light source, such as a quartz-halogen lamp or a high-pressure mercury lamp focused onto the pinhole, is desirable in order to be able to pick up the star image easily. A filter can be placed immediately after the pinhole as shown in Figure 11.14. It is usually not advisable to use a laser as a source for star testing. The main objection is that the coherence length of a helium-neon laser is long enough to produce interference effects between many beams multiply reflected off the surfaces of a refracting system. These effects can be bright enough to obscure the details of the star test, particularly since the workshop testing would probably be done before any antireflection coating was put on the surfaces. Another objection to the use of a laser, for refracting systems at any rate, is that the system may be corrected for a different wavelength and the spherical aberration correction vary appreciably with wavelength. However, for testing large mirror systems a laser is probably the best source, since it provides enough intensity to work in daylight.

11.3.3. The Arrangement of the Optical System for Star Testing

For testing small refracting objectives the simplest arrangement is the one shown in Figure 11.15, with the star at a suitable distance. What this distance needs to be is, of course, arguable, and it is only possible to know for certain by computing the effect of

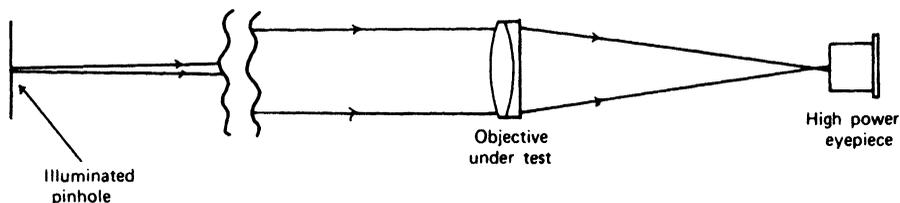


FIGURE 11.15. Arrangement for testing a small refracting objective; the light may or may not be filtered, according to the detailed requirements of the test.

using a finite conjugate from the design specification. For most practical purposes, the rule “infinity is a distance greater than 20 times the focal length of the system under test” is good enough.

The eyepiece for viewing the star must have enough magnification to show all the detail in the point spread function. The facts that the scale of detail is of the order of the radius of the first dark ring, and that the angular resolution limit of the eye under favorable conditions is about one arc-min, lead to the handy rule that the focal length of the eyepiece in millimeters should be equal to or less than the f number of the system under test;[†] thus to test an $f/10$ beam we need a 10-mm eyepiece, that is, magnification $\times 25$. With some experience it is easy to tell when the eyepiece has enough magnification, since only then all the diffraction structure in the star image can be seen.

Needless to say, we must have an eyepiece that itself has negligible aberrations to star-test a good objective. It is rather unlikely that this would not be so since our rule implies that the maximum diameter of the ray pencils in the eyepiece will be 1 mm and most eyepiece designs will be substantially perfect at this aperture. However, there may be accidental defects such as scratches or slight decentering, and these can easily be identified by rotating the eyepiece.

For looking at beams with very short focal ratios, it may be necessary to use a low-power microscope instead of an eyepiece. The rule then takes the form that the overall magnification of the microscope should not be less than 250 divided by the f number.

Sometimes a collimator, either refracting or reflecting, is used to place the star exactly at infinity. This is undesirable, partly because added cost is incurred and partly because the aberrations of the collimator are then added to those of the system under test. However, under some workshop conditions the long air path needed for the arrangement of Figure 11.15 may be subjected to unavoidable thermal turbulence, and it may be necessary to use a collimator instead. Usually, under average conditions an air path exceeding 1 or 2 m is enough to show turbulence effects with a critical optical system, so we must either screen a longer path, put it in a separate test bay, or wait for a particular time of day or night when the disturbances are least. Occasionally, the whole system has been put in a helium atmosphere or even in a vacuum tank for test purposes, but this is not practicable in most workshops. A large concave spherical mirror can be tested in Figure 11.16 since it has no spherical aberration for equal conjugates. However, it is necessary to keep the separation of the two conjugates small to avoid astigmatism. This separation should be less than about

$$\frac{R\sqrt{\lambda R}}{D} \quad (11.3)$$

where R is the radius of curvature of the mirror and D its diameter. For example, if $R = 1$ m and $D = 200$ mm, the separation must be less than 7 mm.

A paraboloidal mirror needs one infinite conjugate. There seems little doubt that star testing is not the best method for short focal ratio paraboloids; on the contrary,

[†] Strictly speaking, these figures give twice the f number, but it is as well to have a factor of 2 at hand in magnification.

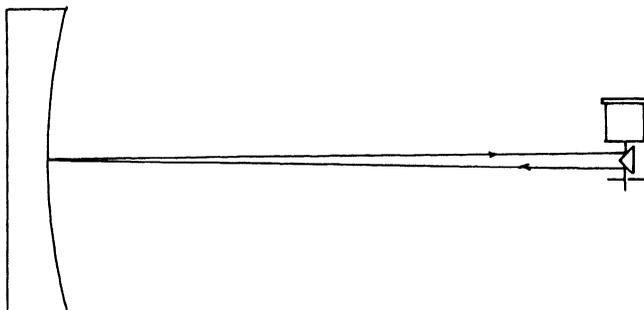


FIGURE 11.16. Testing a concave spherical mirror. The right-angle prism is aluminized on the two faces used as mirrors. For a mirror with a short focal ratio, the eyepiece would have to be replaced by a low-power microscope. Care must be taken that the pinhole is illuminated so as to fill the whole aperture of the mirror; likewise, the microscope objective must have a numerical aperture large enough to collect the light from the aperture.

however, paraboloidal mirrors are very useful as collimators for star testing other systems. In general, to test conicoidal surfaces, any of the null test configurations described in Chapter 12 can be used.

11.3.4. Microscope Objectives

The star test is particularly useful for microscope objectives (Martin and Welford, 1971), and the techniques are rather different from those so far described. The artificial star is made by vacuum-alumizing or silvering a microscope slide and gently wiping it with a clean cloth before cementing a coverslip to it. Unless the slide was cleaned very carefully before coating, the wiping will produce several pinholes, among which will be some that are suitable as artificial stars. An alternative technique described by Slater (1960) will produce pinholes of precisely determined size down to about $0.1 \mu\text{m}$ diameter. The test is then carried out with a bright lamp and an efficient substage condenser.

As mentioned at the beginning of this chapter, the star test is used in the rapid final adjustment of a critical air space to balance the spherical aberration correction for objectives of high numerical aperture. The spherical aberration of these objectives is also sensitive to both the thickness of the slide cover glass (except for oil immersion objectives) and the magnification; the latter is set by the length of the microscope draw tube, if there is one. Thus for critical work it is necessary to determine the cover glass thickness for which a given objective has least spherical aberration (usually a combination of high orders); alternatively, we may be able to find the magnification at which the objective performs best with a given cover glass thickness.

The star test is also convenient for checking the flatness of allegedly flat field objectives and for checking the chromatic correction of apochromatic objectives. Finally, it is known that some specimens of fluogic scatter light strongly, and the star test shows whether the fluorite in a given apochromat or semiapochromat has this

undesirable property: the scattered light shows up as a haze, easily distinguishable from the diffraction rings, around the star image.

11.4. THE STAR TEST WITH LARGE ABERRATIONS

Very simple and useful quantitative tests can be made on systems such as photographic objectives, projection objectives, and television camera objectives that do not have diffraction-limited aberration correction. The system will normally be set up on a nodal slide optical bench with a collimator (Kingslake, 1932), as shown in Figures 11.17 and 11.18, so that off-axis aberrations can be measured. Many detailed descriptions of such benches are available; see, for example, Leistner et al. (1953).

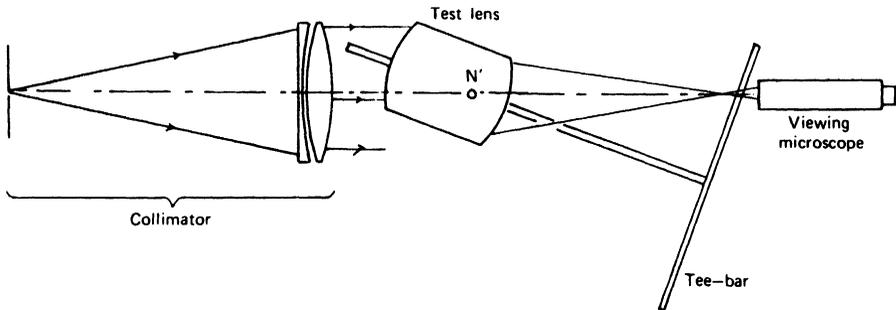


FIGURE 11.17. A nodal slide optical bench as used for star testing photographic objectives and similar systems. The T-bar represents the optical axis and the focal plane of the lens under test; as it turns off axis it pushes the viewing microscope away so that it is always focused on the correct focal plane. (But in some versions the microscope does not move, and the lens is moved forward by the linkage instead; in yet other versions there is no T-bar, and the microscope must be moved back by a calculated amount.)

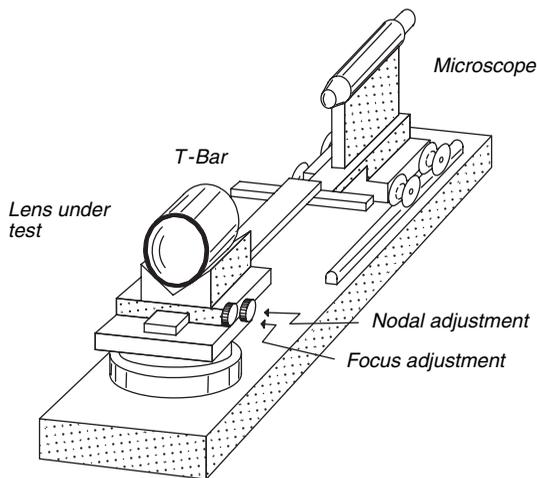


FIGURE 11.18. Schematic perspective view of a nodal lens bench.

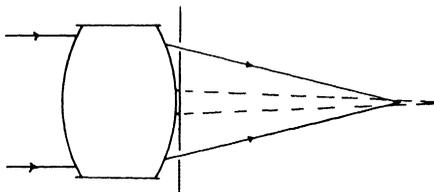


FIGURE 11.19. Measuring longitudinal spherical aberrations by the star test with a zonal diaphragm. The broken lines indicate a paraxial pencil of rays.

The principles in star testing for large aberrations are quite different from those for small aberrations; for the latter, we examine the diffraction structure of the point spread function and try to estimate the types and amounts of aberrations on the basis of experience, but for large aberrations, we essentially use concepts of geometrical optics and determine ray aberrations. Wandersleb (1952) gave some good photographs of heavily aberrated point spread functions.

11.4.1. Spherical Aberration

Here we set the objective under test on axis and use a filter for the appropriate wavelength. Then, we use a series of annular diaphragms in the aperture of the system and determine the focus of the hollow pencil of rays from each as shown in Figure 11.19. In this way, the longitudinal spherical aberration curve is obtained, and this can be treated in the usual ways to obtain any other form of description of the aberration; that is to say, when we know the longitudinal spherical aberration, we can calculate the transverse spherical aberration.

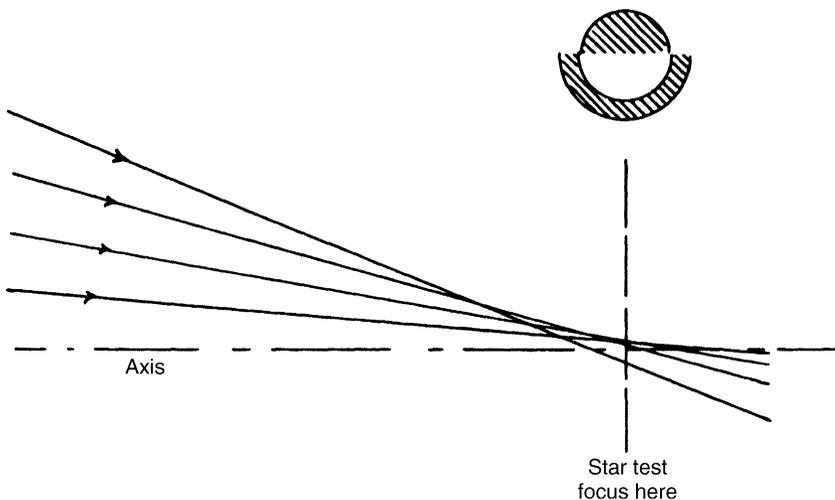


FIGURE 11.20. Testing for spherical aberration by masking half of the aperture; the hatched pattern will appear bright at the focal plane indicated.

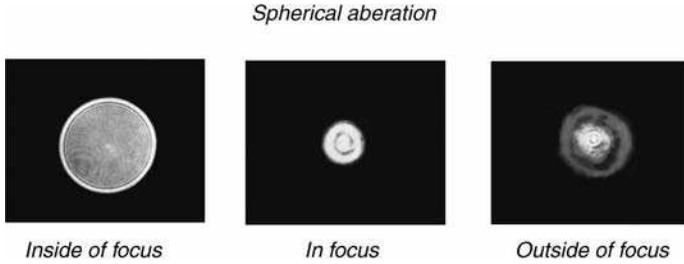


FIGURE 11.21. Inside of focus, in focus, and outside of focus images with spherical aberration.

In this procedure, as in the others described in Section 11.4, there is some advantage in arranging such that the object pinhole is not quite diffraction limited; this not only gives a little more light flux, a useful bonus in most cases, but also makes it rather easier to judge the ray intersections.

A quicker procedure, which gives only a qualitative picture of the aberrations, is simply to put an opaque straightedge across half the aperture as shown in Figure 11.20 and look at the image in various focal planes. As indicated, the general nature of the aberration appears as semicircular strips on either side of the axis, and with a little practice this image can easily be interpreted. The inside of focus, in focus, and outside of focus images with spherical aberration are shown in Figure 11.21.

11.4.2. Longitudinal Chromatic Aberration

Suitable filters are used, and the aperture is stopped down until the spherical aberration is negligible; the longitudinal chromatic aberration can then be plotted directly from the measured focal positions for the different wavelengths. If we use annular diaphragms, we can also measure the chromatic variation of spherical aberration; however, this refinement is rarely of interest in a practical workshop situation.

Again the straightedge across half the aperture can be used as for spherical aberration, without filters in the light source. The distribution of colors then gives a rapid indication of the balancing of the chromatic correction.

11.4.3. Axial Symmetry

In the systems with which Section 11.4 is concerned, it is usual to find some asymmetry on axis due to the building up of constructional tolerances. This is particularly so with zoom lenses, where the decentering may vary over the zoom range. The axial star image provides an instant check on the centering and may be used as a control in improving it, if this is required. To aid in this, it is useful to have an accurate rotation mechanism on the nodal slide mount for the lens under test.

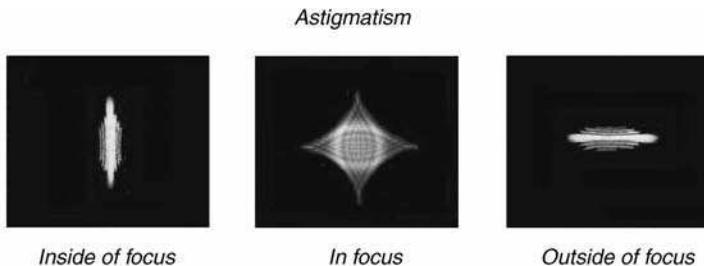


FIGURE 11.22. Inside of focus, in focus, and outside of focus images with astigmatism.

11.4.4. Astigmatism and Coma

We can plot the positions of the sagittal and tangential focal lines by turning the lens off axis and focusing in turn on each line. Of course, in this process we also see all the other off-axis aberrations (coma, transverse chromatic aberration, etc.). We therefore close down the lens iris enough to eliminate or nearly eliminate spherical aberration and coma, but not, of course, to the extent that the astigmatism disappears as well.

In a large system, it may be as well to check the symmetry of the lens field by rotating the lens and watching the astigmatic focal lines at a chosen field angle. Sometimes, more asymmetry is found off axis than on axis. A star test measurement of the astigmatic field curvature, together with some assessment of coma, is often used to tune the air spacings of large photographic objectives. Figure 11.22 shows the inside of focus, in focus and outside of focus images in the presence of astigmatism. The corresponding images for coma are in Figure 11.23.

11.4.5. Distortion

The accurate calibration of, for example, photogram metric objectives for distortion is a very large topic that is not considered in this book. However, it is worth noting that the nodal slide optical bench used for star testing enables us to make a simple measurement of distortion that is useful in many cases.

When the lens under test is being set up, we set the axis rotation of the nodal slide under the second nodal point by the following procedure. We find the star image in the microscope and turn the slide a small amount off axis. If the star image moves,

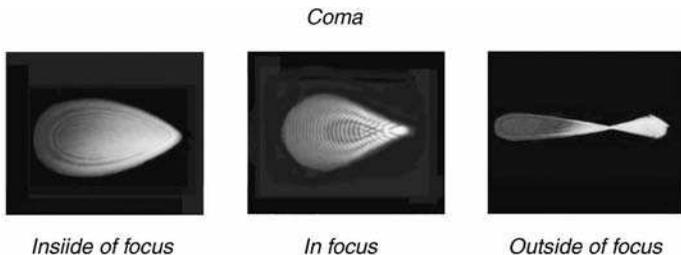


FIGURE 11.23. Inside of focus, in focus, and outside of focus images with coma.

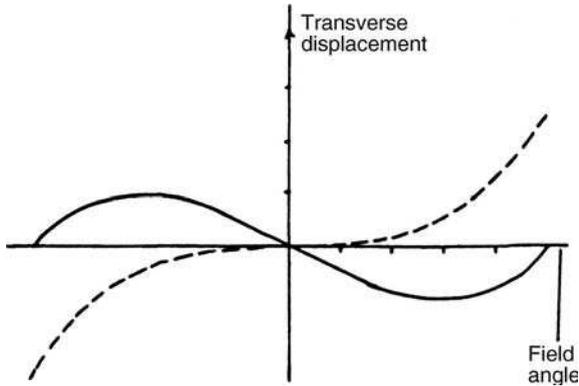


FIGURE 11.24. Distortion measurement. The full-line graph represents transverse displacement of the star image as a function of the field angle, assuming that a compromise position for the nodal axis has been found. The broken line shows the effect of adding a linear term; this amounts to a shift of the nodal point to the paraxial position.

say, to the left, we move the lens along the nodal slide in the direction that brings the image back to the center of the microscope field of view, at the same time keeping it in focus by following it with the microscope. This procedure is continued until the image does not move sideways at all as the nodal slide is rotated; the lens is then correctly positioned. However, it is almost invariably found that there is no position of the lens along the nodal slide that will give no movement at all of the star image; the best we can do is, rather, as shown in Figure 11.24, where the image returns to the axis after a short excursion. The residual displacement is, of course, a manifestation of the distortion of the lens. The full-line graph indicates what might be measured. It would be more usual to subtract from this a suitable linear term to give the broken line graph for distortion, which depends on cubic and higher powers of the field. The subtraction of a linear term merely implies a slightly different choice of nominal focal length for the system under test.

A detailed description of this procedure, with a discussion of errors, has been given by Washer and Darling (1959).

11.4.6. Non-Null Tests

By implication, we have so far considered only null tests, that is, tests of systems that ought to have no aberrations. Sometimes, however, we need to test a system to see whether it has prescribed nonzero aberrations. The classical example is, of course, the paraboloid mirror with equal conjugates, which is geometrically more convenient than one conjugate at infinity. Another example might be an aspheric singlet lens for use in the infrared, since its aberrations would be different if tested in visible light.

Since the star test is not very accurately quantitative, even when used as described in Section 11.3, for actually measuring large aberrations, it is best to convert such a

test into a null test by adding an auxiliary optical system with the required aberrations (see Chapter 14). This is a designed optical system that should be easy to be made accurately; Dall (1947) described the widely used system of an auxiliary lens for null-testing paraboloids, †and many other examples, which have been published.

A more recent alternative, which can be used if it is permissible to test the system with a laser as light source, is to employ a computer-generated hologram as the auxiliary system. A useful description of this method is given by Birch and Green (1972). The computer-generated hologram can synthesize a wavefront of any desired shape, provided that the available computer graphics system has enough resolution to draw it. As usually described, the computer-generated hologram corrector is part of an interferometric testing scheme, but it could be used equally well to synthesize an aberrated wavefront for transmission through the system under test; if the system had the desired aberrations, the result, would be an aberration free star image.

11.5. WAVEFRONT RETRIEVAL WITH SLOPE AND CURVATURE MEASUREMENTS

The wavefront shape can be retrieved from wavefront slopes and curvature measurements by integration of the Poisson equation as we will describe in the next few sections. On the contrary, the local slopes and the local curvatures can be obtained by any of many possible methods. For example, the slopes can be measured by means of some classic procedures, like the Hartmann and Ronchi tests or by using lateral shearing interferometry. It is interesting, however, that the local slopes and curvatures can also be obtained by measuring the irradiance of defocused images, making the star test a quantitative procedure.

11.5.1. The Laplacian and Local Average Curvatures

Roddier (1988) and Roddier et al. 1988, based on a proposal by Beckers (1994) in 1979, developed a method to evaluate the wavefront shape by measuring the local curvatures from defocused star images. The local curvatures c_x and c_y of an almost flat wavefront in the x and y directions are given by

$$c_x = \frac{\partial W^2(x, y)}{\partial x^2} \quad \text{and} \quad c_y = \frac{\partial W^2(x, y)}{\partial y^2} \quad (11.4)$$

The Laplacian, which is defined by

$$\nabla^2 W(x, y) = 2\rho(x, y) = \frac{\partial W^2(x, y)}{\partial x^2} + \frac{\partial W^2(x, y)}{\partial y^2} \quad (11.5)$$

† The paper cited refers to knife-edge testing, but the principle of the auxiliary lens is, of course, the same.

is equal to twice the average of the local curvature $\rho(x, y)$. This is the well-known Poisson equation. To solve it in order to obtain the wavefront deformations $W(x, y)$, we need to know: (a) the average local curvatures distribution $\rho(x, y)$, which is a scalar function and (b) the radial wavefront slopes at the edge of the circular pupil, which are the so called Neumann boundary conditions.

To solve the Poisson equation as described by Roddier et al. (1988), we can use the Jacobi iteration algorithm or, alternatively, we can use iterative Fourier methods as will be shown next.

11.5.2. Wavefront Determination with Iterative Fourier Transforms

Hardy et al. (1977) have measured slope differences to obtain the curvatures, and from them the Poisson equation is solved to obtain the wavefront. The curvature in the x direction is taken as the difference between two adjacent tilts in this direction, and in the same manner the curvature along the y axis is obtained. Then, the average of these curvatures is calculated. They have used the Hudgin (1977) algorithm to obtain this solution.

Roddier and Roddier (1991) and Roddier et al. (1990) have reported a method to obtain the wavefront deformations $W(x, y)$ from a knowledge of the Laplacian, solving the Poisson equation or from local slope measurements, by using iterative Fourier transforms. To understand this method, let us take the Fourier transform of the Laplacian of the wavefront as follows:

$$\mathcal{F}\{\nabla^2(x, y)\} = \mathcal{F}\left\{\frac{\partial^2 W(x, y)}{\partial x^2}\right\} + \mathcal{F}\left\{\frac{\partial^2 W(x, y)}{\partial y^2}\right\} \quad (11.6)$$

But on the contrary, from the derivative theorem in Fourier theory, we have

$$\mathcal{F}\left\{\frac{\partial W(x, y)}{\partial x}\right\} = i2\pi f_x \mathcal{F}\{W(x, y)\} \quad (11.7)$$

and similarly for the partial derivative with respect to y . In identical manner, we may also write

$$\mathcal{F}\left\{\frac{\partial^2 W(x, y)}{\partial x^2}\right\} = i2\pi f_x \mathcal{F}\left\{\frac{\partial W(x, y)}{\partial x}\right\} = -4\pi^2 f_x^2 \mathcal{F}\{W(x, y)\} \quad (11.8)$$

Thus, it is easy to prove that

$$\mathcal{F}\{\nabla^2(x, y)\} = -4\pi^2 \mathcal{F}\{W(x, y)\} (f_x^2 + f_y^2) \quad (11.9)$$

Hence, in the Fourier domain the Fourier transform of the Laplacian operator translates into a multiplication of the Fourier transform of the wavefront $W(x, y)$ by $f_x^2 + f_y^2$.

The wavefront may be calculated if measurements of the slopes along x and y are available, as in the case of the Hartmann and Ronchi tests as well as in lateral shearing interferometry, by means of this expression:

$$W(x,y) = -\frac{i}{2\pi} \mathcal{F}^{-1} \left\{ \frac{f_x \mathcal{F} \left\{ \frac{\partial W(x,y)}{\partial x} \right\} + f_y \mathcal{F} \left\{ \frac{\partial W(x,y)}{\partial y} \right\}}{f_x^2 + f_y^2} \right\} \quad (11.10)$$

This simple approach works for a wavefront without any limiting pupil. In practice, however, the Laplacian is multiplied by the pupil function, to take into account its finite size. Thus, its Fourier transform is convolved with the Fourier transform of the pupil function. As a result, this procedure does not give correct results. An apodization in the Fourier space, that is, a filtering of the frequencies produced by the pupil boundaries is needed, as in the Gershberg (1974) algorithm, to extrapolate the fringes outside of the pupil. The division by $f_x^2 + f_y^2$ produces this filtering. As a result of this filtering, just like in the Gershberg algorithm, after taking the inverse Fourier transform, the wavefront extension is not restricted to the internal region of the pupil, but extends outside the initial boundary.

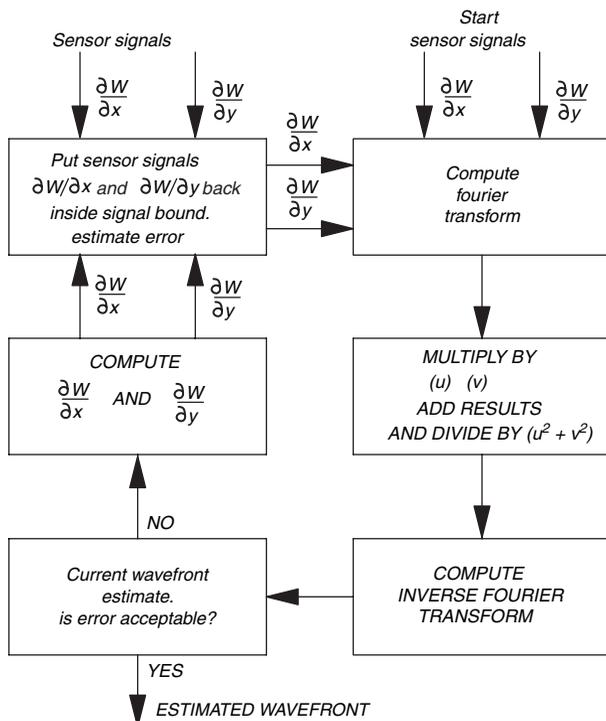


FIGURE 11.25. Flow chart of iterative Fourier transform algorithm used to find the wavefront from the measured slopes. (After Roddier and Roddier, 1991b.)

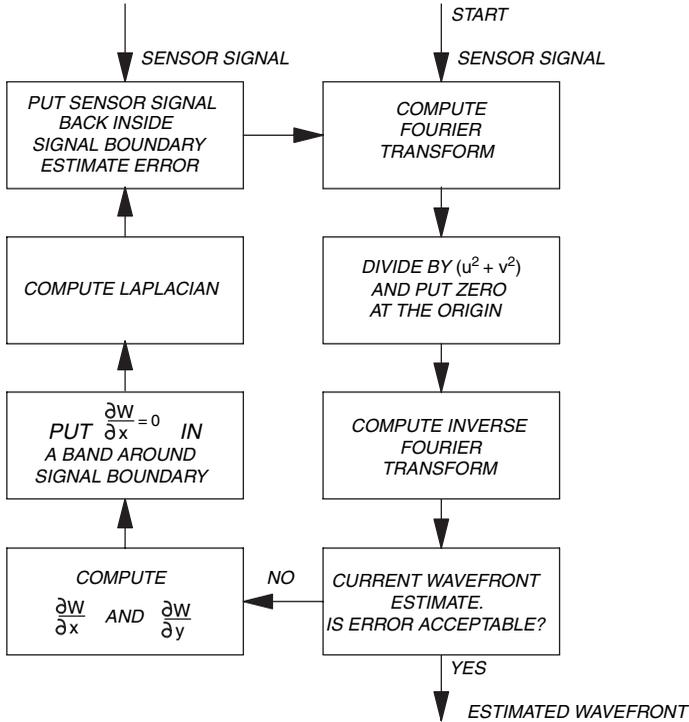


FIGURE 11.26. Flow chart of iterative Fourier transform algorithm used to find the wavefront from the measurement of the Laplacian. (After Roddier and Roddier, 1991b.)

The complete procedure to find the wavefront is thus an iterative process, as described in Figure 11.25.

We can also retrieve the wavefront by taking the Fourier transform of the Laplacian wavefront, dividing it by $f_x^2 + f_y^2$ and taking the inverse Fourier transform, as follows:

$$W(x, y) = -\frac{i}{4\pi} \mathcal{F}^{-1} \left\{ \frac{\mathcal{F} \{ \nabla^2 W(x, y) \}}{f_x^2 + f_y^2} \right\} \quad (11.11)$$

An iterative algorithm quite similar to the one just described, based on this expression has been also proposed by Roddier and Roddier (1991a and 1991b), as shown in Figure 11.26. The Laplacian is measured with a method to be described here, with two defocused images. The Neumann boundary conditions are taken by setting the radial slope equal to zero within a narrow band surrounding the pupil. To better understand this boundary condition, we may consider the wavefront curvature on the edge of the pupil as the difference between the slopes on each side of the pupil's edge. If the outer slope is set to zero, the curvature has to be equal to the inner slope. In other words, the edge radial slope is not arbitrarily separated from the inner curvature, if this external slope is made equal to zero.

11.5.3. Irradiance Transport Equation

To understand how the local curvatures and hence the Laplacian can be obtained from defocused star images, we have to model the image formation by using diffraction theory. Let us consider a light beam passing through a diffracting aperture (pupil) in the x - y plane and propagating with its average direction along the z axis. The irradiance as well as the wavefront shape (phase) continuously change along the trajectory. If the diffracting aperture is much wider than the wavelength, the diffraction Huygens wavelets are emitted in a narrow cone. Thus, we can use a parabolic approximation for the wavefront shape of each wavelet. This is equivalent to a geometrical optics approximation. Then, as shown by Teague (1983) and Steibl (1984), the disturbance at any plane parallel to the x - y plane and with any value of z may be found with the differential equation

$$\nabla^2 u(x, y, z) = 2k^2 u(x, y, z) + 2ik \frac{\partial u(x, y)}{\partial z} = 0 \quad (11.12)$$

instead of the Helmholtz diffraction equation, where $u(x, y, z)$ is the wave amplitude disturbance with its amplitude and phase, and the wave number is $k = 2\pi/\lambda$. We may assume a solution to this equation in the form

$$u(x, y, z) = I^{1/2}(x, y, z) \exp(ikW(x, y, z)) \quad (11.13)$$

where $I(x, y, z)$ is the irradiance. If we substitute this disturbance $u(x, y, z)$ into the differential equation, after some algebraic steps we can obtain a complex function that should be made equal to zero. Then, equating real and imaginary parts to zero, we obtain two equations. The first one is

$$\frac{\partial W}{\partial z} = 1 + \frac{1}{4k^2 I} \nabla^2 I - \frac{1}{2} \nabla W \bullet \nabla W - \frac{1}{8k^2 I^2} \nabla I \bullet \nabla I \quad (11.14)$$

which is known as the *phase transport equation* that can be used to find the wavefront shape at any point along the trajectory. The second equation is

$$\frac{\partial I}{\partial z} = -\nabla I \bullet \nabla W - I \nabla^2 W \quad (11.15)$$

which is known as the *irradiance transport equation*. Ichikawa et al. (1988) reported an experimental demonstration of phase retrieval based on this equation. In these two expressions, the (x, y, z) dependence has been omitted for notational simplicity. The Laplacian ∇^2 and the gradient operators ∇ work only on the lateral coordinates x and y . Following an interesting discussion by Ichikawa et al. (1988), the two terms forming the irradiance transport equation can be interpreted as follows:

- (a) The gradient $\nabla W(x, y, z)$ represents the direction and magnitude of the local slope (tilt) of the wavefront and the gradient $\nabla I(x, y, z)$ is the direction in which

the irradiance value changes with maximum speed. Thus, their scalar product $\nabla I(x, y, z) \cdot \nabla W(x, y, z)$ is the irradiance variation along the optical axis z , due to the local wavefront slope. Ichikawa et al. (1988) call this a prism term.

- (b) The second term $I(x, y, z)\nabla^2 W(x, y, z)$ can be interpreted as the irradiance along the z axis, caused by the local wavefront average curvature. Ichikawa et al. (1988) called this a lens term.

Thus, these two terms describe the variation of the beam irradiance caused by the wavefront deformations, as it propagates along the z axis. The transport equation, as pointed out before, is a geometrical optics approximation, valid away from sharp apertures, as long as the aperture is large enough compared with the wavelength of the light. The defocused image structure can be computed with this equation only if the diffraction effects are negligible, which is true only with a large defocusing.

To gain even more insight into the nature of this equation, it can be rewritten as

$$-\frac{\partial I(x, y, z)}{\partial z} = \nabla \bullet [I(x, y, z)\nabla W(x, y, z)] \quad (11.16)$$

and since ∇W is a vector representing the wavefront local slope, we can see that the transport equation represents the light energy conservation, which is analogous to the laws of mass or charge conservation, frequently expressed by

$$-\frac{\partial \rho}{\partial t} = \nabla \bullet (\rho v) \quad (11.17)$$

with ρ and v being the mass or charge density and the flow velocity, respectively.

11.6. WAVEFRONT DETERMINATION WITH TWO IMAGES USING THE IRRADIANCE TRANSPORT EQUATION

With two images with a large defocusing, one inside of focus and one outside of focus the local curvatures, the Laplacian and hence also the wavefront can be obtained by using the irradiance transport equation (Roddier, 1990).

Following Roddier et al. (1990), let $P(x, y)$ be the pupil function equal to one inside the pupil and zero outside the pupil. We also assume that the illumination at the plane of the pupil is uniform and equal to a constant I_0 inside the pupil. Hence, the irradiance gradient $\nabla I(x, y, 0)$ at the plane of the pupil is equal to zero everywhere except at the pupil's edge where

$$\nabla I(x, y, 0) = -I_0 \mathbf{n} \delta_c \quad (11.18)$$

where δ_c is a Dirac distribution around the pupil's edge and \mathbf{n} is unit vector perpendicular to the edge and pointing outward. On the contrary with this pupil function the gradient of W at this plane becomes

$$\nabla W = \frac{\partial W}{\partial \rho} \quad (11.19)$$

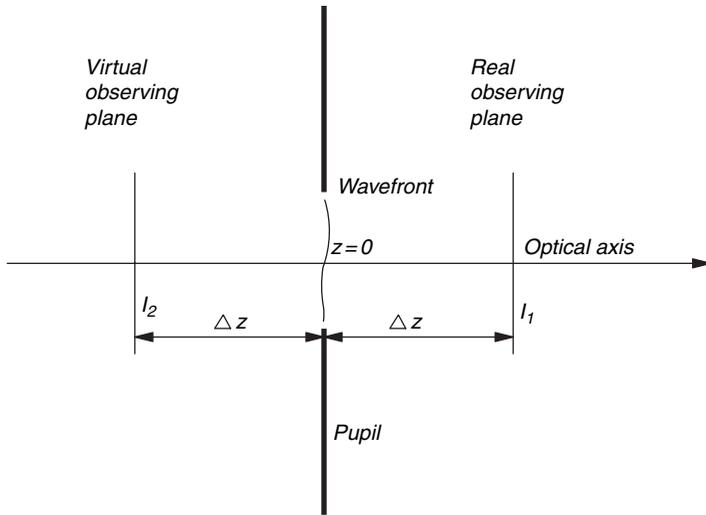


FIGURE 11.27. Irradiance measured in the two planes, symmetrically placed with respect to the pupil.

where ρ is the radial coordinate given by $\rho^2 = x^2 + y^2$. Substituting these values of the gradient of the irradiance and the wavefront into the irradiance transport equation evaluated at the pupil plane, we obtain

$$\left(\frac{\partial I(x, y, z)}{\partial z}\right)_{z=0} = -I_0 \cdot \left(\frac{\partial W(x, y, z)}{\partial \rho}\right)_{z=0} \delta_c - I_0 P(x, y) \nabla^2 W(x, y, z) \quad (11.20)$$

where the derivative on the right-hand side of this expression is the wavefront derivative in the outward direction, perpendicular to the pupil’s edge. The curvature sensing is performed by taking the difference between the illumination observed in two planes, symmetrically located with respect to the diffracting stop, as shown in Figure 11.27. Thus, the measured irradiances at these two planes are

$$\begin{aligned} I_1(x, y, +\Delta z) &= I_0 + \left(\frac{\partial I(x, y, z)}{\partial z}\right)_{z=0} \Delta z \\ I_2(x, y, -\Delta z) &= I_0 - \left(\frac{\partial I(x, y, z)}{\partial z}\right)_{z=0} \Delta z \end{aligned} \quad (11.21)$$

When the wavefront is perfectly flat at the pupil, both the Laplacian at all points inside the pupil and the radial slope at the edge of the pupil are zero. Then, $I_2(x, y, -\Delta z)$ is equal to $I_1(x, y, \Delta z)$. Having obtained these data, one may form the so called sensor signal as

$$s(x, y, \Delta z) = \frac{I_1 - I_2}{I_1 + I_2} = \frac{1}{I_0} \left(\frac{\partial I(x, y, z)}{\partial z}\right)_{z=0} \Delta z \quad (11.22)$$

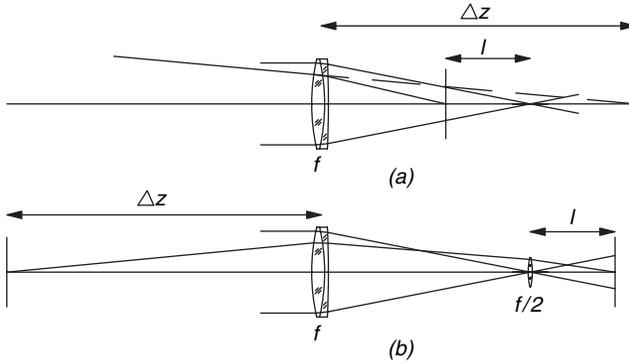


FIGURE 11.28. Two conjugate planes, one plane before refraction on the optical system, at a distance Δz from the pupil and, the second plane after refraction, at a distance l from the focus of the system. (a) With the first plane at the back of the pupil and the second plane inside of focus. (b) With the first plane at the front of the pupil and the second plane outside of focus, using an auxiliary small lens with focal length $f/2$.

Substituting Eq. (11.19) into Eq. (11.21) yields

$$\frac{I_1 - I_2}{I_1 + I_2} = \left(\frac{\partial W(x, y, z)}{\partial \rho} \delta_c - P(x, y) \nabla^2 W(x, y) \right) \Delta z \quad (11.23)$$

Thus, with the irradiance I_1 and I_2 in two planes symmetrically located with respect to the pupil ($z = 0$), we obtain the left term of this expression. This gives us the Laplacian of $W(x, y)$ (average local curvature) for all points inside the aperture and the wavefront slope $\partial W/\partial n$ around the pupil's edge $P(x, y)$, as a Neumann boundary condition, to be used when solving Poisson's equation.

The two planes on which the irradiance has to be measured are symmetrically located with respect to the diffracting pupil. In other words, one plane is real because it is located after the pupil, but the other plane is virtual because it is located before the pupil. In practice this problem has an easy solution because the diffracting aperture is the pupil of a lens to be evaluated, typically a telescope objective.

As we see in Figure 11.28, a plane at a distance l inside the focus is conjugate to a plane at a distance Δz after the pupil. On the contrary, if a small lens with focal length $f/2$ is placed at the focus of the objective, a plane at a distance l outside the objective focus is conjugate to a plane at a distance Δz before the pupil. In both cases the distance Δz and the distance l are related by

$$\Delta z = \frac{f(f - l)}{l} \quad (11.24)$$

Roddiier and Roddiier (1991) have pointed out that the small lens with focal length $f/2$ is not necessary if l is small compared with f . We must take into account that one defocused image is rotated 180° with respect to the other and also any possible

difference in the magnification of the two images. The important consideration is that the subtracted and added irradiances in the two measured images must correspond to the same point (x, y) on the pupil.

The measurements of the irradiance have to be made close enough to the pupil, so that the diffraction effects are negligible, and the geometric approximation remains valid. Let us assume that the wavefront to be measured has some corrugations and deformations of scale r_0 (maximum spatial period). With the diffraction grating equation, we see that these corrugations spread out the light over a narrow cone with an angular diameter $\alpha = \lambda/r_0$. Thus, the illumination in the plane of observation can be considered a blurred pupil image. Let us now impose the condition that the maximum allowed blurring at a distance Δz is equal to $r_0/2$. With this condition it is possible to show that the geometrical optics approximation implied in the transport irradiance equation is valid only if Δz is sufficiently small, so that the following condition is satisfied:

$$\Delta z \ll \frac{r_0^2}{2\lambda} \quad (11.25)$$

It is interesting to see that this distance Δz is one fourth of the Rayleigh distance in Talbot autoimaging. This result is to be expected, since then the shadow of the grating is geometrical. If the light angular diameter spread α is known, for example, if this is equal to the atmospheric light seeing in a telescope, we may also write

$$\Delta z \ll \frac{\lambda}{2\alpha^2} \quad (11.26)$$

When measuring the converging beam, this condition implies that the defocusing distance l should be large enough, so that

$$l \gg \frac{f}{1 + \frac{r_0^2}{2\lambda f}} \quad (11.27)$$

In conclusion, the minimum defocusing distance depends on the maximum spatial frequency of the wavefront corrugation we want to measure. This frequency also determines the density of sampling points to be used to measure the irradiance in the defocused image.

11.7. WAVEFRONT DETERMINATION WITH A SINGLE DEFOCUSED IMAGE USING FOURIER TRANSFORM ITERATIONS

If the defocusing distance cannot be made large enough, the geometrical optics approximation assumed by the irradiance transport equation is not satisfied. In that case, diffraction effects are important like in the classical star test, and the method described in the preceding sections cannot be applied in such a test. Thus, a different iterative methods have to be used.

Gershberg and Saxton (1972) described an algorithm using a single in-focus image rather than two defocused images, with the following steps:

- (a) An arbitrary guess of the wavefront deformations (phase and pupil transmission) is taken. The pupil transmission is frequently equal to one and the phase can be anything.
- (b) The in-focus image (amplitude and phase) in the observation plane is computed with a fast Fourier transform.
- (c) The calculated amplitude is replaced by the observed amplitude (square root of the observed intensity), keeping the calculated phase.
- (d) An inverse Fourier transform gives a new estimate of the incoming wavefront amplitude and phase (deformations) at the pupil plane.
- (e) The calculated input amplitude is replaced by the known input amplitude (pupil transmission), keeping the calculated phase.

These steps are iterated until a reasonably small difference between measured and calculated amplitudes is obtained. This algorithm quickly converges at the beginning, but then tends to stagnate.

11.8. WAVEFRONT DETERMINATION WITH TWO OR THREE DEFOCUSED IMAGES USING FRESNEL TRANSFORM ITERATIONS

Fienup and Wackermann (1987) described an improved method that converges more easily and avoids stagnation, reconstructing the images from the modulus of their Fourier transform. Misell (1973a and 1973b) found that by using two images, one in-focus and one out-of-focus, the results are much improved. Just like in the Gershberg and Saxton algorithm, iterations with the fast Fourier transform are carried out

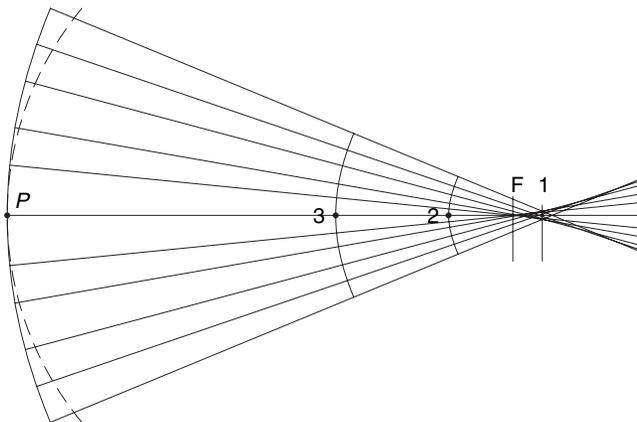


FIGURE 11.29. Aspheric wavefront and three planes where images are observed. (a) Best focus position *I*, (b) paraxial focus *F*, (c) Defocused images 2 and 3.

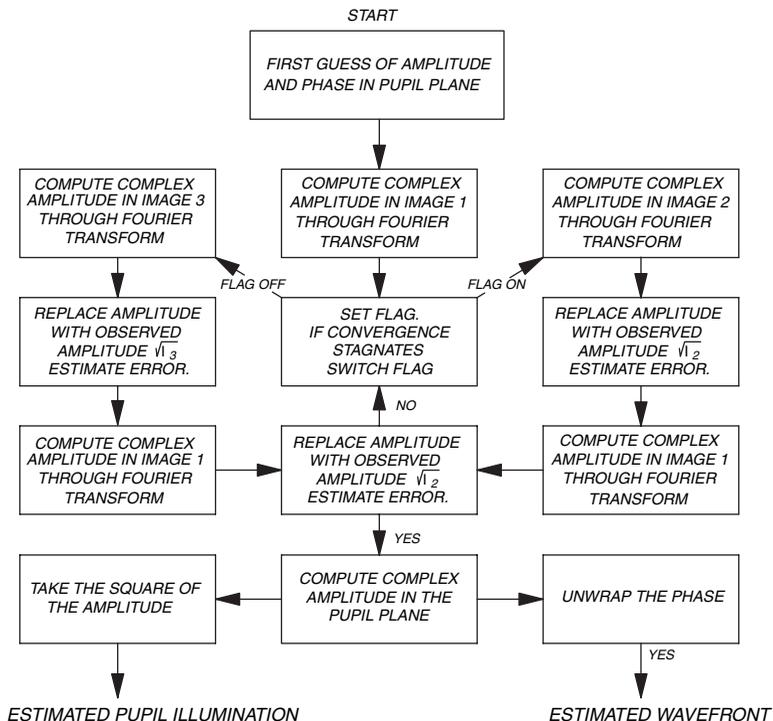


FIGURE 11.30. Flow chart of iterative Roddier and Roddier (1991a) used to find the wavefront and the pupil illumination for the Hubble Space Telescope.

between the pupil and the in-focus image, alternating with iterations between the pupil and the out-of-focus image of the pupil. To calculate the out-of-focus image, a quadratic phase term is added to the phase of the complex amplitude at the pupil function. This term introduces the desired defocusing. This procedure is called a Fresnel transform since the image is not in-focus.

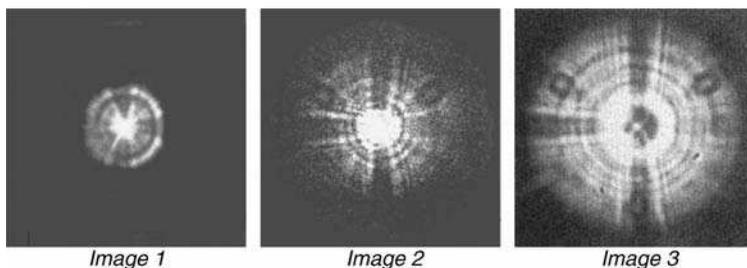


FIGURE 11.31. Three images used in the Roddier and Roddier (1991a) algorithm used to evaluate the Hubble Space Telescope. (From Roddier and Roddier, 1991a)

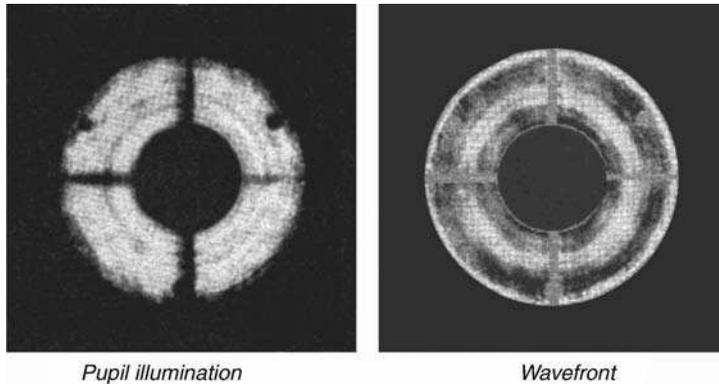


FIGURE 11.32. Results obtained for the evaluation of the Hubble Space Telescope. (a) Pupil illumination and (b) computed wavefront after defocusing and spherical aberration terms had been removed. (From Roddier and Roddier, 1991a).

Roddier and Roddier (1991b) developed an improved version of Misell method to compute the shape of the defective primary mirror of the Hubble telescope. They pointed out that a relatively large defocusing is more accurate than a short defocusing. The defocusing should be large, but not so large that the geometric approximation is possible. To understand how this defocused image may be obtained, let us consider Figure 11.29 where the pupil of the optical system, located at P produces an aberrated (aspherical) wavefront with paraxial focus at F . If we take the Fourier transform of the complex amplitude at the pupil, with the phase given by the separation between the wavefront and a reference sphere with center at F , we obtain the complex amplitude of the image at the paraxial focus. Now, if we take the Fourier transform of the complex amplitude at the pupil, with the phase given by the separation between the wavefront and a reference sphere with center at 3 , illustrated by the dotted curve, we obtain the complex amplitude of the image at the plane 3 . This is what we call a Fresnel transform. Roddier and Roddier noticed that this procedure may bring some errors, since the phase difference with respect to the reference sphere has so large slopes in the radial direction that this may cause undersampling errors.

The proposed alternative is to go from the pupil to the desired defocused observation plane in two steps; first from the pupil P to the best focus position 1 . Going to the best focus minimizes the wavefront deviations with respect to the reference sphere by introducing a small defocusing. Then, by taking the inverse Fourier transform with a different focusing factor we obtain the image at the desired defocused plane. This procedure implies two transforms instead of one, but the final result is better. Three observation planes are used in this method, as illustrated in Figure 11.29. The first step is to make a first guess of the pupil complex amplitude and to compute the image at the best focus 1 . The computed amplitude is replaced by the measured amplitude, keeping the calculated phase. Then, the complex amplitude at the plane 2 is calculated with an inverse Fourier transform using a different focus factor. Then, iterations are performed in this manner between planes 1 and 2 without

going back to the pupil system. When the convergence slows down, image 2 is left and iterations between planes 1 and 3 are made. Figure 11.30 shows the flow chart for this algorithm, using the three images. Figure 11.31 shows the three images taken with the Hubble telescope and Figure 11.32 shows the results.

REFERENCES

- Barakat R. and A. Newman, "Measurement of Total Illuminance in a Diffraction Image. I: Point Sources," *J. Opt. Soc. Am.*, **53**, 1965 (1963).
- Beckers J. M., "Interpretation of Out-of-Focus Star Images in Terms of Wave-Front Curvature," *J. Opt. Soc. Am. A.*, **11**, 425-427 (1994).
- Beiser L., "Perspective Rendering of the Field Intensity Diffracted at a Circular Aperture," *Appl. Opt.*, **5**, 869 (1966).
- Birch K. G. and F. J. Green, "The Application of Computer-Generated Holograms to Testing Optical Elements," *J. Phys., D: Appl. Phys.*, **5**, 1982 (1972).
- Born M. and E. Wolf, *Principles of Optics*, 5th Ed., Pergamon Press, Oxford and New York, 1975.
- Cagnet, M., M. Francon, and J. C. Thierr, *Atlas of Optical Phenomena*, Springer Verlag, Heidelberg and New York, 1962.
- Dall, H. E., "A Null Test for Paraboloids," *J. Br. Astron. Assoc.*, **57** (1947); reprinted in *Amateur Telescope Making*, vol. 3, A. E. Ingalls, Ed., Scientific American, New York, 1953.
- Fienup J. R. and C. C. Wackermann, "Phase-retrieval Stagnation Problems and Solutions," *J. Opt. Soc. Am. A.*, **3**, 1897-1907 (1986).
- Gershberg R. W. and W. O. Saxton, "A Practical Algorithm for the Determination of Phase from Image and Diffraction Plane Pictures," *Optik*, **35**, 237 (1972).
- Gershberg R. W., "Super-Resolution Through Error Energy Reduction," *Optica Acta*, **21**, 709-720 (1974).
- Hardy J. W., J. E. Lefevre and C. L. Koliopoulos, "Real-time Atmospheric Compensation," *J. Opt. Soc. Am.*, **67**, 360-369 (1977).
- Hudgin R. H., "Wave-front Reconstruction of Phase Values from Phase Differences," *J. Opt. Soc. Am.*, **67**, 375-378, 1977.
- Ichikawa K., A. W. Lohmann, and M. Takeda, "Phase Retrieval Based on the Irradiance Transport Equation and the Fourier Transport Method: Experiments," *Appl. Opt.*, **27**, 3433-3436 (1988).
- Kingslake R., "A New Bench for Testing Photographic Lenses," *J. Opt. Soc. Am.*, **22**, 207-222 (1932).
- Kogelnik H. and T. Li, "Laser Beams and Resonators," *Appl. Opt.*, **5**, 1550-1567 (1966).
- Leistner K., B. Marcus, and B. W. Wheeler, "Lens Testing Bench," *J. Opt. Soc. Am.*, **43**, 44 (1953).
- Li Y. and E. Wolf, "Three-Dimensional Intensity Distribution Near Focus in Systems of Different Fresnel Numbers," *J. Opt. Soc. Am. A*, **1**, 801-808 (1984).
- Linfort E. H. and E. Wolf, "On Telescopic Star Images," *Mon. Not. Ry. Astron. Soc.*, **112**, 452 (1952).

- Linfort E. H. and E. Wolf, "Diffraction Images in Systems with an Annular Aperture," *Proc. Phys. Soc.*, **B66**, 145 (1953).
- Martell L. C., *Technical Optics*, vol. 2, 2nd Ed., Pitman, London, 1961.
- Martin L. C. and W. T. Welford, in *Physical Techniques in Biological Research*, vol. I, Part A, 2nd ed., G. Oster, Ed., Academic Press, New York and London, 1971.
- Misell D. L., "An Examination of an Iterative Method for the Solution of the Phase Problem in Optics and Electron Optics. I: Test Calculation," *J. Phys. D., (Appl. Phys.)* **6**, 2200 (1973a).
- Misell D. L., "An Examination of an Iterative Method for the Solution of the Phase Problem in Optics and Electron Optics. II: Sources of Error," *J. Phys. D., (Appl. Phys.)* **6**, 2217 (1973b).
- Nienhuis K., *On the Influence of Diffraction on Image Formation in the Presence of Aberrations*, Thesis, J. B. Wolters, Groningen, 1948.
- Roddiier F., "Curvature Sensing and Compensation: A New Concept in Adaptive Optics," *Appl. Opt.*, **27**, 1223–1225 (1988).
- Roddiier F., "Wavefront Sensing and the Irradiance Transport Equation," *Appl. Opt.*, **29**, 1402–1403 (1990).
- Roddiier F., C. Roddiier and N. Roddiier, "Curvature Sensing: A New Wavefront Sensing Method," *Proc. SPIE.*, **976**, 203–209 (1988).
- Roddiier C., F. Roddiier, A. Stockton and A. Pickles, "Testing of Telescope Optics: A New Approach," *Proc. SPIE.*, **1236**, 756–766 (1990).
- Roddiier C. and F. Roddiier, "Wavefront Reconstruction Using Iterative Fourier Transforms," *Appl. Opt.*, **30**, 1325–1327 (1991b).
- Roddiier C. and F. Roddiier, "Reconstruction of the Hubble Space Telescope Mirror Figure from Out-of Focus Stellar Images," *Proc. SPIE.*, **1494**, 11–17 (1991a).
- Slater P. N., in *Optics and Metrology*, P. Mollet, Ed., Pergamon Press, Oxford and New York, 1960.
- Stammes J. J., *Waves in Focal Regions*, Adam Hilger, Bristol and Boston, 1986.
- Steibl N., "Phase Imaging by the Transport Equation of Intensity," *Opt. Commun.*, **49**, 6–10 (1984).
- Taylor C. A. and B. J. Thompson, "Attempt to Investigate Experimentally the Intensity Distribution near the Focus in the Error-Free Diffraction Patterns of Circular and Annular Apertures," *J. Opt. Soc. Am.*, **48**, 844 (1958).
- Taylor H. D., *The Adjustment and Testing of Telescope Objectives*, Sir Howard Grubb, Parsons & Co., Newcastle-upon-Tyne, 1891; 4th Ed., 1946, 5th Ed. 1983, Adam Hilger Ltd, Bristol, UK.
- Teague M. R., "Deterministic Phase Retrieval: A Green's Function Solution," *J. Opt. Soc. Am.*, **73**, 1434–1441 (1983).
- Twyman F., *Prism and Lens Making*, Adam Hilger, London, 1942. 2nd Rev. Ed. 1988, Adam Hilger, Bristol, UK.
- Wandersleb E., *Die Lichtverteilung in Der Axialen Kaustik eines mit Sphoiiirischer Aberration Behafteten Objektivs*, Akademie-Verlag, Berlin, 1952.
- Washer F. E. and W. R. Darling, "Factors Affecting the Accuracy of Distortion Measurements Made on the Nodal Slide Optical Bench," *J. Opt. Soc. Am.*, **49**, 517 (1959).
- Welford W. T., "On the Limiting Sensitivity of the Star Test for Optical Instruments," *J. Opt. Soc. Am.*, **50**, 21 (1960).

12

Testing of Aspheric Wavefronts and Surfaces

D. Malacara, K. Creath, J. Schmit and J. C. Wyant

12.1. INTRODUCTION

Aspheric wavefronts with spherical aberration are produced by optical systems using spherical as well as aspherical surfaces. Aspheric surfaces are used in optical systems in order to improve aberration correction and, frequently, to decrease the number of optical elements needed to make this correction satisfactorily. However, if these surfaces are tested while being isolated from the rest of the optical system to which they belong, they frequently produce aspherical wavefronts. The interferometric testing and measurement of aspherical wavefronts are not as simple as in the case of spherical or flat wavefronts.

To test aspherics, often a null test is issued. The usual definition of a null test is that which produces a fringe-free field when the desired wavefront is obtained. Then, if a tilt between the wavefront under test and the reference wavefront is added and the paraxial curvature of them are equal, straight and parallel fringes are obtained. Under these conditions, any deviation from straightness of the fringes is a graphical representation of the wavefront deformation. This is the ideal testing procedure because the desired wavefront is easily identified and measured with high accuracy. There are several methods to obtain this null test, but sometimes this is not simple and may even be a source of possible errors.

Typically, if a quantitative retrieval of the wavefront is desired, the interferogram is imaged onto a CCD detector. Then, the straightness of the fringes for a perfect wavefront is useful but not absolutely necessary. However, the minimum fringe spacing should be larger than twice the pixel size in the detector. This is the well-known Nyquist condition, which may be impossible to satisfy if the wavefront has a strong asphericity.

In a Fizeau or Twyman–Green interferogram, a strong rotationally symmetric aspheric wavefront has many fringes when taken at the paraxial focus setting as shown in Figure 12.1(a). By adding a small curvature to the wavefront, that is, by adding defocusing, the minimum fringe spacing can be slightly reduced. For

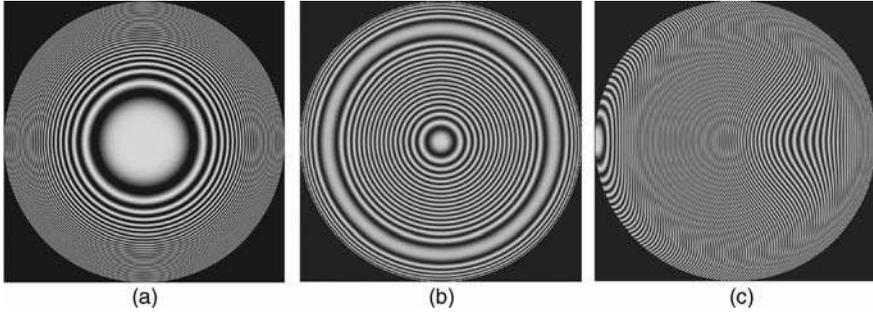


FIGURE 12.1. Interferogram of an aspheric wavefront with primary spherical aberration (a) at the paraxial focus, without tilt; (b) at the best focus, without tilt; and (c) at the best focus, with large tilt.

example, let us assume that the aspheric wavefront with rotational symmetry can be represented by

$$W(S) = aS^2 + bS^4 \tag{12.1}$$

where S is the radial distance from the optical axis ($S^2 = x^2 + y^2$), the first term to the right of the equal sign is the defocusing, and the second term is the asphericity (primary aberration). The defocusing coefficient a should be chosen so that the minimum fringe spacing on the pupil aperture is as large as possible. The fringe spacing is defined by the wavefront slope, so the maximum wavefront slope has to be minimized. This condition is satisfied if (see Fig. 12.2)

$$a = -\frac{3}{2} bS_{max}^2 \tag{12.2}$$

This increases the minimum fringe spacing in the interferogram by a factor of 4, as illustrated in Figure 12.1(b).

To simplify the analysis of a single fringe pattern with closed fringes, it is sometimes necessary to introduce a large tilt so that only open fringes appear as shown in Figure 12.1(c). Then, we say that a linear carrier has been introduced along

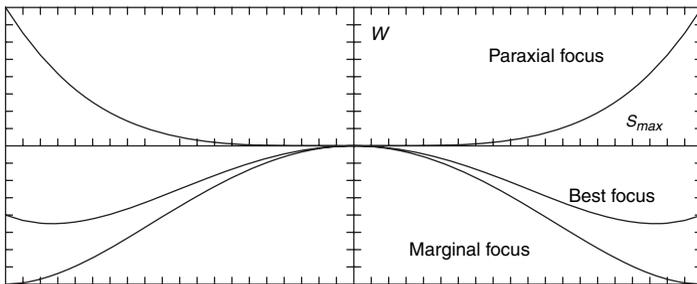


FIGURE 12.2. Wavefront profiles for primary spherical aberration at three different focus settings.

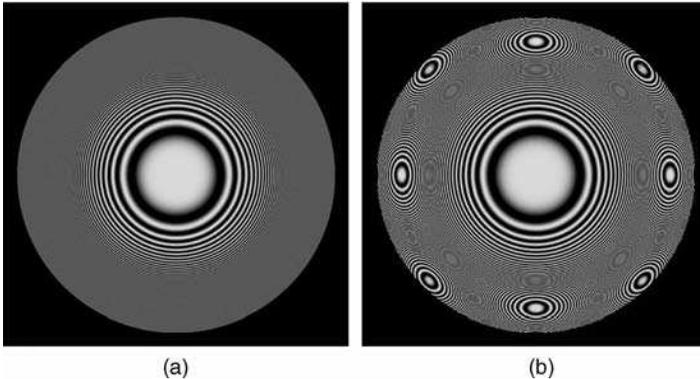


FIGURE 12.3. Interferogram of an aspheric wavefront with primary spherical aberration, at the paraxial focus, with fringe spacing smaller than twice the pixel size. (a) With pixel size equal to pixel separation and (b) with pixel size smaller than pixel separation.

the x axis. With these conditions, when scanning the wavefront along horizontal lines the same fringe is not scanned twice. The problem with this method is that the introduction of the large tilt that opens the fringes also decreases the fringe spacing. The minimum fringe spacing occurs where the sum of the slope of the wavefront under test plus the slope of the reference wavefront tilt is the largest.

In a rotationally symmetric aspheric wavefront, the fringes are concentric rings and also in many other interferograms the fringes may close, forming loops. This kind of interferograms are difficult to analyze and frequently some special complicated techniques have to be used. A second and more important disadvantage is that the fringe spacing is quite small near the edge of the pupil.

When sampling an interferogram where the Nyquist condition is violated, each pupil detector measures the average intensity of the light falling over its small area. If the fringe spacing is smaller than twice the pixel size, the image of the fringes is washed out or the contrast will be greatly reduced if these two dimensions are similar as shown in Figure 12.3(a). If the pixel size is much smaller than its spacing, instead of a low fringe contrast, some false spurious fringes will appear, as shown in Figure 12.3(b). This effect is known as aliasing. Most CCD detectors have a pixel spacing almost equal to the pixel size. So, fringe aliasing is not common.

12.2. SOME METHODS TO TEST ASPHERIC WAVEFRONTS

There are several methods that had been used to test aspheric surfaces, which will be described in this chapter. The methods for testing the quality of aspheric surfaces that have been developed can be classified into the following categories:

1. One of the several possible non-null classic tests can be quantitatively used, measuring the wavefront with the Foucault, Ronchi or Hartmann tests, mathematically

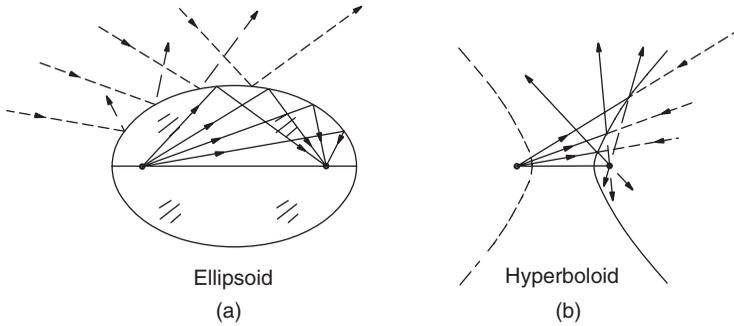


FIGURE 12.4. Reflective concave and convex ellipsoidal and hyperboloidal surfaces.

in calculating the surface deformations with respect to the closest sphere. This process in general is time consuming, and frequently the accuracy of the results is not high enough if the aspheric deformation is strong, due to their relatively smaller sensitivity as compared with interferometry. These methods are described in Chapters 8–10.

2. For conicoidal surfaces with symmetry of revolution, there is always a pair of conjugate foci that are free of spherical aberration as illustrated in Figure 12.4. Then, if the surface under test is illuminated with a convergent or divergent wavefront with the center of curvature at the proper focus, the reflected wavefront is spherical and thus is easier to test. To produce the necessary illuminating wavefront, some auxiliary optical elements are nearly always necessary. There are several of these null configurations that will be described here.

3. Some additional optical elements can be added to the testing system to compensate for the spherical aberration of the wavefront reflected from the aspheric surface. Then, an auxiliary optical system is designed so that, in combination with the aspheric surface, it forms a stigmatic image of a point source. The auxiliary optical system is called a null corrector or null compensator. These methods will be described in this chapter.

4. If only one interferogram picture is taken and the asphericity is not very strong, several interferometric methods can be used to evaluate the fringe pattern. However, if the asphericity is strong and the fringe spacing is not larger than twice the pixel separation at the detector (Nyquist condition), the interferogram may become impossible to analyze. Analysis is possible only with a procedure described by Greivenkamp (1987) and only when the following conditions are satisfied:

- (a) The pixel size is smaller than the pixel separation. Then, spurious fringes will appear where the Nyquist condition is not satisfied.
- (b) The wavefront's general shape is known.
- (c) The expected wavefront is smooth. The problem is then solved by proper phase unwrapping until the retrieved wavefront and its slopes are continuous.

5. With phase-shifting techniques, a series of Fizeau or Twyman–Green interferograms can be used without the introduction of tilt, even if closed loop fringes are formed. A series of a minimum of three interferograms have to be taken with different values of the constant term for the phase, also called piston term. Then, the interpretation of the wavefront evaluation becomes relatively simple from a mathematical point of view, but it is much more complicated from the experimental point of view. The phase shifting techniques are described in detail in Chapter 14. When measuring an aspheric wavefront with phase shifting, the defocusing term has to be properly chosen so that the fringe spacing is minimum as described before. As pointed out before, the limitation is that by using the optimum focus setting, the Nyquist condition is not violated.

6. If the wavefront has a strong deviation from sphericity, even phase shifting techniques become impossible. Another possibility under these conditions is to test the wavefront by dividing the complete aperture into small regions where the Nyquist condition is not violated. In other words, in all small regions the fringe spacing should be larger than twice the pixel separation. This technique, sometimes referred to as a wavefront stitching technique, will be described in this chapter.

7. If a longer wavelength is used, the dynamic range is increased by reducing some of its sensitivity. For example, by operating at $10.6\ \mu\text{m}$, the dynamic range is increased twenty times, approximately.

8. Sources and detectors for longer wavelengths may be prohibitively expensive. A solution to this problem may be found in two-wavelength methods, where fringes at longer (synthetic) wavelength are created by simultaneous exposure at two different wavelengths. These fringes, which are resolvable for high slopes at a longer wavelength, may be analyzed with a phase shifting method. Different synthetic wavelengths can be obtained by combining different visible wavelengths, as described in Chapter 15 and Section 12.13 in this chapter.

9. Lateral or radial shear interferometry, as described in Chapters 4 and 5, provides a larger dynamic range for testing strong aspherics, but a smaller sensitivity.

12.3. IMAGING OF THE INTERFERENCE PATTERN IN NON-NULL TESTS

An aberrated wavefront continuously changes its shape as it travels; therefore, if the wavefront is aspherical, the interference pattern will also continuously change as the beam advances as shown in Figure 12.5. Since the errors of an instrument are represented by wavefront distortions on the pupil, the interferogram should represent the wavefront deformations at that place.

The problem is even worse if the wavefront travels twice through the optical system during the test. For example, when testing a lens with any of the configurations as described in Chapter 2, the wavefront travels twice through the lens; the second time, it travels after being reflected from the small mirror in front of the lens. If the aberration is small, the total wavefront deformation is twice the deformation

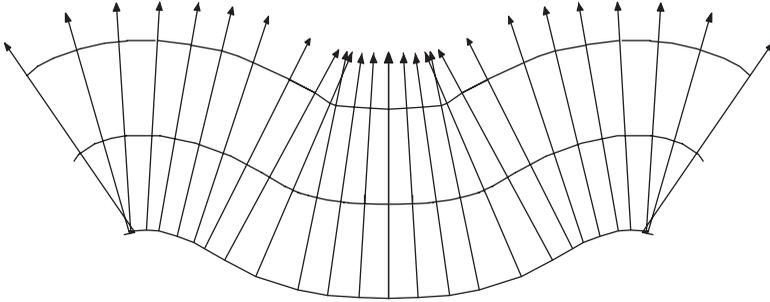


FIGURE 12.5. Change in the wavefront profile as it travels.

introduced in a single pass through the lens. However, if the aberration is large, this condition is not followed, because the wavefront changes while traveling from the lens to the mirror and back to the lens. Then, the spot in the surface on which the defect is located is not imaged back onto itself by the concave or convex mirror, and the ray will not pass through this defect the second time. Great confusion then results with regard to the interpretation of the interferogram, since the defect is not precisely duplicated by the double pass through the lens.

It may be shown that the image of the lens is formed at a distance L from the lens, which is given by

$$L = \frac{2(F - r)^2}{2F - r} \quad (12.3)$$

where F is the focal length and r is the mirror radius of curvature ($r > 0$ for a convex mirror and $r < 0$ for a concave mirror). We may see that the ideal mirror is convex and very close to the lens ($r \sim F$).

An adequate optical arrangement has to be used if the lens under test has a large aberration, in order to image its pupil back on itself. Any auxiliary lenses or mirrors must preserve the wavefront shape. Some examples of these arrangements are shown in Figure 12.6 (Malacara and Menchaca, 1985).

However, for microscope objectives this solution is not satisfactory because the ideal place to observe the fringes is at the back focus. In this case the Dyson (1959) system illustrated in Figure 12.7 is an ideal solution. It is interesting to point out that Dyson's system may be used to place the self-conjugate plane at concave or convex surface, while maintaining the concentricity of the surfaces.

Even if the wavefront passes only once through the optical system under test, the second problem is to image the interference pattern on the observing detector, screen, or photographic plate. The imaging lens does not need to preserve the wavefront shape if it is placed after the beam splitter, and thus both interfering wavefronts pass through this lens. However, this lens has to be designed in such a way that the interference pattern is imaged without any distortion, assuming that the pupil of the system is at an image of the point light source as shown in Figure 12.8(a). A rotating

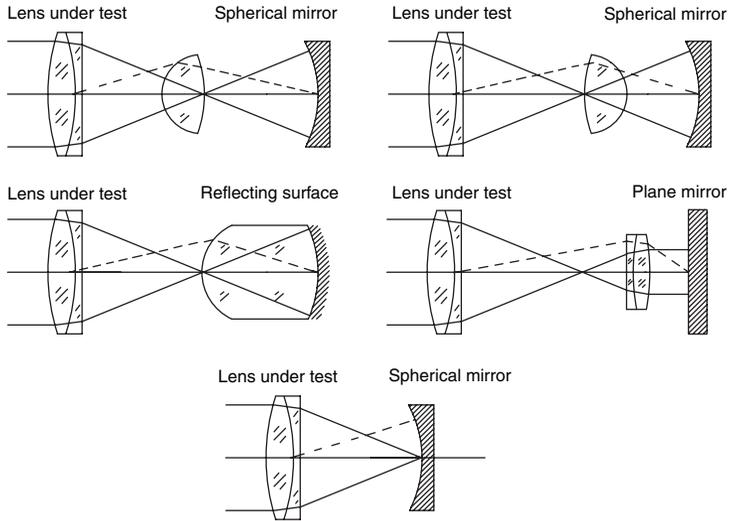


FIGURE 12.6. Systems to image the pupil of the optical element under test back on itself after reflection at the mirror.

ground glass in the plane of the interferogram might be sometimes useful in order to reduce the noise due to speckle and dust in the optical components. Ideally, this rotating glass should not be completely ground in order to reduce the loss of brightness and to keep the stop of the imaging lens at the original position, as shown in Figure 12.8(b). If the rotating glass is completely ground, the stop of the imaging

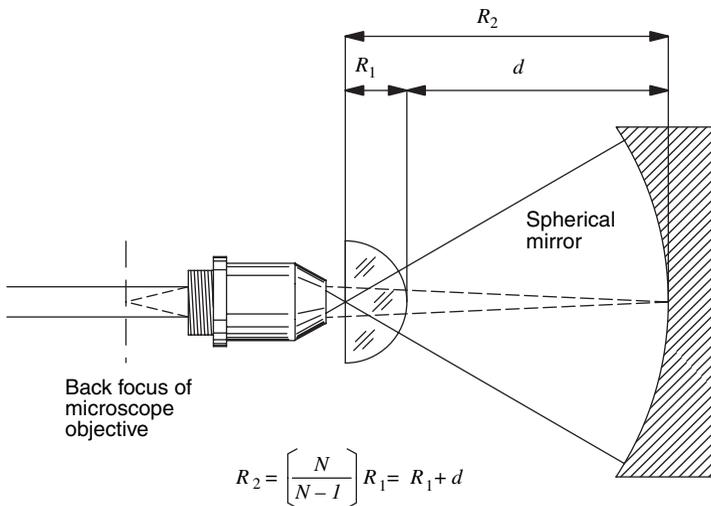


FIGURE 12.7. Dyson system to image the back focus of the microscope objective back on itself, after reflection on a concave mirror.

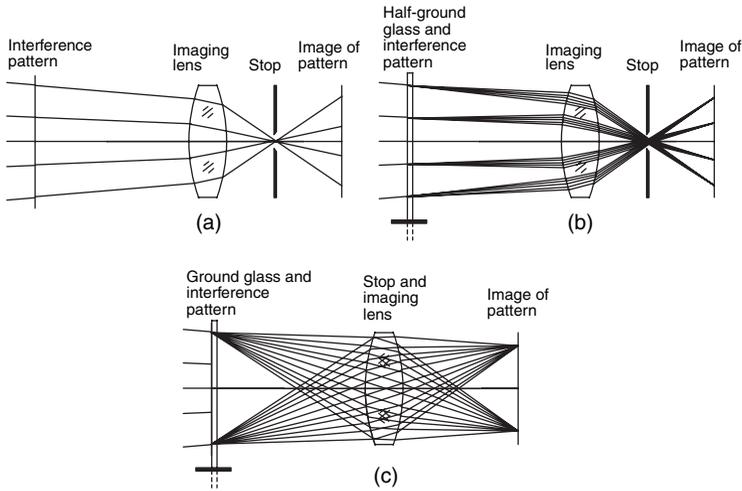


FIGURE 12.8. Imaging of the interferogram on the observation plane by means of a lens (a) without any rotating ground glass, (b) with a rotating half-ground glass, and (c) with a rotating fully-ground glass.

lens should be shifted to the lens in order to use all available light, but then the lens must be designed taking this new stop position into consideration as shown in Figure 12.8(c).

12.4. SOME NULL TESTING CONFIGURATIONS

Now some of the many possible configurations that allow a null test of an aspheric surface are reviewed.

12.4.1. Flat and Concave Spherical Surfaces

Not only aspheric surfaces are sometimes difficult to test, even spherical surfaces are difficult if the radius of curvature is too short or too large with respect to its diameter. Some null test configurations appropriate for flat or concave spherical surfaces are shown in Figure 12.9, with the relevant parameters and dimensions (Ritchey, 1904).

12.4.2. Telescope Refracting Objectives

A telescope doublet may easily be tested by autocollimation against an optical flat as shown in Figure 12.10. The flat must, of course, be as good as the rest of the system is intended to be. It is also necessary to keep the light source and the testing point as close as possible to each other to avoid astigmatism. When a system is double passed (including the case of the testing of a single concave surface at the center of curvature), as in this configuration, any antisymmetric wavefront aberration, like

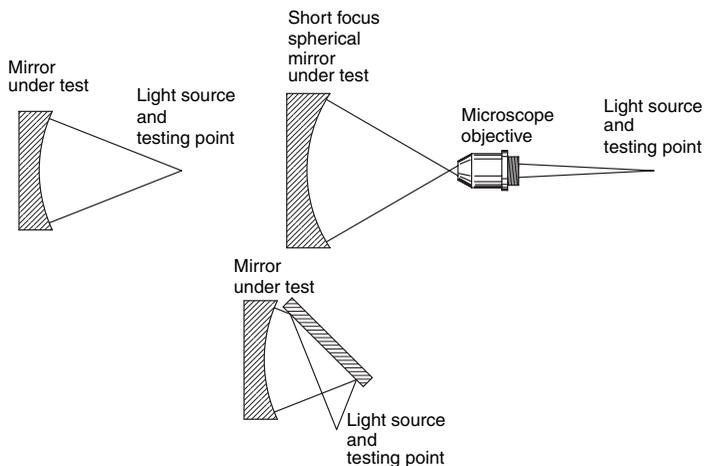


FIGURE 12.9. Testing concave spherical and flat mirrors.

coma, transverse color, and distortion, is canceled due to the the symmetry of the system.

12.4.3. Concave Paraboloidal Surfaces

If the paraboloid is not too large popular configurations use autocollimation with an optical flat as shown in Figure 12.11. The amount of spherical convexity or concavity permissible in the flat mirror used in autocollimation tests was shown by Burch (1938) to be given by

$$\delta = 64 \left(\frac{f}{D} \right)^2 \frac{\epsilon}{4Q - 0.5} \tag{12.4}$$

where f is the effective focal length and D is the aperture diameter of the system under test. The symbol δ represents the depth in fringes of the spherical concavity or convexity of the “flat” mirror, and ϵ represents the tolerance, also in fringes, of the

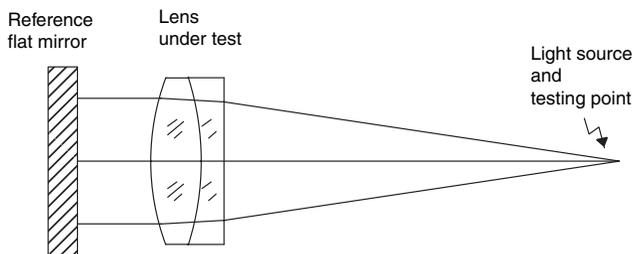


FIGURE 12.10. Testing a lens by autocollimation against a flat mirror.

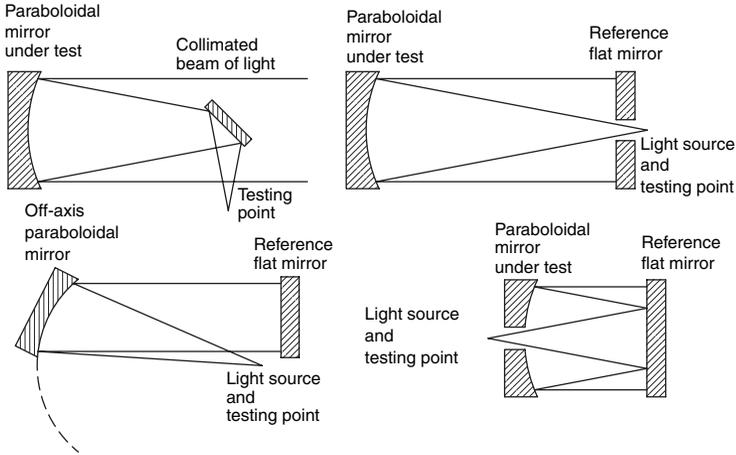


FIGURE 12.11. Testing paraboloidal concave mirrors.

zonal effect error introduced. If the system under test is refracting, the zonal error is $-2\varepsilon/(N - 1)$, where N is the refractive index. The parameter Q is defined by

$$Q = -\frac{\text{OSC}}{\sin^2 \theta} = \frac{1}{\sin^2 \theta} \left(\frac{Y}{F \sin \theta} - 1 \right) \tag{12.5}$$

where OSC is the “offense against the sine condition” and θ is the angle at which a marginal ray with height Y at the entrance pupil converges to the focus of the system. As Burch pointed out, a paraboloid and an aplanatic system are the two cases of practical interest: for these, $Q = 0.25$ and $Q = 0$, respectively, giving

$$\delta = \pm 128 \left(\frac{F}{D} \right)^2 \varepsilon \tag{12.6}$$

If the paraboloid has a large aperture with respect to its radius of curvature, a point light source may be placed at its focus. Then, the collimated beam may be examined with another paraboloid with the same diameter but much larger focal length, as proposed by Parks (1974).

12.4.4. Concave Ellipsoidal or Spheroidal Surfaces

An ellipsoidal sometimes also called a prolate spheroid surface, obtained by rotation of the ellipse about its major axis may be tested with conjugates at finite but at different distances (Kirkham, 1953) as shown in Figure 12.12(a). In a Twyman–Green interferometer, a configuration, like the one illustrated in Figure 12.13 and as suggested by Schwider (1999), can be used. Two identical lenses have to be used in both the arms of the interferometer, so that the aberration introduced by the lenses is the same.

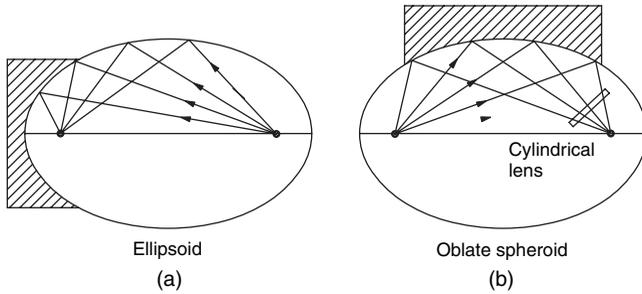


FIGURE 12.12. Testing concave ellipsoidal and oblate spheroidal mirrors.

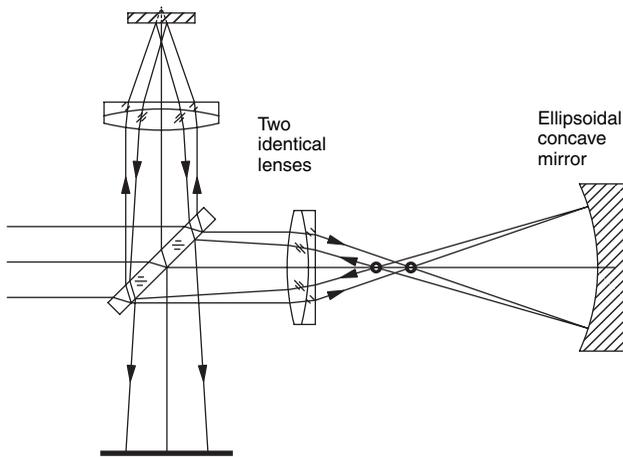


FIGURE 12.13. Testing and ellipsoidal concave mirror in a Twyman–Green interferometer.

An oblate spheroid is obtained by rotating the ellipse about its minor axis. In this case the images at the foci of the ellipse are astigmatic. Coma aberration is not present due to the symmetry of the system. A small cylindrical lens may be placed near its focus in order to correct this astigmatism if necessary (Everhart, 1966) as shown in Figure 12.12(b). Several different arrangements to test oblate spheroids using refractive compensators have been proposed, as described by Rodgers and Parks (1984).

12.5. TESTING OF CONVEX HYPERBOLOIDAL SURFACES

12.5.1. Hindle Type Tests

The testing of convex hyperboloids is very important for astronomical instruments. The most common test for these surfaces has been implemented by using a Hindle sphere as described in the following section. The problem with this method is that a

very large concave spherical surface, much larger than the surface under test, is required. Various other methods using compensators have been reported for testing convex hyperboloids, as described by Parks and Shao (1988).

We have mentioned that when testing conicoids a null test is obtained, then the center of curvature of the illuminating wavefront is at the proper focus; but since at least one of the geometrical foci is inaccessible, additional optical elements are required. A convex hyperboloidal surface can be tested with the method proposed by Hindle (1931), who showed how an autostigmatic arrangement for testing a convex hyperboloid can be implemented by retro-reflection from a sphere whose center is at the inaccessible focus of the hyperboloid as shown in Figure 12.14(a). Concave ellipsoidal surfaces can be examined with a Hindle arrangement, as shown in Figure 12.14(b). Small concave hyperboloids are also tested in a similar way (Silvertooth, 1940) as shown in Figure 12.14(c). A complete two mirror Cassegrain or Ritchey–Chrétien telescope can be tested in an autocollimating configuration devised by Ritchey if a large reference flat is available (see Fig. 12.14(d)).

Figure 12.15 shows Hindle arrangements for testing convex paraboloids and convex prolate spheroids. In addition to the Hindle sphere, a collimator is required to test a convex paraboloid. To test an ellipsoid (prolate spheroid), a lens designed for conjugates at finite distances is needed to provide a beam that converges to one of the spheroidal foci.

Although the Hindle and related tests for convex conicoids affords a stigmatic retroreflected image, its implementation is often impractical because keeping the obscuration inherent in the test within permissible bounds results in a prohibitively large spherical mirror. In the case of a hyperboloid of diameter D , the aperture of the Hindle sphere D_H is given by the relation

$$D_H = \frac{D(m+1)}{mr+1} \quad (12.7)$$

where r is the permissible obscuration ratio, and m is the magnification of the hyperboloid for its stigmatic conjugates. Thus, a 0.25-m hyperboloid with $m = 10$ and permissible $r = 0.2$ requires a 0.92-m Hindle sphere.

A modification of the Hindle test that avoids this difficulty was described by Simpson et al. (1974). Their arrangement is shown in Figure 12.16. The concave surface of a meniscus element serves as the Hindle sphere. This surface is half silvered so that it can be placed close to the conicoid without introducing obscuration. The radius of the convex surface can be chosen to compensate for the spherical aberration introduced by the refraction of the light beam passing through the concave surface. This testing configuration has been studied by Robbert (1979) and by Howard et al. (1983).

To test the effect of the meniscus Hindle element on the retroreflected wave, the hyperboloid is removed and the retroreflected image of S from a calibrating sphere with center at F is examined. Any significant aberration introduced by the meniscus can then be subtracted from the measurement of the figure error of the hyperboloid.

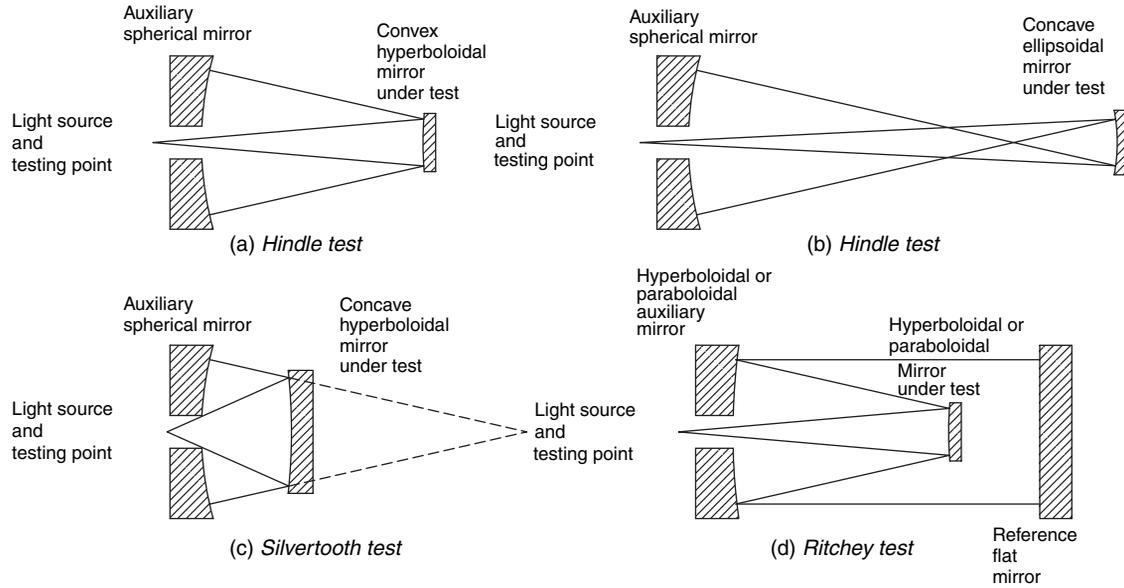


FIGURE 12.14. Hindle, Silvertooth, and Ritchey tests for hyperboloidal mirrors.

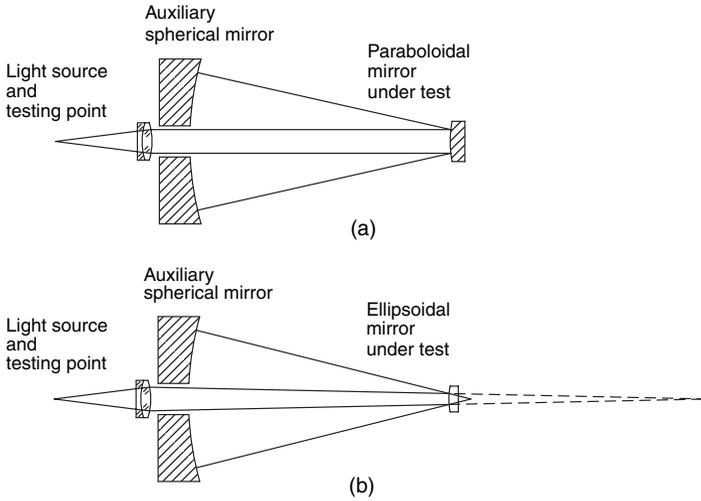


FIGURE 12.15. Testing paraboloidal and ellipsoidal convex mirrors.

A meniscus was designed to test the 0.25-m $10\times$ hyperboloid mentioned earlier. The geometrical foci of the hyperboloid were at 0.6 and 6 m. The base radius of the hyperboloid is thus 1.33333 m, and its conic constant is $K = -1.49383$. Glass of index 1.52 was chosen for the meniscus element. For a 5 cm separation of the meniscus from the hyperboloid, the radius of the concave surface is 65 cm. With a meniscus thickness of 5 cm, the radius of the convex surface that results in aberration compensation at the edge of the aperture is 66.6637 cm. The required clear aperture of the meniscus is 0.254 m. The RMS (root mean square) of the wavefront deformations as given by the OPD (optical path difference) of the retroreflected wave is 0.0016λ at $\lambda = 632.8$ nm. The stigmatic quality of the retroreflected image is thus retained in this modified Hindle test. Its use for testing convex conicoids is limited

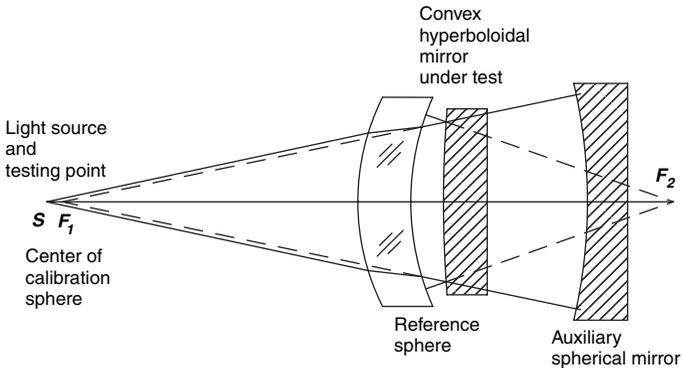


FIGURE 12.16. Simpson-Oland-Meckel modified Hindle arrangement.

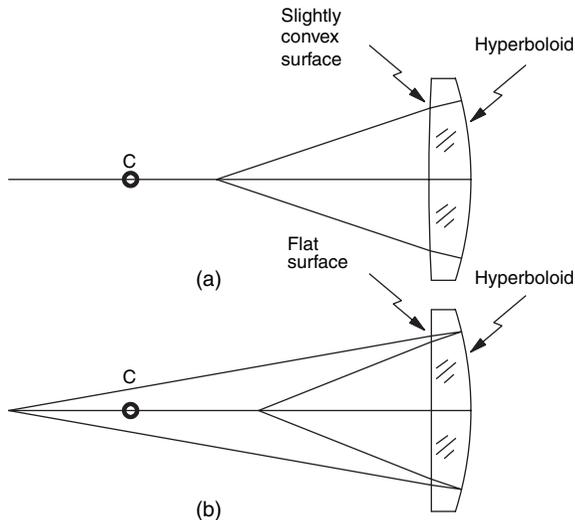


FIGURE 12.17. Meinel’s test for a hyperboloidal convex mirror (a) using equal conjugates, with the light source and the testing point at the same position and (b) using unequal conjugates. The light source and the testing point are at different positions, but they can be made to coincide if a double-pass configuration is used, by placing a small flat mirror on these points.

only by the availability of refracting material of approximately the same size as the surface being tested.

An inconvenience with the Hindle test is that a large spherical mirror is needed. Another solution has been proposed by Meinel and Meinel (1983a, 1983b) in order to test the hyperboloid from the back surface. The mirror has to be made out of fused quartz in order to have good transparency and homogeneity. There are two possible optical arrangements for this test, one is shown in Figure 12.17(a) with the light source and the testing point at the same position. The surface has to be slightly convex with a long radius of curvature. As pointed out by Meinel and Meinel, a better solution is obtained if the back surface is made flat and the spherical aberration is completely corrected by locating the light source and the testing point along the optical axis, separated some distance from each other as shown in Figure 12.17(b). Many interesting variations of this test and some others may be found in a paper by Parks and Shao (1988).

12.5.2. Testing by Refraction

Descartes discovered that a refractive conical can focus an incident collimated beam of light without any spherical aberration, if the conic constant K is equal to

$$K = -\left(\frac{n_1}{n_2}\right)^2 \tag{12.8}$$

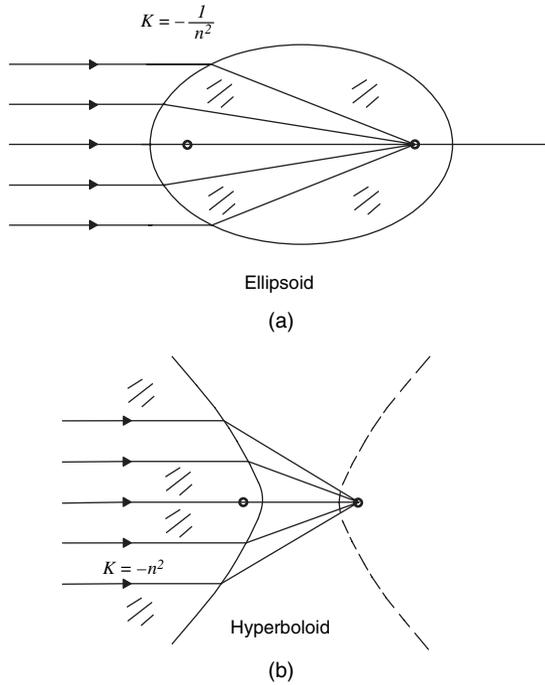


FIGURE 12.18. Cartesian configuration for an ellipsoid and a hyperboloid.

where n_1 is the refractive index of the first medium and n_2 is the refractive index of the second medium. There are two possible solutions:

- (a) If $n_2 > n_1$ and $r > 0$, the conic is an ellipsoid, sometimes called Descartes ovoid, as shown in Figure 12.18(a). If the first medium is air and the second one is glass, we have a conic constant given by

$$K = -\frac{1}{n^2} \tag{12.9}$$

and if the distance from the vertex of the ellipsoid to the second focus, which is the point of convergence, is L , we have

$$Lc = \frac{n}{n-1} \tag{12.10}$$

where c is the vertex curvature.

- (b) If $n_2 < n_1$ and $r < 0$, this is a hyperboloid, as illustrated in Figure 12.18(b). If the first medium is glass and the second one is air, the conic constant is

$$K = -n^2 \tag{12.11}$$

TABLE 12.1. Main parameter for the focusing of a collimated beam using a refractive conicoid with two common optical glasses.

	Refractive index	Ellipsoid		Hyperboloid	
	n	K	Lc	K	Lc
BK7	1.5168	-0.4347	2.9350	-2.3007	-1.9350
F2	1.6200	-0.3810	2.6129	-2.6244	-1.6129

and if the distance from the vertex of the hyperboloid to the focus, which is the point of convergence, is L , we have

$$Lc = -\frac{1}{n - 1} \tag{12.12}$$

where the vertex curvature is negative.

Table 12.1 gives the main parameters for two common optical glasses.

Using this property, a convex hyperboloid may be tested with the arrangement given in Figure 12.19, if the index of refraction of the glass is of the proper value. The glass has to be clear and homogeneous; so optical glass or fused silica have to be used. Pyrex or glassceramic materials are not adequate. Since this test is made by transmission instead of reflection, the wavefront deformations OPD and the surface deformations W are related by $W = OPD/2(n - 1)$, which is about half the sensitivity obtained in a reflective arrangement.

If the refractive index is not adequate to produce an aplanatic image, an object distance may be selected so that the image is free of primary spherical aberration. This may be done by ray tracing, but an approximate solution may be calculated which requires that the third-order spherical aberration in Eq. (16.26) be zero. By making $n_{-1} = 1$ and defining a distance $l = y/u_{-1}$, this condition is

$$(8A_1 + Kc^3) + \frac{(n+1)}{n^2} + c \left(c + \frac{1}{l} \right)^2 = 0 \tag{12.13}$$

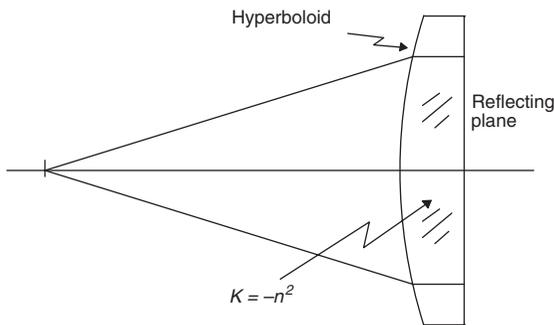


FIGURE 12.19. Testing a hyperboloid with a Cartesian configuration.

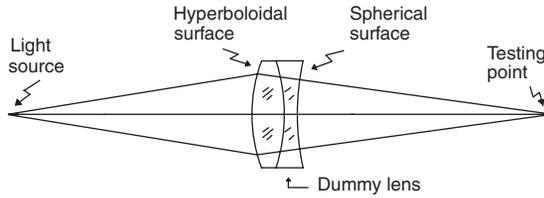


FIGURE 12.20. Testing an aspheric surface by selecting the conjugate distance that produces the minimum amount of spherical aberration.

Then, as suggested by James and Waterworth (1965), a lens with a convex hyperboloidal surface may be tested as shown in Figure 12.20. The spherical aberration of the second surface is eliminated by bringing this surface in contact with another lens using oil, whose second surface is spherical and is concentric with the testing point. In this case, however, the compensation is not perfect, since some small residual aberrations may remain. The sensitivity is about one fourth of that obtained in a reflective arrangement.

This property of the lenses with one hyperbolic convex surface can be used to test a convex surface with another procedure suggested by Bruns (1983). A convergent lens is placed in front of the convex surface under test to compensate for the spherical aberration introduced by the convex hyperbolic surface under test as illustrated in Figure 12.21. Unfortunately, no lens with spherical surfaces may correct the spherical aberration of the hyperboloid. So, Bruns uses a lens with a hyperboloidal surface. This hyperboloid in the front lens has conic constant K_L related to the selected glass refraction index. This makes the lens very simple to test when a collimated beam enters the lens through the flat face, since for this lens orientation spherical aberration is corrected. When this lens is used in reverse, that is, with the collimated beam entering the convex surface, it has a spherical aberration with the proper sign to cancel that of the convex mirror.

If the desired conic constant for the convex mirror is K_M , with radius of curvature R_M , Bruns has shown that the spherical aberration of this system is

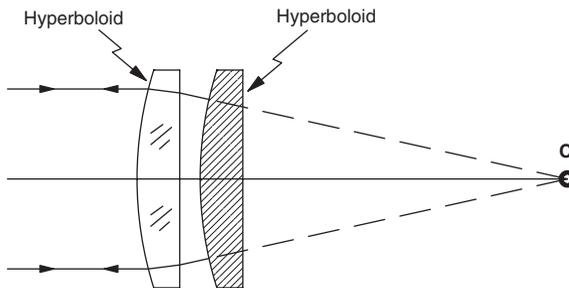


FIGURE 12.21. Testing a hyperboloidal surface by using an aspherical lens to compensate the spherical aberration.

compensated for when the refractive index n_L of the lens in front is given by the following relation:

$$K_M = -\frac{2(n_L + 1)(R_M + d)}{n_L(n_L - 1)R_M} \quad (12.14)$$

where d is the separation between the flat surface of the lens and the mirror, and an infinitely thin lens is assumed. The focal length of the lens is equal to $R_M + d$. The value of R_M is fixed, thus n_L and d must be chosen so that this relation is satisfied for the desired value of K_M .

12.6. TESTING OF CYLINDRICAL SURFACES

A concave cylindrical surface can be tested as illustrated in Figure 12.22, where the axis of the cylindrical test is in the vertical direction (Shnurr and Mann, 1981). A flat reference surface that reflects partially is located at the focus of the cylinder where the image is formed like a bright and narrow vertical line. The retroreflected wavefront is flat. It has to be noticed, however, that the wavefront is reversed about a vertical axis. Since the light is reflected twice on the surface under test, the antisymmetric mirror deformations are cancelled out while the symmetrical components are duplicated. Also, a highly spatially coherent illumination like that produced by a gas laser is needed.

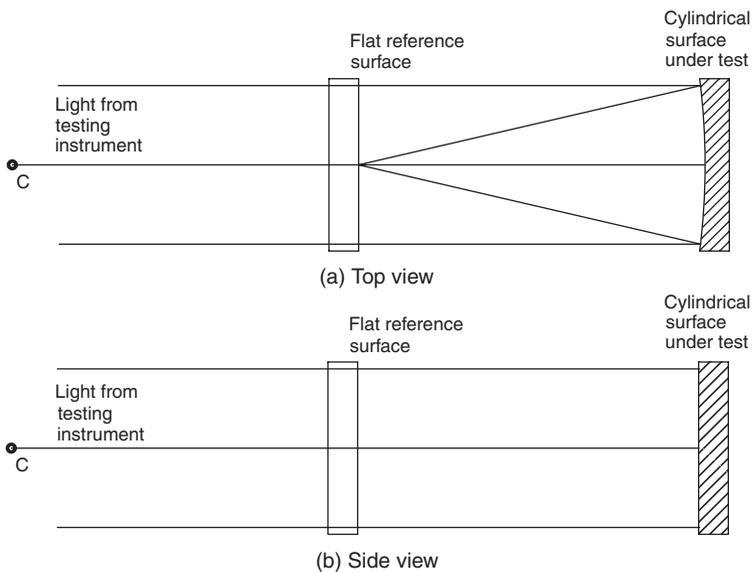


FIGURE 12.22. Testing a concave cylindrical surface in Twyman–Green interferometer using an auxiliary flat surface.

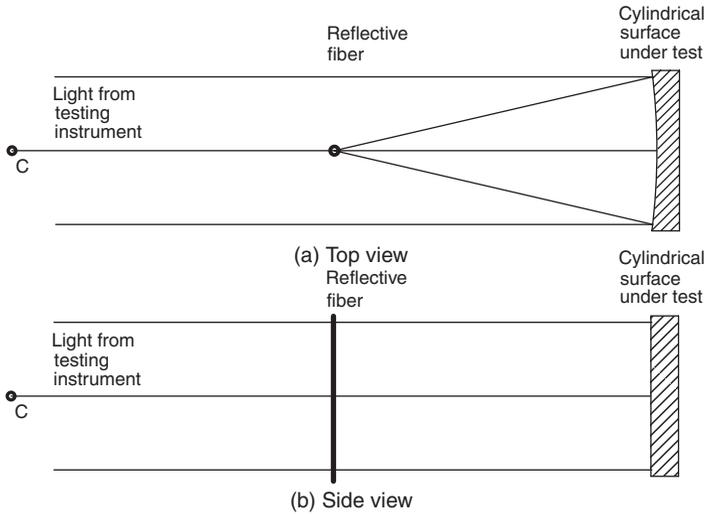


FIGURE 12.23. Testing a concave cylindrical surface in Twyman–Green interferometer using a reflective fiber.

Since the light being reflected is quite narrow, resembling a bright line, the reference mirror can be substituted by a very narrow mirror covering just the image and not the whole aperture. It must be remembered that this narrow image is not a perfect line due to the aberrations of the mirror under test. Geary and Parker (1987) and Geary (1991) substituted the narrow mirror by a thin fiber as shown in Figure 12.23. The fiber is coated to make it highly reflective and thinner than the width of the image due to the aberrations. The wavefront reflected back to the cylindrical surface on the fiber is cylindrical. Then, the aberrations are not duplicated due to the double pass, and the antisymmetrical components of the aberration are eliminated.

A cylindrical lens can also be tested using similar arrangements with a flat mirror as shown in Figure 12.24, or with a wire. The same conclusions as described in the last paragraph apply to this test.

A convex cylindrical mirror can be tested with the arrangement in Figure 12.25. The difference is that the beam from the interferometer has to be convergent. The retro-reflected wavefront is also spherical.

Another approach to test cylindrical lenses has been proposed by Lamprecht et al. (2003). The cylindrical wavefront produced by the cylindrical lens is transformed back to a plane wavefront by means of a diffractive optical element, which is made by optical e-beam lithography.

12.7. EARLY COMPENSATORS

This section on null compensators is adapted from the original chapter in earlier versions of this book, written by late Abbe Offner. The corresponding test of a

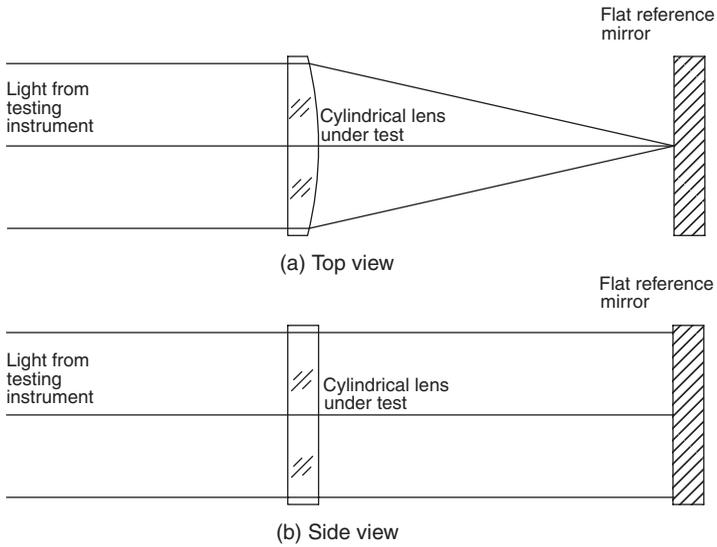


FIGURE 12.24. Testing a convergent cylindrical lens in Twyman–Green interferometer.

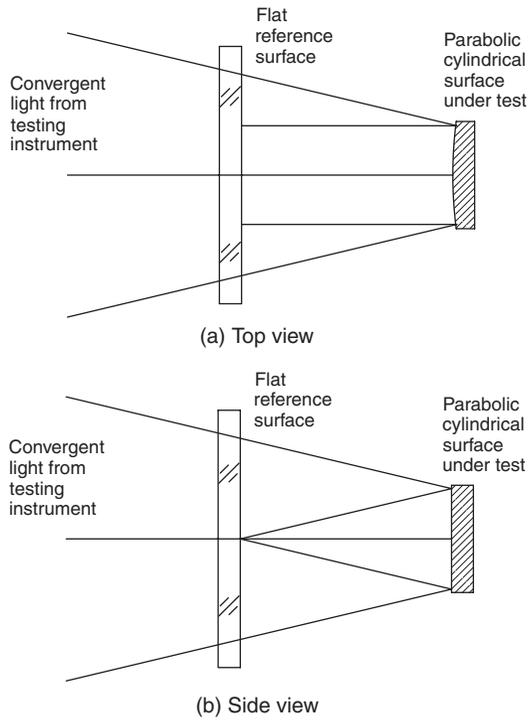


FIGURE 12.25. Testing a convex paraboloidal cylindrical surface with an auxiliary flat surface.

concave paraboloidal mirror is a lengthy and less accurate process. In this case, the figure errors must be deduced from measurements of the ray transverse aberrations (surface slopes) for a large number of zones, since a concave aspheric or conic surface tested at the center of curvature has a spherical aberration wavefront deformation given in third-order approximation by

$$W(\rho) = \frac{(8A_1 + Kc^3)\rho^4}{4} \quad (12.15)$$

where A_1 is the first aspheric deformation term, K is the conic constant, and $c = 1/R$ is the curvature. For a conic surface we may write

$$W = \frac{1}{4}KR\left(\frac{\rho}{R}\right)^4 \quad (12.16)$$

For many years the only alternative to the method of using the knife-edge test during the manufacture of a paraboloidal mirror was to test the mirror by autocollimation with the aid of an optical flat, which had to be as large and as accurately figured as the mirror being manufactured.

12.7.1. Couder, Burch, and Ross Compensators

Couder (1927) pointed out that departure from stigmatism of the image of a point source at the center of curvature of a paraboloidal mirror can be removed by interposing a small compensating lens between the image and the mirror. He used a two-element compensator in the arrangement shown in Figure 12.26. Two elements were necessary because he required a null corrector of zero total power to conveniently carry out the manufacturing process desired in his paper. To manufacture a 30-cm $f/5$ paraboloidal mirror, he used a null corrector whose aperture (scaled from his drawing) was about 4 cm.

The use of a spherical mirror beyond the center of curvature of a paraboloid to compensate for the aberrations of a paraboloid used with source and knife edge near its center of curvature was described by Burch (1936). He derived the fifth-order aberration of the null systems of this type and showed that with the two-mirror arrangement of Figure 12.27, the residual aberration of the paraboloid is less than one-fortieth of a wave for paraboloids as fast as $f/5$ and with apertures up to 80 cm,

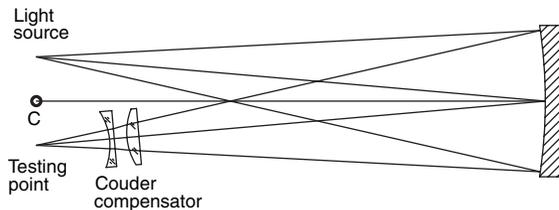


FIGURE 12.26. Couder two-element compensator.

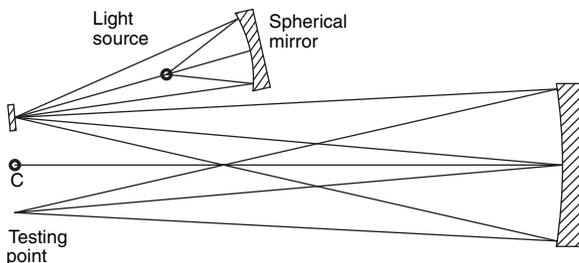


FIGURE 12.27. Burch two-mirror compensator.

when the compensating mirror aperture is one fourth of that of the paraboloid. For larger apertures and/or faster paraboloids, he suggested figuring the convex mirror with a down edge that departed from the base sphere by an eighth power law so that the seventh-order spherical aberration could be balanced. He calculated that an asphericity of about 2.8λ would be required to compensate for the aberrations of the 5-m $f/3.33$ Mount Palomar mirror.

A simple third-order solution for a refracting compensator was also published by Burch (1938). The refractor was a planoconvex lens of focal length f and refractive index n , used in the arrangement shown in Figure 12.28, in which its plane surface is reflecting. For an paraboloid of base radius R , the third-order aberration of the image at the center of curvature is balanced when $f = Rn^2/(n - 1)^2$, so that for $n = 1.52$ the lens has an aperture of about one eighth of that of the paraboloid. Burch expected the residual aberration with this null corrector to be negligible for paraboloids of aperture ratio not exceeding $f/8$. He added, “Anyone with an aptitude for analytical optics or for ray-tracing could earn the gratitude of practical opticians by calculating the secondary aberrations for this and other kinds of compensating lens systems.” This plea was answered 30 years later by Holleran (1968).

During the manufacture of the 5-m $f/3.33$ Mount Palomar mirror, a 25-cm diameter compensator was used to form a stigmatic retro-reflected image near the center of curvature of the mirror (Ross, 1943). The arrangement used is shown in Figure 12.29.

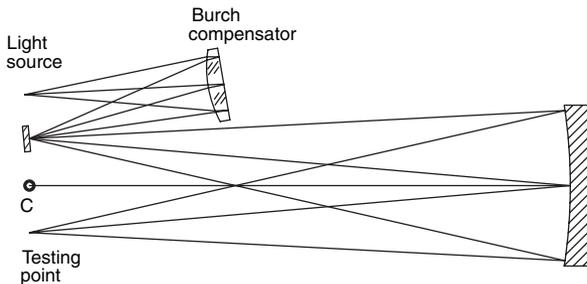


FIGURE 12.28. Burch planoconvex compensator.

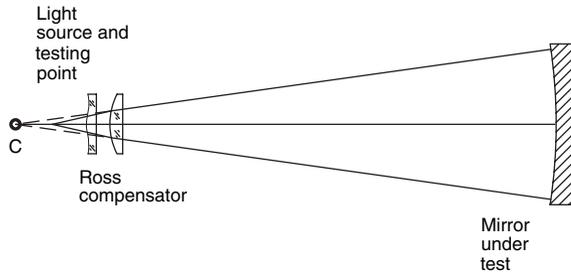


FIGURE 12.29. Ross aspheric compensator used during the fabrication of the 5-m Mount Palomar mirror.

To obtain a degree of compensation such that the residual zonal aberration resulted in a disk of confusion that was small compared to that caused by the atmosphere, Ross found it necessary to add an aspheric corrector plate to a refractive element with spherical surfaces, which by itself balanced the spherical aberration of the paraboloidal mirror. The retroreflective arrangement used by Ross has the advantage that it is coma-free and therefore insensitive to departures of the source and knife edge from the axis of the system. Moreover, since the compensator is used twice, it has to contribute only one-half as much aberration as is required in Couder's arrangement.

12.7.2. Dall Compensator

The planoconvex lens of Burch is a convenient and an easily used solution to the problem of making a null corrector for a paraboloidal mirror of moderate aperture. However, with this method a planoconvex lens used in the manufacture of a paraboloidal mirror can serve as a compensator only for other paraboloidal mirrors of the same focal length.

Dall (1947, 1953) noted that, since the spherical aberration of a lens is a function of its conjugates, the same planoconvex lens can be used as a compensator for many paraboloids. Dall employed the arrangement of Figure 12.30, which is quite similar to the arrangement used by Couder. The major difference is that, the Dall compensator is placed in front of the light source and not in front of the observer's eye. The reason for this is that if it is used on the convergent beam, the convergence of the

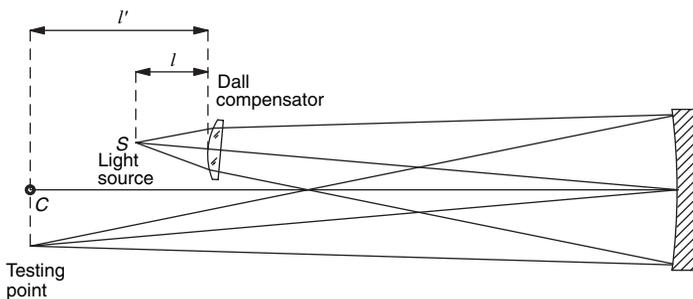


FIGURE 12.30. Dall planoconvex compensator.

beam may become so large that an observation of the whole surface under test may not be possible.

The Dall compensator is normally used in an off-axis arrangement, with the light passing only once through the compensator. Two double-pass configurations, with the divergent as well as the convergent beams passing through the compensator, may also be used: (a) exactly on axis, using a beam splitter and (b) slightly off axis. Then, the compensating lens needs to compensate only half the total aberration on each pass, making it less strong. Then, an additional factor of 2 has to be used in front of Eq. (12.15). Another advantage of the double-pass symmetrical arrangement is that it is coma free. Thus, any coma introduced by the off-axis lateral displacement of the light source and the observer or any misalignment of the lens compensator does not introduce any coma. Unfortunately, the double-pass arrangement may be used only if the radius of curvature of the surface under test is very large compared with its diameter.

Dall found that proper choice of the short conjugate of the lens provides adequate compensation if the ratio of the radius of curvature R of the paraboloid to the focal length f of the lens is between 10 and 40. The relation required to balance the third-order aberration of the parabola at its center of curvature is

$$\frac{R}{f} = \frac{1}{2}(m-1)^2 \left[\frac{n^2(m-1)^2}{(n-1)^2} + \frac{(3n+1)(m-1)}{n-1} + \frac{3n+2}{n} \right] \quad (12.17)$$

where m is the ratio of the long conjugate distance l' to the short conjugate distance l , and n is the index of refraction of the plano-convex lens. (The sign convention is such that in the Dall arrangement $m > 1$.)

The Dall compensator has been widely used, especially by amateur telescope makers. The degree of compensation that can be attained with this extremely simple null corrector is illustrated by the following example.

A Dall compensator is desired for a 0.6-m $f/5$ paraboloidal mirror. Taking $m = 2$, $N = 1.52$, and $F = 3$ m in Eq. (12.17), we find $R/f = 11.776$, which is within the bounds prescribed by Dall. The compensator specifications are then $f = 50.950$ cm, $l = -25.475$ cm, and $l' = -50.950$ cm. With this null corrector the computed root-mean-square (RMS) departure from the closest sphere [RMS optical path difference (OPD)] of the wavefront that converges to form an image of the light source is 0.048λ at $\lambda = 632.8$ nm. A paraboloid fabricated to give a null test with this compensator would have an RMS figure error of 0.024λ . The Strehl intensity resulting from this figure error is 0.91. The diameter of the plano-convex lens required to achieve this compensation is about one-twelfth of that of the paraboloidal mirror.

The paraboloid of the examples is about the largest for which a Dall compensator is adequate. Since the arrangement suggested by Dall is not coma-free, the light source must be located accurately on the axis of the convex lens, and this axis should be directed to pass through the pole of the paraboloid.

Practical instructions for making and using a Dall null tester are given in two papers by Schlauch (1959) and by Stoltzmann and Hatch (1976). By restricting the

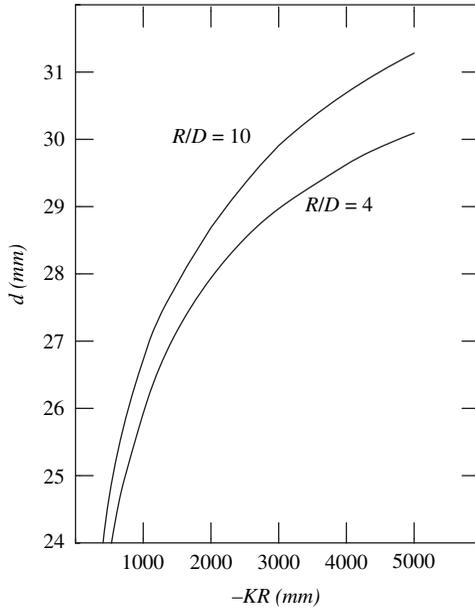


FIGURE 12.31. Distance d from the light source to the Dall compensator lens in Figure 12.32 for different values of the product of the radius of curvature times the conic constant.

refractive index of the plano-convex lens to 1.52, the computation of Eq. (12.17) can be avoided with the help of a curve in Schlauch’s paper, which is adapted from the one published by Dall (1953).

A Dall lens made with BK-7 glass and using red light can also be calculated with the curves in Figure 12.31. The radius of curvature is 25 mm and the thickness is 5 mm. A difference with Figure 12.30 is that it has been assumed that the plane surface of the lens and the testing point (knife edge) are in the same plane, as illustrated in Figure 12.32. Assuming only the presence of third-order spherical aberration, from Eq. (12.16) we may say that the distance d from the light source to the vertex of the convex surface of the lens has to be a function only of the product KR . This is true only for small apertures (large ratio R/D). If the aperture is large (ratio R/D small), the fifth-order spherical aberration is present in the Dall lens, and it

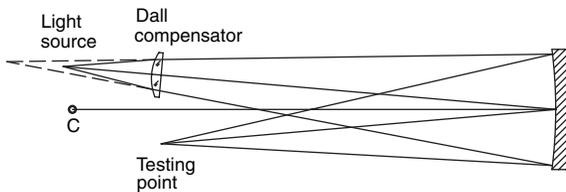


FIGURE 12.32. A Dall compensator. The testing point is assumed to be at the same plane as the flat surface of the Dall lens.

has to be partially compensated with some overcompensation of the third-order aberration. Then, a zonal spherical aberration remains uncorrected. This remaining aberration is not a severe problem if the aperture is small or if there is a large central obscuration, as pointed out by Rodgers (1986).

A Dall compensator can be modified to be coma-free by using the planoconvex lens on axis like in the retro-reflective arrangement of Ross. The relation to be satisfied with this arrangement is a modification of Eq. (12.17) in which the factor $1/2$ is eliminated since the lens is traversed twice. With this arrangement, for example, the compensator for the $0.6-mf/5$ paraboloid has $R/f = 23.552$, $f = 25.475$ cm, $l = -12.737$ cm, and $l' = -25.475$ cm. The diameter of the planoconvex lens is then that of the paraboloid. The residual figure error of the paraboloid that gives a null test with this arrangement is exactly the same as that computed in the preceding example. The retroreflective arrangement has the advantage, however, that since it is coma-free, it is not affected by small departures of the light source from the axis of the planoconvex lens.

In an interesting variation of the Dall compensator proposed by Puryayev (1973), an afocal meniscus whose concave surface is conicoidal is substituted for Dall's planoconvex lens in the autostigmatic arrangement used by Ross (Fig. 12.29). For an afocal meniscus,

$$r_1 - r_2 = \frac{d(n-1)}{n} \quad (12.18)$$

where r_1 is the radius of the concave surface of the meniscus, r_2 is the radius of its convex surface, d is its thickness, and n is its index of refraction.

The third-order value of the conic constant K of the concave surface required to compensate for the aberration of a paraboloid of radius R is

$$K = \frac{R}{(N-1) \left(\frac{r_2}{r_1} \right)^2 l} \quad (12.19)$$

where l is the distance from the light source to the meniscus. (The sign convention is such that K is negative.)

With the same 20-cm-diameter meniscus compensator, Puryayev achieved compensation for any paraboloidal or near-paraboloidal surface whose focal length does not exceed 24 m and whose aperture ratio does not exceed 1:4. The maximum residual wave aberration of the retro-reflected wave for any paraboloid in this range is about $\lambda/2$ at 632.8 nm. This residual can be computed and taken into account when the figure of the test mirror is being determined.

12.8. REFRACTIVE COMPENSATORS

As described earlier (Burch, 1936; Ross, 1943), compensation for the spherical aberration of a paraboloid or other aspheric concave mirror can be achieved to any

desired degree of accuracy by the incorporation of an aspheric element in the null corrector. This method is limited to cases in which the figure of the aspheric element can be ascertained with an accuracy better than that desired for the aspheric mirror. Primary mirrors that are to be incorporated into diffraction-limited space-borne optical systems are now required to have RMS figure errors as small as one-hundredth of the wavelength of visible light. It is therefore desirable that the components of a null corrector for these mirrors be spherical or flat so that their figure errors can be measured to the required accuracy.

In the design of his null corrector, Ross found that the farther he put the lens from the center of curvature of the mirror, the less residual aberration there was when the compensation was exactly at the center and the edge. This is so because, although the longitudinal spherical aberration S of the normals to the paraboloid follows the simple law $S = y^2/2R$, where y is the distance of the normal from the axis of the paraboloid and R is its radius, additional terms of a power series would be required to describe the same spherical aberration distribution in a coordinate system with its origin at the null corrector. For a null corrector in contact with the paraboloid, the compensating spherical aberration would be described by the same simple law. Unfortunately, this null corrector would be as large as the paraboloid.

12.8.1. Refractive Offner Compensator

Offner (1963) pointed out that a small lens that forms a real image of a point source at the center of curvature of a paraboloid, in combination with a field lens at the center of curvature that images the small lens at the paraboloid, is optically equivalent to a large lens at the paraboloid. The use of a field lens in this way was first suggested by Schupmann to control secondary spectrum (Schupmann, 1899; Offner, 1969). This kind of compensator has been widely used to test astronomical optics of many different characteristics (Sasian, 1988).

With a field lens that images the compensating lens c at the paraboloid in Offner's arrangement, the spherical aberration of the compensating lens must follow the same law for aperture as do the normals to the paraboloid.

This restriction on the compensating lens is not necessary. All that is required is that lens c provide sufficient third-order spherical aberration to compensate for that of the normals of the paraboloid. The power of the field lens (and thus the location of the image of lens c) is then varied to minimize the high-order aberration.

To balance the third-order aberration of the normals to a conicoidal mirror with conic constant K and base radius R , a planoconvex lens of focal length f and index of refraction n must satisfy the relation

$$-\frac{KR}{f} = (1 - m)^2 \left[\frac{n^2(1 - m)^2}{(n - 1)^2} + \frac{(3n + 1)(1 - m)m}{n - 1} + \frac{(3n + 2)m^2}{n} \right] \quad (12.20)$$

where m is the ratio l'/l (Fig. 12.33). (The conicoid must have $K < 0$ if the aberration of its normals are to be balanced by the spherical aberration of a planoconvex lens. The sign convention is such that $m < 0$.)

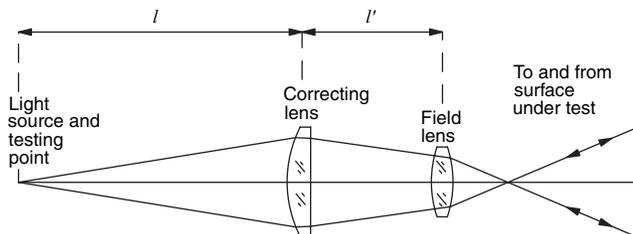


FIGURE 12.33. Refracting compensator with field lens.

Since the mirror under test in general has a large spherical aberration, the field lens does not have an ideal position. It may be placed between the mirror and its caustic, inside the caustic region or outside the caustic. Any of these possible locations produce good compensations, with slight variations, as explained by Sasian (1989). Different techniques for designing this compensator have been explored by several authors, among others, by Landgrave and Moya (1987) and by Moya and Landgrave (1987).

The main purpose of the field lens is to avoid fifth-order spherical aberration, but another equally important consequence of its presence is that the wavefront at the mirror under test is imaged on the plane where the aberrations are observed, that is, at the compensating lens.

The importance of the field lens in the Offner arrangement is evident from the example of the design of a null corrector for a 1-m $f/4$ paraboloid using planoconvex lenses with refractive index 1.52. The quantity m that results in the desired convergence angle of the retroreflected wavefront is first chosen.

Choosing a convergence of $f/12$ leads to $m = -0.6667$. The value of the focal length of the compensating lens required to balance the third-order aberration of the paraboloid normals is then seen from Eq. (12.20) to be 20.9115 cm, since $K = -1$ and $R = 800.0$ cm. The conjugates for $m = -0.6667$ are $l = -52.2772$ cm and $l' = -34.8532$ cm. The retroreflective system formed by placing the source at the long conjugate of this lens and the paraboloidal mirror center of curvature at its short conjugate is corrected for third-order spherical aberration, but has fifth-order lateral spherical aberration of -0.0205 mm. The RMS OPD of the retroreflected wavefront is 0.23λ at $\lambda = 632.8$ nm.

A field lens of focal length 33.3976 cm at the center of curvature of the paraboloid forms an image of the compensating lens on the paraboloidal mirror. With this addition, the sign of the fifth-order spherical aberration is reversed, its value being $+0.0207$ mm. The RMS OPD of the retroreflected wave is increased slightly to 0.26λ .

The focal length of the field lens that minimizes the high-order spherical aberration is found to be 66.8900 cm. With this field lens, the computed RMS OPD of the retroreflected wavefront is reduced to 0.0003λ , a value well below what can be measured. The diameter of the compensating lens required for this degree of correction is one-twentieth that of the $f/4$ paraboloid.

In Eq. (12.20), it is assumed that the curved surface of the planoconvex field lens is at the center of curvature of the paraboloid. It is sometimes convenient to move the field lens to a position close to, but not at the center of curvature. In this case the field lens introduces an additional magnification m_f . The condition for compensation of third-order aberration then becomes

$$-\frac{KR}{f_f^2} = (1 - \bar{m})^2 \left[\frac{n^2(1 - \bar{m})^2}{(n - 1)^2} + \frac{(3n + 1)(1 - m)\bar{m}}{n - 1} + \frac{(3n + 2)\bar{m}^2}{n} \right] \quad (12.21)$$

where $\bar{m} = m_f^2/m$.

Like Ross's arrangement, the retroreflective arrangement of Offner is inherently coma-free so that the correction of the retroreflected wavefront is maintained when the source is near the axis but not exactly on it.

The high degree of stigmatism that can be achieved by the use of the Offner corrector has led to its application for the quantitative assessment of the figures of large aperture concave conicoidal mirrors. For this purpose, the retroreflected wavefront can be compared with a reference sphere in a spherical wave interferometer (Houston et al., 1967). A multipass version of the spherical wave interferometer in which the retroreflected wavefront and the reference sphere are optically conjugate (Heintze et al., 1967) is particularly useful for making this measurement with the greatest accuracy. With this interferometer, which has been given the acronym SWIM (spherical wave interferometer multibeam), individual points of a wavefront have been measured with an accuracy of 0.003λ (private communication).

12.8.2. Shafer Compensator

This compensator is a triplet, designed by Shafer (1979), so that the following three conditions are satisfied:

1. For a certain distance from the light source to the compensator, the spherical wavefront from the light source preserves its spherical shape after passing through the compensator. Then, positive or negative compensations may be achieved by displacing the system along the optical axis.
2. The system is afocal (effective focal length infinite), so that the angles with respect to the optical axis of the light rays entering the system are preserved after passing through the compensator.
3. The angular magnification of the afocal system, (lateral magnification for a near object) must be equal to $+1$. Thus the apparent position of the light source does not change when moving the compensator along the optical axis.

A system with a negative lens between two positive lenses, as in Figure 12.34, is appropriate for positive conic constants (oblate spheroids) of any magnitude and negative conic constants (paraboloids or hyperboloids) of moderate magnitude as

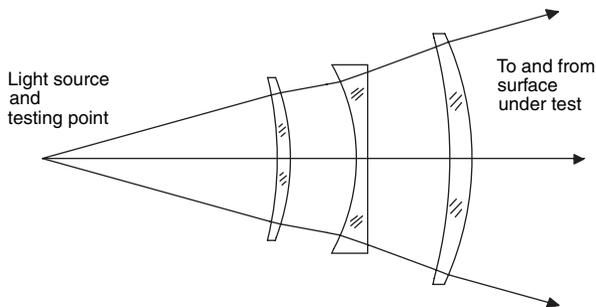


FIGURE 12.34. Shafer compensator with all lenses made with BK-7 glass.

shown in the graph in Figure 12.35. This graph is inverted when a positive lens is placed between two negative lenses.

12.8.3. General Comments about Refracting Compensators

The main refracting compensators are the Couder, Dall, Shafer, and the Offner compensators. They have the following properties in common:

1. All compensate the spherical aberration but with a different degree of perfection. The best in this respect is the Offner compensator.
2. All can be used in a single or double pass, but the Couder and Dall compensators are typically used in a single pass, generally in front of the light source, while the Shafer and Offner compensators are normally used in double pass.

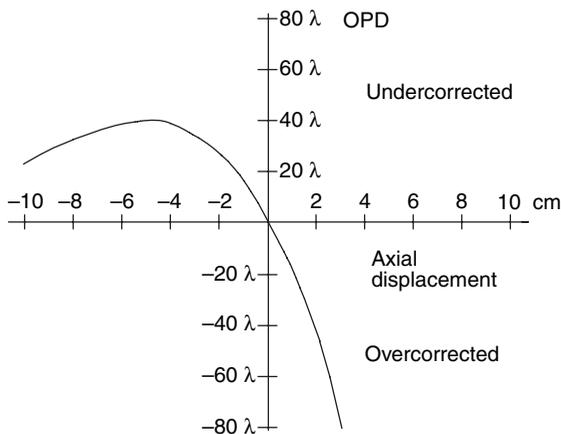


FIGURE 12.35. Wavefront aberration due to third-order spherical aberration vs. lens motion for Shafer compensator and an $f/2$ system with a wavelength of 632.8 nm.

3. The chromatic aberration is not corrected in any of these compensators, hence monochromatic light has to be used. One possibility is to use laser light and another is to use a color filter. It is suggested to use a red filter close to the eye or image detector (after the compensator) so that the wavefront shape is not affected by this filter.

4. The amount of spherical aberration correction depends very critically on the axial position of the compensating lens; so unless its position is very accurately measured, we can never be sure about the exact value of the conic constant of the concave surface under test. This value has to be measured by some other test that does not use any compensator, like the Hartmann test. However, the general smoothness of the surface can be easily determined only with a compensator.

5. In a double-pass configuration, any lateral displacement of the light source and the observer in opposite direction and by equal amounts, with respect to the optical axis, does not introduce any coma. It is assumed, however, that the optical axis of both the conic surface under test and the compensating lens coincide; otherwise some coma is introduced.

12.9. REFLECTING COMPENSATORS

The weak point in making measurements with the Offner null corrector is the difficulty in measuring the index variations of the nulling element to the required degree of accuracy. In the example described (Offner, 1963), the thickness of the 4.5-cm-diameter compensating lens was 1.05 cm. An average index difference of 3×10^{-7} along the paths of the two rays that traverse this lens twice results in an optical path difference of $\lambda/100$ at $\lambda = 632.8$ nm. Faster aspheric mirrors with larger aperture require larger compensating lens diameters and thicknesses. For these, even smaller average index differences result in optical path errors of this magnitude. Fabrication and qualification of large glass elements to this degree of homogeneity is not feasible at present.

These difficulties can be avoided by substituting spherical mirrors for the planoconvex refracting compensating element of Figure 12.33. The figure errors of such elements can be determined with great accuracy. A small field lens can be retained since it is possible to select small pieces of glass with satisfactorily small index variations.

It is well known (Burch, 1936) that the axial aberration of a spherical mirror used at a magnification other than -1 can be used to compensate for the aberration of the normals of a concave conicoid with negative conic constant. The high degree of compensation achieved with the Offner refractive null corrector can also be obtained by a reflecting compensator used with a field lens at the center of curvature of the conicoid, as shown in Figure 12.36. As in the refractive version, the radius of the nulling mirror R_N and its conjugates l and l' are chosen to balance the third-order aberration of the normals to the conicoid of radius R_c and the conic constant K . The power of the field lens is then

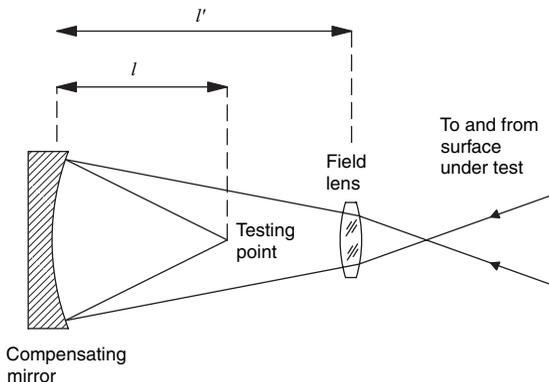


FIGURE 12.36. Single mirror compensator with field lens.

varied to minimize the high-order aberration. The relations to be satisfied are as follows:

$$R_N = -\frac{8KR_c}{(m^2 - 1)^2} \tag{12.22}$$

$$l = \frac{(1 - m)R_c}{2} \tag{12.23}$$

$$m = \frac{l}{l'} = -\frac{2\eta_c}{\eta_N} \tag{12.24}$$

where $2\eta_c$ and $2\eta_N$ are the f numbers of the beam at the center of curvature of the conicoid and at the retroreflected image, respectively.

The ratio of the diameter of the conicoid D_c to that of the null mirror D_N can be computed from the relation

$$\frac{D_c}{D_N} = \frac{(m^2 - 1)^2}{4K(m - 1)} \tag{12.25}$$

Some values of the ratio of the diameter of a paraboloid to that of a single mirror compensator, computed from Eq. (12.25), are listed in Table 12.2. Diameter ratios of more than 10 require values of $-m$ greater than 4. A practical limit on the value of m is set by the resultant value of h_N , the f number at the retroreflected image, which is inversely proportional to m [Eq. (12.24)]. If the compensated wavefront is to be examined interferometrically without transfer optics, the interferometer must be capable of handling f/η_N beam. The single-mirror compensator thus requires large compensating mirrors if η_c , the f number of the conicoidal mirror, is small. The permissible residual aberration of the compensated image also must be taken into account, and in some cases this results in a value of A_c greater than that set by the lower limit on η_N .

TABLE 12.2. Diameter ratio and magnification, single mirror compensator.

Magnification	Diameter ratio
-3	4
-3.5	7
-4	11.25
-4.5	16.8
-5	24

For example, a one-mirror compensator was designed for a 3-m $f/2.45$ hyperboidal primary mirror of a proposed Ritchey–Chrétien system. The conic constant of the mirror was -1.003313 . A value of -4.9 was chosen for m , resulting in an $f/1$ beam at the retroreflected image. The third-order mirror specifications computed from Eqs. (12.18) to (12.21) and the specifications of the optimized design are shown in Table 12.3. The focal length of the field lens that minimizes the high-order aberration is 55.4849 cm. The computed RMS OPD of the retroreflected wavefront is 0.009λ at $\lambda = 632.8$ nm. If a smaller residual had been required, a smaller value of $-m$ would have been chosen for the compensator. The resulting compensator mirror would then have been larger, and the convergence angle of the retroreflected beam would have been smaller.

12.9.1. Reflective Offner Compensator

Although the single-mirror compensator of Figure 12.36 is optically the least complicated of the reflecting compensators, practical implementations require an additional element, such as the folding flat of Figure 12.37, to make the retroreflected image accessible. The quality of the flat must, of course, be comparable to that of the spherical mirror.

The same number of accurately fabricated optical components is required in the two-mirror compensator shown in Figure 12.38. The in-line arrangement facilitates accurate alignment and provides an accessible retroreflected image.

With the two-mirror compensator, the central portion of the aspheric mirror cannot be observed because of the holes in the nulling mirrors. These null correctors should be designed so that the obscured portion of the aspheric mirror in the null test is no larger than the obscured portion of the aspheric mirror in actual use.

The third-order design of a two-mirror compensator is affected by the value of the obscuration ratio at each of the mirrors. The following equations apply when the

TABLE 12.3. One-mirror null corrector for 3-m $f/2.45$ hyperboloid.

Type of design	m	R_N	l	A_N	η_N
Third order	-4.9	22.2849	65.7404	13.42	1.000
Optimized	-4.7	22.2849	65.7518	13.5	1.044

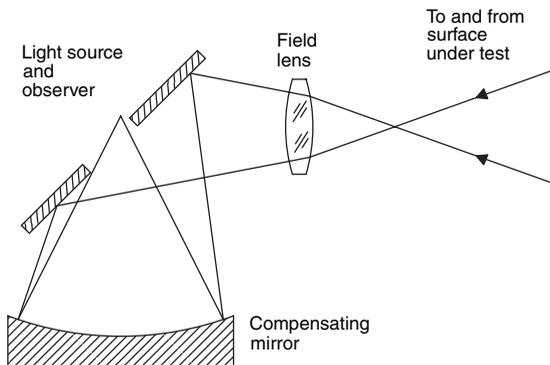


FIGURE 12.37. Practical implementation of single-mirror compensator.

obscuration ratio r is the same at the two mirrors. As in the case of the single-mirror null corrector, the parameters and apertures are functions of a magnification. For the two-mirror compensator, the magnification is that from the intermediate image to the image at the center of curvature of the aspheric mirror, defined by the relation

$$m_1 = -\frac{2\eta_c}{\eta_1} \tag{12.26}$$

where η_1 is the f number at the intermediate image. The ratios of the diameters of the apertures of the two nulling mirrors D_1 and D_2 to the aperture of the conicoid D_c when the third-order aberration of the conicoid normals is compensated for by that of the null corrector can be computed from the following relations:

$$\frac{D_c}{D_1} = \frac{1}{4K} (m_1 + 1) [m_1^2 (1 + 2r - r^2) - 2m_1 (1 - r) - 2] \tag{12.27}$$

$$\frac{D_c}{D_2} = \frac{D_c m_1 r - 1}{D_1 m_1 + 1} \tag{12.28}$$

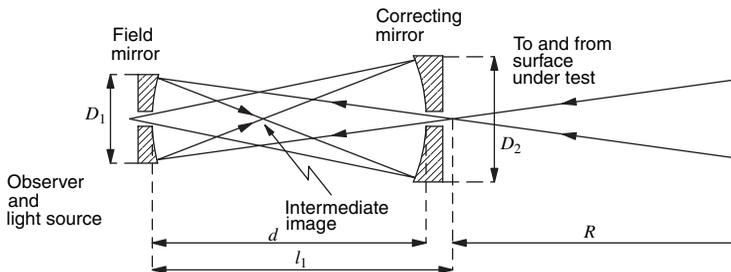


FIGURE 12.38. Two-mirror compensator with field lens.

The other relations required for the third-order design are as follows:

$$R_1 = \frac{4D_1\eta_c}{1 - m_1} \quad (12.29)$$

$$R_2 = \frac{4D_2\eta_c}{m_1(2 - r) + 1} \quad (12.30)$$

$$l_1 = 2\eta_c D_1 \quad (12.31)$$

$$d = \eta_1(D_1 + D_2) \quad (12.32)$$

$$\eta_N = -\frac{2\eta_c}{m_1(1 - r) + 1} \quad (12.33)$$

where l_1 is the distance from the center of curvature of the conicoid to the field mirror, d is the distance between the two mirrors, and η_N is the f number at the retroreflected image.

The way in which the apertures of the two mirrors and the f number of the retroreflected image vary as a function of the magnification m for two values of the obscuration ratio can be seen in Table 12.3. Comparison with Table 12.1 shows that for a given magnification, the larger mirror of the two-mirror compensator is approximately of the same size as the single-mirror compensator. However, the difference between the magnitudes of m_1 and $2\eta_c\eta_N$ in Table 12.3 indicates that, for a given maximum size compensator element, the convergence angle at the retroreflected image with the two-mirror compensator is approximately one half of that with a single-mirror compensator.

The degree of compensation attainable with the two-mirror compensator is extremely high as shown by the following example. A compensator was required for a 3-m $f/1.5$ paraboloid that was to be used with an obscuration ratio of 0.3. A spherical wave interferometer that could accommodate convergence angles up to $f/1.2$ was available. The values $r = 0.25$ and $m_1 = -4$ were chosen to give the safe value $\eta_N = 1.5$. Equations (12.27) and (12.28) led to the acceptable values $A_1 = 14.82$ cm and $A_2 = 27.22$ cm. The parameters obtained from Eqs. (12.29) to (12.33) are listed as third-order design parameters in Table 12.4. The field lens required to optimize this compensator is a meniscus lens of refractive index 1.519, thickness 0.5 cm, and radii 14.619 cm (convex) and 71.656 cm (concave), with the convex surface facing the paraboloid. With this field lens and the slight modifications of the other parameters shown in Table 12.4, the RMS OPD of the retroreflected wavefront is 0.009λ at $\lambda = 632.8$ nm. Had a smaller residual aberration been

TABLE 12.4. Two-mirror null corrector for 3-m $f/1.5$ paraboloid.

Type of design	m_1	R_1	R_2	l_1	d	A_2	η_N
Third order	- 4.00	17.7778	22.2222	44.4444	27.7778	22.22	1.50
Optimized	- 3.99	17.7776	22.2227	44.3868	27.7789	21.34	1.52

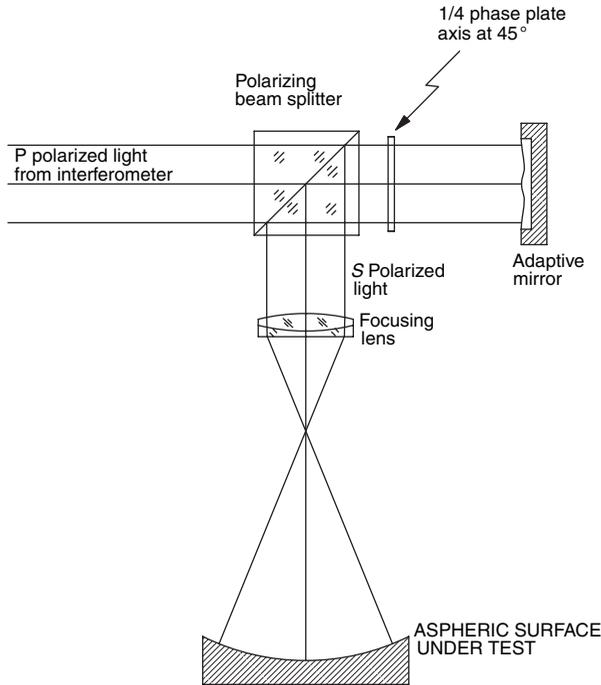


FIGURE 12.39. Aspheric compensator using an adaptive mirror.

required, the value of m_1 would have had to be reduced. This would have resulted in larger values of A_1 , A_2 , and η_N .

12.9.2. Reflective Adaptive Compensator

Another interesting reflective compensator has been described by Tiziani et al. (2001). They use an optical configuration where the reflective compensator is an adaptive mirror whose shape can be changed as desired by controlling it with a computer. The system in Figure 12.39 adds the reflected wavefront from the aspheric surface under test to the shape of the adaptive mirror. The light beam is reflected twice on the adaptive mirror, so that its shape is added twice to the aspheric wavefront from the system under test. The flexible mirror is an aluminum coated nitride membrane that is attracted electrostatically by electrodes on the back of the membrane. This membrane can be deflected up to about $20\ \mu\text{m}$ PV, which corresponds to about 80 wavelengths due to the double reflection on this membrane.

12.10. OTHER COMPENSATORS FOR CONCAVE CONICOIDS

The success of the simple small compensators described in the preceding sections results from a fortunate combination of conditions.

1. The concave aspheric mirrors by themselves transform the divergent wave from a point source to a convergent one though aberrated wave.
2. The greater part of the aberration introduced by the aspheric is of low order.
3. The sense of the aberration is opposite to that introduced by a concave mirror or a simple convex lens.

The first condition must be met if the null corrector is to be smaller than the aspheric mirror being tested. The second condition makes it possible to get good compensation with a single element of convenient form. The third condition makes it possible to use a simple relay lens that provides a position for a field lens. Concave spherical mirrors can be used for compensators only when this condition is met.

A small null corrector of the same general form can be designed for any concave mirror whose surface is generated by rotating a conic section about its major axis.

Concave prolate spheroids do not require compensators since their geometrical foci are accessible and, as is the case with all conicoids that have geometrical foci, their imagery is stigmatic when these are the conjugates. However, when one of the geometrical foci is at a large distance from the mirror, it may be more convenient to perform a null test at the center of curvature with one of the null correctors described in Sections 12.2–12.4.

An oblate spheroid, such as the one that is used as the primary mirror of a Wright–Schmidt system, does not satisfy condition 3. Nevertheless, a null test of the modified Dall type can be obtained by substituting a plano-concave lens for the Dall plano-convex lens (Figure 12.40). Since the curved surface of this lens faces the oblate spheroid, the third-order solution is formally the same as that for the Offner plano-convex compensator. The parameters and conjugates of the plano-convex lens required to balance the third-order aberration of the normals to the prolate spheroid can thus be obtained by the use of Eq. (12.20). The quantity m is the ratio l'/l . In this case it is positive and has a value less than 1. The effect of the choice of m on the ratio of the size of the oblate spheroid to that of the null corrector is shown in Table 12.5; the values were computed for $N = 1.52$ and $K = 1$. The implementation of the value

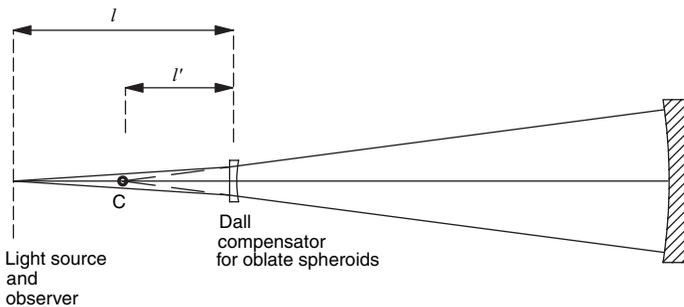


FIGURE 12.40. Modified Dall compensator for oblate spheroids.

TABLE 12.5. Aperture of planoconcave null lens compensator for oblate spheroid ($K = 1, n = 1.52$).

m	Diameter asphere/Diameter compensator
0.0	8.54
0.1	7.13
0.2	5.88
0.3	4.77
0.4	3.8

$m = 0$ requires the addition of a collimator to put the source optically at an infinite conjugate. The resulting arrangement is shown in Figure 12.41.

A null corrector of this type was designed for a 0.6-m aperture $f/5$ oblate spheroid with conic constant $K = 1$. The departure of this aspheric from the base sphere is equal in magnitude but opposite in sense to that of the paraboloid used as an example in Section 12.2. The values $n = 1.52$ and $m = 0$ were chosen. The focal length of the planoconcave lens, obtained from Eq. (12.20), is 70.2216 m. The RMS OPD of the retroreflected plane wave is 0.033λ at $\lambda = 632.8$ nm. The diameter of the null lens is 0.7 cm.

Holleran (1963, 1964) described a null test for concave conicoids that has the virtue that no auxiliary optical elements need be manufactured. The surface to be tested is made level and is immersed in a liquid that forms a planoconvex lens in contact with the surface under test. In the simplest form of the test, a pinhole light source and knife edge are placed at a distance d above the plane surface of the liquid. For liquid lens of thickness t and refractive index n ,

$$d = \frac{R}{n} - t \tag{12.34}$$

where R is the vertex radius of curvature of the conicoid. The retroreflected image is corrected for third-order spherical aberration if

$$n^2 = 1 - \frac{KR}{R - t} \tag{12.35}$$

where K is the conic constant of the mirror.

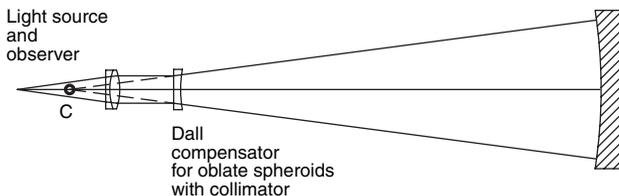


FIGURE 12.41. Compensator for oblate spheroid used with a collimator.

The accuracy of the test is very good for shallow curves of moderate aperture. The peak-to-valley departure ΔW of the surface from the desired conicoid when Eq. (12.35) is satisfied and $t \ll R$ is

$$\Delta W \approx \frac{KR}{41.5\eta^6} \quad \text{waves} \quad (12.36)$$

In Eq. (12.36), η is the f number of the conicoidal mirror, R is its radius in millimeters, and the wavelength is 632.8 nm. The peak-to-valley error of a 0.5-m-diameter $f/3$ paraboloid figured to give a perfect null by this test is 0.10 wave. Decreasing the f number to 2.5 for a mirror having the same aperture increases the figure error to 0.25 wave. The error of an $f/2.5$ mirror reaches 0.10 wave for an aperture of 0.2 m.

The immersion test in this form can also be applied to convex aspherics by observing them through a plane back surface. In this case the optical material replaces the immersion liquid. Puryayev (1971) analyzed an extension of this method in which an immersion fluid is placed above the plane surface. Since the liquid must extend to the retro-reflected image, this extension is practical only for small elements. Puryayev's equations reduce to those of Holleran when the immersion fluid is air.

A related test for convex hyperboloids, described by Norman (1957), makes use of the fact that a planoconvex lens forms a stigmatic image of a collimated source on its axis if the convex surface is a conicoid with eccentricity equal to the refractive index of the material of which the lens is made. The autocollimated image of a point source at a distance of one focal length from the convex surface of lens and reflected from its plane surface, or from a flat mirror parallel to its plane surface, can then be examined to determine the figure of the convex surface. As in the Holleran and Puryayev tests, the range of conicoids that can be tested by this method is limited by the range of refractive indices available. The test can be applied to hyperboloids with magnifications between 3.5 and 5 when the range of glass indices is restricted to 1.5 to 1.8. Holleran (1966) showed that a spherical back surface can provide a null test compensator for a concave oblate spheroid when the latter is tested from the back, through the spherical surface.

12.11. INTERFEROMETERS USING REAL HOLOGRAMS

Many different experimental setups can be used for the holographic testing of optical elements. Because a hologram is simply an interferogram with a large tilt angle between the reference and object wavefronts, holographic tests can be performed either with standard interferometers or with setups having a larger angle between the object and reference beams. Figures 12.42 and 12.43 show interferometers that can be used with a hologram for testing of a concave mirror.

The hologram is made in a plane conjugate to the test mirror. Once the hologram is made, it can be replaced in the same location and reconstructed by illuminating with a plane wave and by imaging onto a viewing screen. When the object beam is blocked and the reference mirror is tilted so that the plane reference wave interferes with the

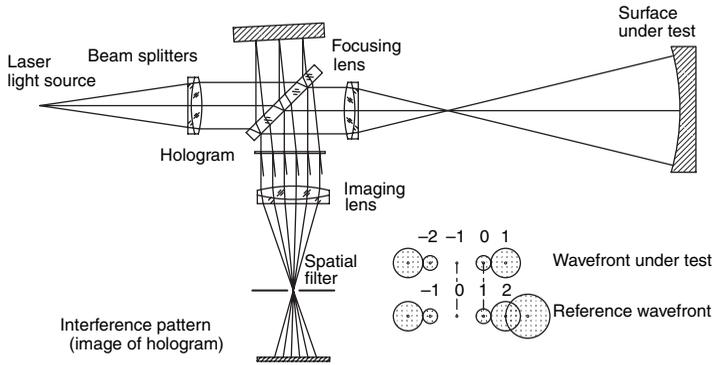


FIGURE 12.42. Twyman-Green interferometer with a holographic compensator.

first-order diffraction from the hologram, the wavefront due to the mirror will be reconstructed. Because several diffraction orders are produced by the hologram, it is usually necessary to select one of the diffraction orders using a spatial filter. The imaging lens and spatial filter are necessary only for the reconstruction of the hologram.

Holograms can be recorded on photographic plates, thermoplastic materials, or in photorefractive crystals. Photographic plates provide the highest resolution; however, they require a lot of chemicals for processing and unless they are processed in situ, they are hard to replace in the correct location for real-time techniques (Biedermann, 1975). Thermoplastic materials provide up to 1000 lines/mm resolution and can be erased and reprocessed hundreds of times (Leung et al., 1979; Friesem et al., 1980). They also have a very fast turnaround time. Photorefractive crystals have a lot of potential as a high-resolution recording medium (Lam et al., 1984; Urich and Hesselink 1988); however, the optical setup is more complex, and getting a high-quality crystal is not as easy as getting other recording materials. A very promising recording medium is the use of high-resolution detector arrays or charge-coupled device (CCD) cameras to record a hologram directly.

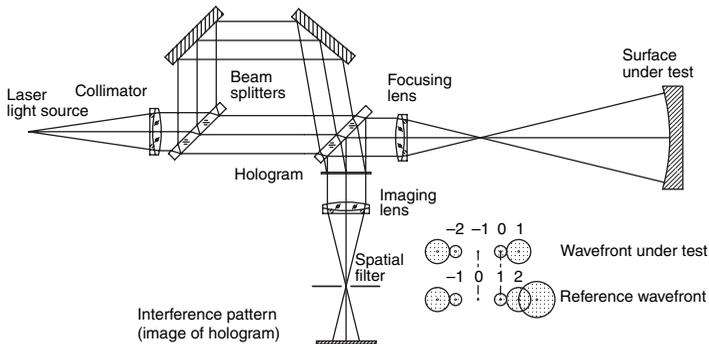


FIGURE 12.43. Mach-Zehnder interferometer with a holographic compensator.

12.11.1. Holographic Wavefront Storage

Sometimes it is convenient to holographically store a wavefront produced by an optical system and analyze the wavefront later without the test system present (Hildebrand et al., 1967; Hansler, 1968).

Care must be taken to ensure that the reconstructed wavefront is identical to the wavefront used to record the hologram. Errors are possible from differences between the reconstructing geometry and the recording geometry, recording material deformation, and aberrations introduced by the recording-material substrate. Errors introduced because of the differences in the recording and reconstructing geometries are greatly reduced if the reference wavefront is collimated and the object wavefront is collimated as well as possible. Collimated wave fronts are particularly important if the reconstruction wavelength is different from the recording wavelength. Recording-material deformation can change the shape of the recorded interference fringes and thereby change the shape of the reconstructed wavefront. It is possible to keep the root-mean-square (rms) wavefront error less than $\lambda/40$ by using Kodak 649F photographic plates (Wyant and Bennett 1972). Glass substrates used for photographic plates generally show optical-thickness variations of at least one fringe per inch. For wavefront storage, this magnitude of error is not acceptable. This problem can be solved by either putting the hologram in a fluid gate or index matching the two surfaces to good optical flats. Thickness variations can also be minimized by sending both beams through the hologram to cancel the errors.

Stored holograms can be used to test for symmetry in optical components (Greivenkamp, 1987). A hologram is made of the test surface, and then the test surface is rotated with respect to the hologram. The fringes from the interference between the stored wavefront and the wavefront produced by the rotated test surface will correspond to symmetry deviations in the test surface. Care must be taken not to translate the test surface while it is rotated because unwanted fringes will affect the test.

12.11.2. Holographic Test Plate

If an ideal optical system is available, the wavefront produced by the system can be stored holographically and used to test other optical systems. This is very similar to the use of a test plate in the testing of optical components (Hansler, 1968; Pastor, 1969; Snow and Vandewarker, 1970; Lurionov et al., 1972; Broder-Bursztyn and Malacara, 1975). After processing the hologram, it is placed in its original position and the master mirror is replaced with a test mirror. The wavefront stored in the hologram is interferometrically compared with the wavefront produced by the mirror under test. The secondary interference between the stored wavefront and the test wavefront should be recorded in a plane conjugate to the exit pupil of the test surface. If the hologram is made in a plane conjugate to the exit pupil of the mirror under test, the amount of tilt in the resulting interferogram can be selected by simply changing the tilt of the wavefront used in the hologram reconstruction process. If the hologram is not made in a plane conjugate to the exit pupil of the test surface, tilting the reference beam

not only adds tilt to the interferogram, it also causes a displacement between the image of the exit pupil of the test surface and the exit pupil stored by the hologram.

The holographic test plate interferometer can also be thought of in terms of moiré patterns (Pastor, 1969). Interference fringes, resulting from the wavefront stored in the hologram and the wavefront coming from the optics under test, can be regarded as the moiré pattern between the interference fringes recorded on the hologram plate (formed by the wavefront produced by the master optics and a plane wave) and the real-time interference fringes formed by the wavefront under test and a plane wavefront. The contrast in this moiré pattern is increased with spatial filtering by selecting only the wavefront produced by the mirror under test and the diffraction order from the hologram, giving the stored wavefront produced by the master optics.

In addition to the error sources already mentioned, there can be error due to improper positioning of the hologram in the interferometer. Any translation or rotation of the hologram will introduce error. If the hologram is made conjugate to the exit pupil of the master optical system, the exit pupil of the system under test must coincide with the hologram. If the test wavefront in the hologram plane is described by the function $\phi(x, y)$, a displacement of the hologram by a distance Δx in the x direction produces an error

$$\Delta\phi(x, y) \approx \frac{\partial\phi(x, y)}{\partial x} \Delta x \quad (12.37)$$

where $\partial\phi/\partial x$ is the slope of the wavefront in the x direction. Similarly, for a wavefront described by $\phi(r, \theta)$, the rotational error $\Delta\theta$ is given by

$$\Delta\phi(r, \theta) \approx \frac{\partial\phi(r, \theta)}{\partial\theta} \Delta\theta \quad (12.38)$$

Phase-shifting techniques can be used to measure the phase of the secondary interference fringes by placing a phase shifter in the reference beam of the interferometer and shifting the phase of the secondary interference fringes (Hariharan et al., 1982; Hariharan, 1985). Because the secondary interference fringe spacing corresponds to one wavelength of OPD between the stored wavefront and the live test surface wavefront, a $\pi/2$ phase shift of the fringes for the test surface will cause a $\pi/2$ phase shift in the secondary interference fringes. The calculated phase surface will correspond to the difference between the master optical component and the test optical component. To ensure that the fringes actually correspond to the test surface, the hologram must be made in the image plane of the test surface, and the hologram plane must be imaged onto the detector array when the phase measurement is performed.

12.12. INTERFEROMETERS USING SYNTHETIC HOLOGRAMS

When master optics are not available to make a real hologram, a computer generated (or synthetic) hologram (CGH) can be made (Pastor, 1969; Lee, 1970, 1974;

MacGovern and Wyant, 1971; Wyant and Bennett, 1972; Schwider and Burow, 1976; Caulfield et al., 1981; Ono and Wyant, 1984; Dörband and Tiziani, 1985; Beyerlein et al., 2002; Reichelt et al., 2004; Pruss et al., 2004). A CGH is a binary representation of the actual interferogram (hologram) that would be obtained if the ideal wavefront from the test system is interfered with a tilted plane wavefront. The test setup is the same as that for a real hologram used as a holographic test plate. CGHs are an alternative to null optics when testing aspheric optical components.

12.12.1. Fabrication of Computer-Generated Holograms (CGHs)

To make a CGH, the test setup must be ray traced to obtain the fringes in the hologram plane that result from the interference of the tilted plane wave and the wavefront that would be obtained if the mirror under test were perfect. Just like a real hologram used as a test plate, the CGH should be made in a plane conjugate to the exit pupil of the system under test. These fringes are then represented as a binary grating, commonly having a 50% duty cycle. Methods for calculating these fringes are outlined by Wyant and Bennett (1972), Lee (1974), and Arnold (1989). A procedure that encodes the fringes as a series of exposure rectangles is discussed by Leung et al., (1980). The process of breaking fringes into rectangles or polygons reduces the amount of computer storage necessary and the time needed to plot the CGH. The procedure used to make a CGH can be employed for any general optical system as long as all the optics in the interferometer are known and can be ray traced.

A typical CGH is shown in Figure 12.44. An example of an interferometer used to test a steep aspheric optical element is shown in Figure 12.45. If the deviation of the test surface from a sphere is substantial, the marginal rays (not the paraxial rays) will follow a different path back through the divergent lens after they have reflected off the test surface. When the surface under test is properly imaged at the observation

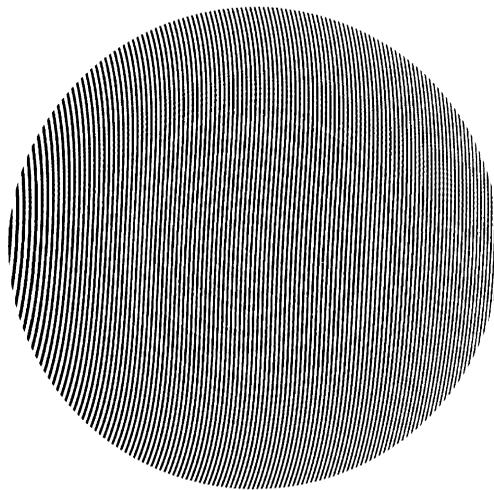


FIGURE 12.44. Computer generated hologram.

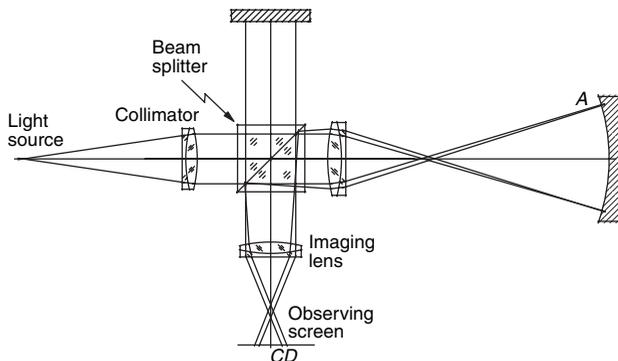


FIGURE 12.45. Aspheric test surface in interferometer showing rays not following same path after reflection off the test surface.

plane, the marginal rays will arrive at the observation plane at the same point where the corresponding ray following the path through the reference mirror arrives, as when there is no aspherical deformation, even if the path followed by the ray is different. By Fermat's principle, the optical path from any point on the surface under test to the observation plane is the same for all possible paths, automatically correcting the error. In conclusion, when there is no proper imaging of the surface being tested over the detector, there will be an additional aberration added to the final interferogram. However, in this case, when the entire system is ray traced, the hologram can be designed to correct these extra aberrations if a null test is performed. Another important consideration is that the test element may deviate the rays so much that the light reflected from the test surface will not get back through the interferometer. In this case, a partial null lens must be used to ensure that light will get back through the system. This is discussed in more detail in the section on the combination of CGHs with null optics.

Once the fringes are calculated, they are either plotted directly on a holographic substrate, or plotted and photographically reduced onto a holographic substrate. The techniques of plotting have been substantially improved over the years. Early work utilized pen plotters to make an enlarged version of the hologram that was then photographically reduced to the appropriate size (MacGovern and Wyant, 1971; Wyant and Bennett, 1972; Wyant and O'Neill, 1974). The large format enabled a high-resolution CGH to be formed. However, problems due to plotter irregularities such as line thickness, pen quality, plotter distortion, and quantization caused errors in the reconstructed wavefront. Nonlinearities inherent in the photographic process and distortion in the reduction optics caused further degradation. With the advent of laser-beam recorders, resolution improved due to machine speed and an increased number of distortion-free recording points (Wyant et al., 1974). The most recent advances in the recording of CGHs have been made using the electron beam (e-beam) recorders used for producing masks in the semiconductor industry (Emmel and Leung, 1979; Leung et al., 1980; Arnold, 1989). These machines write onto photoresist deposited on an optical quality glass plate and currently produce the highest quality

CGHs. Patterns with as many as 10^8 data points can be produced in a hologram of the desired size. Typical e-beam recorders will write a 1-mm area with a resolution of $0.25\ \mu\text{m}$. Large patterns are generated by stitching a number of 1-mm scans together. Errors in this technique are due to aberrations in the electron optics, beam drift, instabilities in the controlling electronics, and positioning of the stepper stage. For in-line holograms a thermochemical method with selective oxidation of a Cr coated substrate has been used by Burge et al. (1994). Then, a special writing machine with a rotary air-bearing spindle rotates the substrate to selectively oxidize some rings with an Ar^+ laser. After writing the pattern, the substrate with its coating is immersed in a liquid that removes all non oxidized Cr.

Many of the errors in CGHs are reproducible and can be compensated for in the software controlling the recorder (Chang and Burge 1999 and Arnold and Kestner 1995). Plotter errors can be evaluated by generating a hologram that is composed of straight lines in orthogonal directions forming a grid (Wyant et al., 1974). This test hologram is then illuminated with two plane waves as shown in Figure 12.46 to interfere the $+N$ and $-N$ orders. Deviations of the fringes from straight lines will correspond to errors in the plotting process. The resulting aberration in the interferogram is $2N$ times that of the first order.

12.12.2. Using a CGH in an Interferometer

A CGH test is performed by interfering the test wavefront with a reference wavefront stored in the hologram. This entails overlapping the zero-order test beam and the first-order reference beam from the hologram in the Fourier plane of the hologram. The test can also be performed by interfering the minus first-order test beam with the zero-order reference beam to compare the two plane waves instead of the two aspheric wavefronts. When the test wavefront departs from the reference wavefront, fringes corresponding to the difference between the wavefronts appear. In the Fourier plane of the hologram, the zero- and first-order diffracted spots of the reference wavefront will overlap the minus first- and zero-order diffracted spots of the test wavefront when the interferometer is correctly aligned. Both outputs yield the same interferogram. Spatial filtering can be used to improve the fringe contrast if the tilt of the plane reference wavefront used for the hologram is large enough.

The CGH compensator can be used off axis as well as in line, with different advantages and disadvantages, as we will show in the next sections.

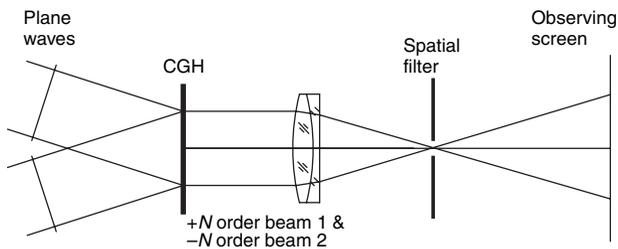


FIGURE 12.46. Test setup with $\pm N$ orders of hologram interfering to test the quality of a CGH plotter.

Also, the position where the CGH is located can be different. There are two possibilities, one is to place the hologram in the testing space where only the wavefront under test passes through the hologram. A second possibility is to place it in the observation space, where both interfering wavefronts pass through the hologram. Let us analyze these two cases with some detail. A CGH is sensitive to the same errors as real holograms. Because of this, the CGH should be placed in the interferometer, so that thickness variations in the hologram substrate have no effect on the results. The importance of this will depend on where the CGH is located.

(a) If the hologram is placed in the observation space, both the reference wavefront and the wavefront under test pass through the hologram. Each wavefront generates its own diffracted wavefronts. Proper spatial filtering should be performed so that the interference pattern is generated with the diffracted reference wavefront (aspheric) and the wavefront under test or with the nondiffracted reference wavefront (flat) and the compensated wavefront under test (nearly plane). Since the CGH is traversed by both interfering wavefronts, any small imperfection in the glass plate of the CGH is unimportant.

In this case, the superposition of the hologram and the interference pattern can be considered as a moiré pattern. If this pattern is spatially filtered, the interferogram representing the wavefront deviations with respect to the aspheric surface (not with respect to a plane) is obtained. If desired, this procedure can even be carried out by the superposition of two transparencies of the interferogram and the hologram, as illustrated in Figure 12.47.

(b) If the hologram is placed in the interferometer test space only the wavefront under test will pass through the hologram, but it will pass twice through it. Let us represent the diffracted beams after the first pass by their order of diffraction number ($\dots -2, -1, 0, 1, 2 \dots$). Since all these beams come back to the hologram after reflection on the aspheric surface under test, they can now be represented by a pair of numbers, the first one being the order of diffraction on the first pass and the second one the order of diffraction on the second pass. These beams are ordered as follows:

$$\begin{array}{ccccccc}
 (-1, -1) & & (-1, 0) & & (-1, 1) & & \\
 & & (0, -1) & & (0, 0) & & (0, 1) \\
 & & & & (1, -1) & & (1, 0) & & (1, 1)
 \end{array}$$

Each column here corresponds to a different diffracted angle. Proper spatial filtering should then be performed to isolate and observe only the desired beams, for example $(1, 0)$ with $(0, -1)$ or $(0, 1)$ with $(1, 0)$. Since only the wavefront under test passes through the hologram, the CGH must be made in a good glass plate, with flat faces and homogeneous refractive index, otherwise its aberrations will be added to the diffracted wavefronts.

In addition to the error sources associated with a real hologram, a CGH has additional error sources due to plotter distortion, incorrect hologram size, and photoreduction distortion if the hologram is photographically reduced in size. These

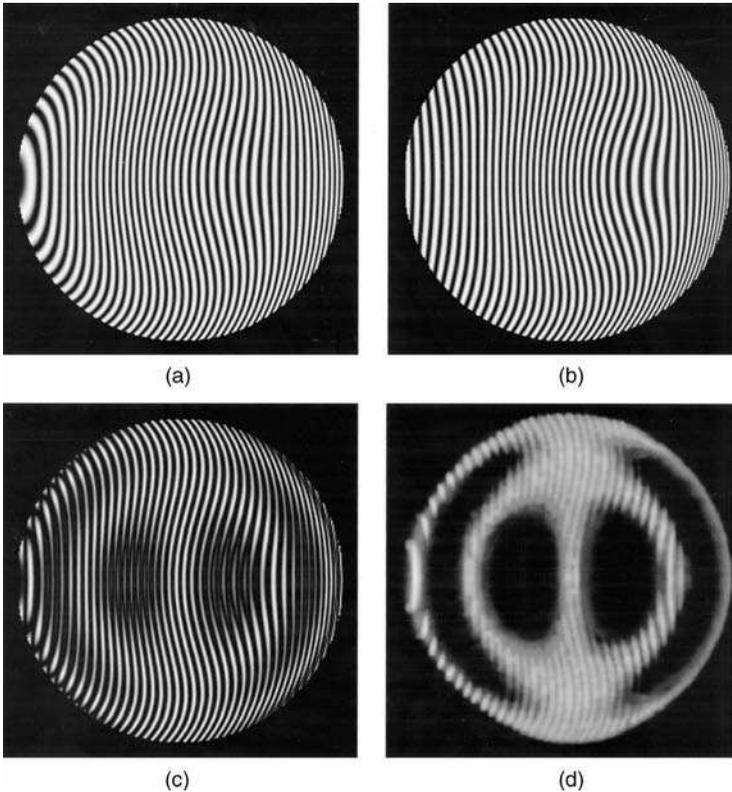


FIGURE 12.47. Moiré analysis of an interferogram with a large tilt (linear carrier). (a) Interferogram to be analyzed, (b) ideal computer generated interferogram, (c) superposition of the interferogram with the ideal reference interferogram and, (d) low pass filtered interferogram combination.

errors are proportional to the maximum slope of the departure of the test wavefront from a spherical or plane wavefront. These errors can be minimized when the test wavefront is calculated relative to the spherical wavefront, which minimizes the slope of the test wavefront departure from a spherical wavefront. Errors due to photographic reduction can be eliminated by writing a hologram of the correct size directly onto a glass substrate using an e-beam recorder.

One source of error is incorrect hologram size. If the aberrated test wavefront in the plane of the hologram is given by $\phi(r, \theta)$, a hologram of incorrect size will be given by $\phi(r/M, \theta)$, where M is a magnification factor. The error due to incorrect hologram size will be given by the difference $\phi(r/M, \theta) - \phi(r, \theta)$ and can be written in terms of a Taylor expansion as

$$\begin{aligned} \phi\left(\frac{r}{M}, \theta\right) - \phi(r, \theta) &= \phi\left[r + \left(\frac{1}{M} - 1\right)r, \theta\right] - \phi(r, \theta) \\ &= \left[\frac{\partial\phi(r, \theta)}{\partial r}\right] \left(\frac{1}{M} - 1\right)r + \dots \end{aligned} \quad (12.39)$$

where terms higher than first order can be neglected if M is sufficiently close to 1 and a small region is examined. Note that this error is similar to a radial shear. When the CGH is plotted, alignment aids, which can help in obtaining the proper hologram size, must be drawn on the hologram plot.

The largest source of error is distortion in the hologram plotter (Wyant et al., 1974). The CGH wavefront accuracy depends on the number of plotter resolution points and the maximum slope of the aspheric wavefront being tested. Assuming that the plotter has $P \times P$ resolution points, there are $P/2$ resolution points across the radius of the hologram. Since the maximum error in plotting any point is half a resolution unit, any portion of each line making up the hologram can be displaced by a distance equal to $1/P$. If the maximum difference between the slope of the test wavefront and the tilted plane wave is $4S$ waves per hologram radius, the phase of the plane wave at the hologram lines can differ from that of the required wavefront at the same lines by as much as $4S/P$ waves (1980). The maximum error in the reconstructed wavefront will be $4S/P$ waves, and since the final interferogram is recorded in the image plane of the hologram, the quantization due to the finite number of resolution points causes a peak error in the final interferogram of $4S/P$ waves. It is important to note that the peak wavefront error of $4S/P$ waves is really a worst-case situation; it occurs only if, in the region of the hologram where the slope difference is maximum, the plotter distortion is also a maximum. This systematic error due to plotter distortion can be calibrated when the plotter distortion is known (Wyant et al., 1974). When the maximum plotting error is equal to one-half of the resolution spot size, the sensitivity of the CGH test ΔW is given by $4S/P$, where P is the number of distortion-free plotter points. Using an e-beam recorder with 0.25- μm resolution over a 10-mm-diameter hologram would enable the measurement of an aspheric wavefront with a maximum wavefront slope of 1000 waves per radius to be tested to a sensitivity of $\lambda/10$ (assuming a perfect plotter).

12.12.3. Off-Axis CGH Aspheric Compensator

To ensure that there is no overlapping of the first and second orders in the Fourier plane (where the spatial filter is located), the tilt angle of the reference plane wave needs to be greater than three times the maximum wavefront slope of the aberrated wave. (Note that there are no even orders for a grating with a 50% duty cycle.) A photograph of the diffracted orders from the hologram in the Fourier plane is shown in Figure 12.48, and a diagram detailing the necessary separations of the orders in the Fourier plane is shown in Figure 12.49. The bandwidth of the N^{th} order is given by $2NS$, where S is the maximum wavefront slope in waves per radius of the wavefront to be reconstructed. This bandwidth (diameter of the diffracted beam in the Fourier plane) determines the size of the spatial-filtering aperture. By moving the spatial-filtering aperture, the output of the interferometer can either be the interference of two plane waves or two aspheric waves.

If the hologram is recorded on a very high-quality optical flat or used in reflection, it can be placed in a single beam of the interferometer, as illustrated in Figure 12.50 using a Fizeau configuration. As we mentioned before, with a pinhole we can isolate

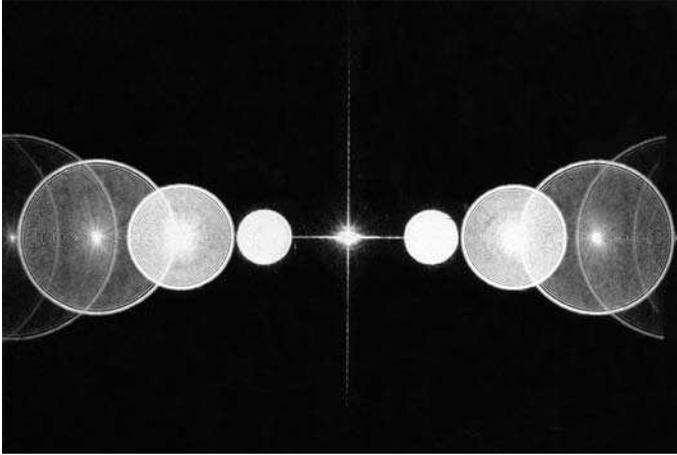


FIGURE 12.48. Diffracted order in Fourier plane of CGH.

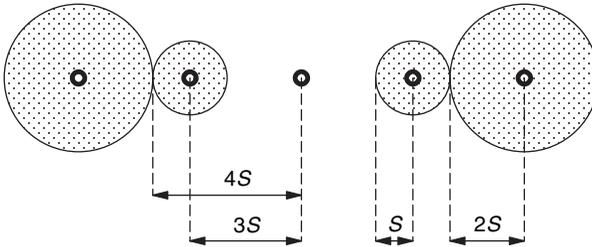


FIGURE 12.49. Diffracted orders in Fourier plane of CGH.

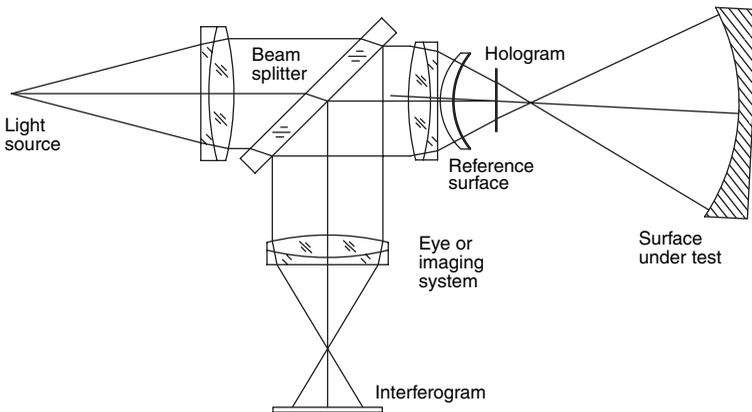


FIGURE 12.50. Fizeau interferometer utilizing a CGH in one beam.

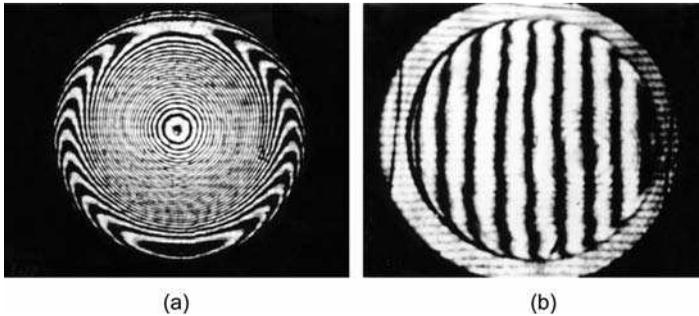


FIGURE 12.51. Results of testing a 10-cm diameter $F/2$ paraboloid (a) without and (b) with a CGH made using an e-beam recorder.

either the diffracted orders $(-1, 0)$ with $(0, -1)$ or the orders $(0,1)$ with $(1,0)$. The whole wavefront compensation is performed in only one of the two passes through the CGH.

Figure 12.51 shows the results of measuring a 10-cm-diameter $F/2$ parabola using a CGH generated with an e-beam recorder (Leung et al., 1980). The fringes obtained in a Twyman–Green interferometer using a helium–neon source without the CGH present are shown in Figure 12.51(a). After the CGH is placed in the interferometer, a much less complicated interferogram is obtained as shown in Figure 12.51(b). The CGH corrects for about 80 fringes of spherical aberration and makes the test much easier to perform.

12.12.4. In-Line CGH Aspheric Compensator

In-line CGHs have the disadvantage with respect to off-axis CGHs that high-order images cannot be completely separated. In this hologram, the separation is made taking advantage that they are focused at different planes, but the isolation of the desired first order cannot be complete. However, they have two great advantages. First, that the alignment is much simpler and second, that since the lateral carrier is not introduced, the degree of asphericity that can be measured is higher. In-line CGH compensators have been described by Mercier et al., (1980), Fercher (1976), and Tiziani et al. (2001). They can be used in a Twyman–Green or Fizeau interferometer, as shown in Figure 12.50.

To avoid asymmetry, so that the paths in and out of the system formed by the CGH and surface under test are the same, the beam to be isolated to produce the interference pattern with the reference beam is the $(1, -1)$.

To test an aspheric convex surface, the arrangement shown in Figure 12.52 can be used. However, to reduce the required size of the auxiliary reference concave surface, the separation between the reference surface and the convex surface can be reduced to a small distance, as suggested by Burge (1995), by engraving the hologram rings on the concave reference surface, as illustrated in Figure 12.53. Notice that the light rays arrive perpendicularly to the concave reference surface, and then, after diffraction, they become perpendicular to the surface under test. Thus, even if the reference

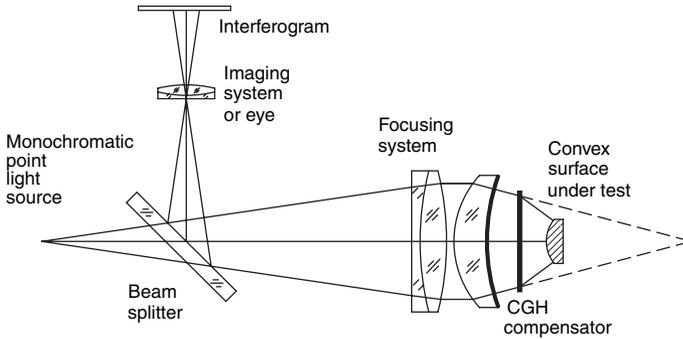


FIGURE 12.52. Testing a convex aspheric surface with an in-line CGH compensator used in a Fizeau interferometer.

surface and the surface under test are placed in contact at the vertex, they would not have exactly the same curvature.

12.12.5. Combination of CGH with Null Optics

Although a CGH can be designed for any optical system, a point is reached where the time and expense required to make a CGH are unreasonable. Also, given enough time and money, null optics, either reflective or refractive, can be designed and built to test almost any arbitrarily complicated optical system.

It is possible to replace the complicated CGH or the complicated null optics required to test complicated optical surfaces (notably aspherical surfaces) with a combination of relatively simple null optics and a relatively simple CGH.

To illustrate the potential of the combined test, results for a CGH-null-lens test of the primary mirror of an eccentric Cassegrain system with a departure of approximately 455 waves (at 514.5 nm) and a maximum slope of approximately 1500 waves per radius are shown in Figure 12.54 (Wyant and O'Neill, 1974). The mirror was a

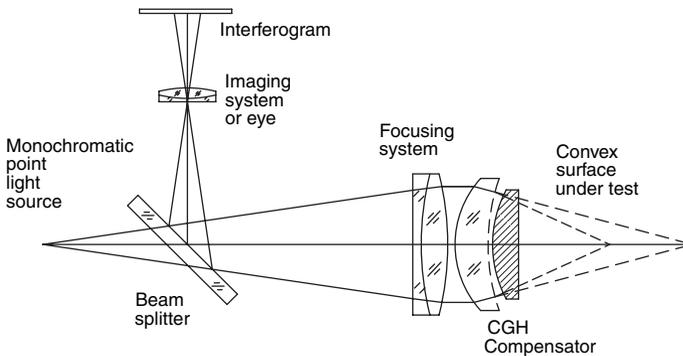


FIGURE 12.53. Testing a convex aspheric surface with a test plate with an in-line CGH compensator on its concave surface.

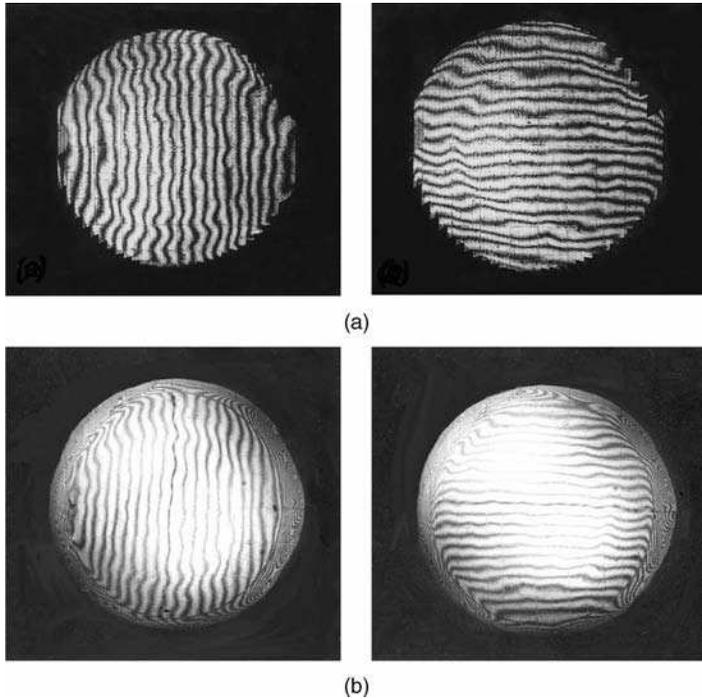


FIGURE 12.54. Setup to test the primary mirror of a Cassegrain telescope using a Maksutov sphere as a partial null with a CGH.

69-cm-diameter off-axis segment whose center lies 81 cm from the axis of symmetry of the parent aspheric surface. The null optics was a Maksutov sphere (as illustrated in Figure 12.55), which reduces the departure and slope of the aspheric wavefront from 910 to 45 waves and from 3000 to 70 waves per radius, respectively. A hologram was then used to remove the remaining asphericity.

Figure 12.55(a) shows interferograms obtained using the CGH-Maksutov test of the mirror under test. Figure 12.55(b) shows the results when the same test was performed using a rather expensive refractive null lens. When allowance is made for the fact that the interferogram obtained with the null lens has much more distortion than the CGH-Maksutov interferogram, and for the difference in sensitivity ($\lambda = 632.8$ nm for the null lens test and 514.5 nm for the CGH-Maksutov test), the results for the two tests are seen to be very similar. The “hills” and “valleys” on the mirror surface appear the same for both tests, as expected. The peak-to-peak surface error measured using the null lens was 0.46 waves ($\lambda = 632.8$ nm), while for the CGH-Maksutov test it was 0.39 waves (514.5 nm). The rms surface error measured was 0.06 waves (632.8 nm) for the null lens, while the CGH-Maksutov test gave 0.07 waves ($\lambda = 514.5$ nm). These results certainly demonstrate that expensive null optics can be replaced by a combination of relatively inexpensive null optics and a CGH.

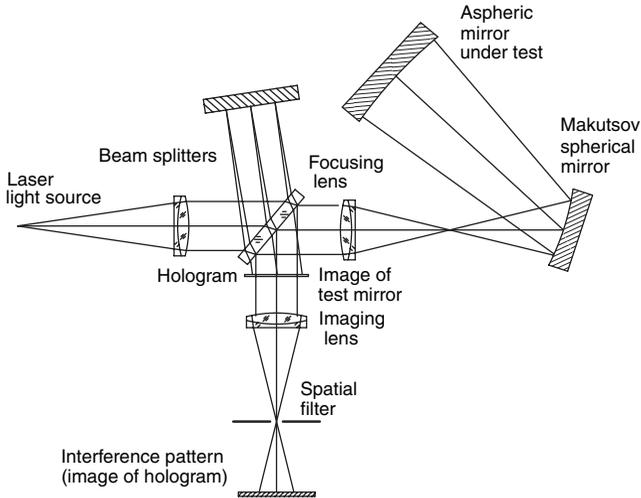


FIGURE 12.55. Results of testing using setup of Figure 12.55 (a) CGH-Maksutov test ($\lambda = 514.5$ nm) and using null lens ($\lambda = 632.8$ nm). (From Wyant and O'Neill, 1974.)

The difficult problem of testing aspheric surfaces, which are becoming increasingly popular in optical design, is made easier by the use of CGHs. The main problem with testing aspheric optical elements is reducing the aberration sufficiently to ensure that light gets back through the interferometer. Combinations of simple null optics with a CGH to perform a test enable the measurement of almost any optical surface. The making and use of a CGH are analogous to using an interferometer setup that yields a large number of interference fringes and measuring the interferogram at a large number of data points. Difficulties involved in recording and analyzing a high-density interferogram and making a CGH are very similar. In both cases, a large number of data points are necessary, and the interferometer must be ray traced so that the aberrations due to the interferometer are well known. The advantage of the CGH technique is that once the CGH is made, it can be used for testing a single piece of optics many times or for testing several identical optical components (Greivenkamp, 1987). In addition, it is much easier for an optician to work with a null setup.

12.13. ASPHERIC TESTING WITH TWO-WAVELENGTH HOLOGRAPHY

The surface under test is often not known close enough to perform a null test. Even if a null test is attempted, the resulting interferogram may contain too many fringes to analyze. Since high accuracy may not be needed, a longer wavelength light source could be used in the interferometer to reduce the number of fringes. Unfortunately, a long wavelength light source creates problems because film and detector arrays may not be available to record the interferogram directly, and the inability to see the

radiation causes considerable experimental difficulty. Two-wavelength and multiple-wavelength techniques provide a means of synthesizing a long effective wavelength using visible light to obtain an interferogram identical to the one that would be obtained if a longer wavelength source were used as shown by several authors, among others by Wyant (1971); Wyant et al. (1984); Cheng and Wyant (1984, 1985); Creath et al. (1985); Creath and Wyant (1986); and Wyant and Creath (1989).

Two-wavelength holography is performed by first photographing the fringe pattern obtained by testing an optical element using a wavelength λ_1 in an interferometer such as the interferometer shown in Figure 12.55. This photographic recording of the fringe pattern (hologram) is then developed, replaced in the interferometer in the exact position it occupied during exposure, and illuminated with the fringe pattern obtained by testing the optical element using a different wavelength λ_2 . The resulting two-wavelength fringes can either be thought of as the moiré between the interference fringes stored in the hologram (recorded at λ_1 and replayed at λ_2) and the live interference fringes (at λ_2), or as the secondary interference between the test wavefront stored in the hologram and the live test wavefront. These fringes are identical to those that would be obtained if the optical element was tested using a long effective wavelength given by (Wyant, 1971)

$$\lambda_e = \frac{\lambda_1 \lambda_2}{|\lambda_1 - \lambda_2|} \quad (12.40)$$

Table 12.6 lists the values of λ_e that can be obtained using various pairs of wavelengths from an argon-ion and a helium–neon laser. By the use of a dye laser, a large range of equivalent wavelengths can be obtained. Tunable helium–neon lasers with four or five distinct wavelengths ranging from green to red are also available.

The contrast in the final interferogram can be increased by spatial filtering. If the filtering is to be effective, the angle between the two interfering beams in the interferometer must be such that only the object beam, not the reference beam, passes through the spatial filter (aperture) shown in Figure 12.55. The spatially filtered pattern yields the interference between the wavefront produced by illuminating (with wavelength λ_2) the hologram recorded at wavelength λ_1 and the wavefront

TABLE 12.6. Possible effective wavelengths in μm obtainable with argon-ion and helium–neon lasers.

λ in nm	459.7	476.5	488.0	496.5	501.7	514.5	632.8
459.7	—	11.73	9.95	5.89	5.24	4.16	1.66
476.5	11.73	—	20.22	11.83	9.49	6.45	1.93
488.0	9.95	20.22	—	28.50	17.87	9.47	2.13
496.5	5.89	11.83	28.50	—	47.90	14.19	2.30
501.7	5.24	9.49	17.87	47.90	—	20.16	2.42
514.5	4.16	6.45	9.47	14.19	20.16	—	2.75
632.8	1.66	1.93	2.13	2.30	2.42	2.75	—

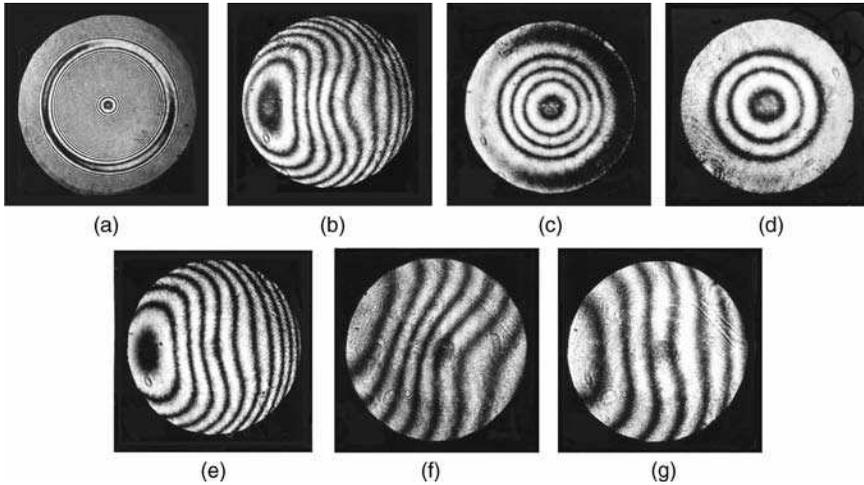


FIGURE 12.56. Interferograms of an optical element at a number of different effective wavelengths: (a) $\lambda = 488$ nm, (b) $\lambda_e = 6.45$ μm , (c) $\lambda_e = 6.45$ μm , (d) $\lambda_e = 9.47$ μm , (e) $\lambda_e = 9.47$ μm , (f) $\lambda_e = 20.22$ μm , (g) $\lambda_e = 28.50$ μm . (From Wyant, 1971.)

obtained from the optical element using wavelength λ_2 . Since the two-wavelength holography interferogram provides the difference between the two interfering beams only in the plane of the hologram, the fringe pattern (hologram) must be recorded in the plane conjugate to the test surface. The final photograph of the interferogram should be recorded in the image plane of the hologram.

Figure 12.56(a) shows an interferogram of an optical element tested using a wavelength of 488.0 nm. The other interferograms were obtained using two-wavelength holography to test the same optical element. The interferograms in Figures 12.56(b) to 12.56(e) were obtained by recording an interferogram (hologram) using a wavelength of 514.5 nm and illuminating the recording with the fringe pattern obtained using a wavelength of 476.5 nm for Figures 12.56(b,c) and of 488.9 nm for Figures 12.56(d,e). The interferograms were spatially filtered, and the amount of tilt shown was adjusted in real time by changing the angle at which the reference wavefront was incident on the hologram during the reconstruction. The interferograms in Figures 12.56(f,g) were obtained by recording an interferogram using a wavelength of 488.0 nm and illuminating this recording with the fringe pattern obtained using a wavelength of 476.5 nm and of 496.5 nm, respectively.

In two-wavelength holography, the final interferogram gives the difference between a fringe pattern recorded at one instant of time and a fringe pattern existing at some later instant of time. If the two fringe patterns are different for reasons other than wavelength change (e.g., air turbulence), incorrect results are obtained. This means that if air turbulence causes a one-fringe change between the fringe pattern obtained using $\lambda_1 = 488.0$ nm and the fringe pattern obtained using $\lambda_2 = 514.5$ nm, the final interferogram will contain one fringe of error corresponding to 9.47 μm .

The effect of air turbulence can be reduced by simultaneously recording the interferograms resulting from the two wavelengths. If the recording process is sufficiently nonlinear and the interferograms have sufficiently high contrast, the interferogram obtained shows the moiré pattern described earlier. Generally, this moiré pattern is too low in contrast to be useful. However, when this interferogram (hologram) is illuminated with a plane wave, is spatially filtered, and is reimaged in the manner shown in Figure 12.55, the result is a high-contrast interferogram, identical to that obtained using the two-wavelength holography method described earlier. Since both fringe patterns are recorded simultaneously, and air dispersion is small, the sensitivity of the interferometer to air turbulence is essentially the same as if a long wavelength light source were used in the interferometer.

More details on the applications of two-wavelength and multiple-wavelength interferometry will be given in Chapter 15.

12.14. WAVEFRONT STITCHING

When the wavefront is strongly aspheric and even with zero tilt in the reference wavefront the minimum fringe spacing is too small, an option to measure the wavefront is segmenting the complete aperture in small regions where the Nyquist condition is not violated, so that the minimum fringe spacing is larger than twice the pixel spacing. There are many approaches to dividing the aperture, but they can be classified in three broad categories.

12.14.1. Annular Zones

One obvious procedure to is to use several different defocussing values, as described by Liu et al. (1988) and by Melozzi et al. (1993). Then, for several rings where the fringe spacing never exceeds, the Nyquist limit are obtained. This is clearly illustrated in Figure 12.57 where three different defocussing values are used. The useful ring where the fringes can be processed has a much larger contrast than the rest.

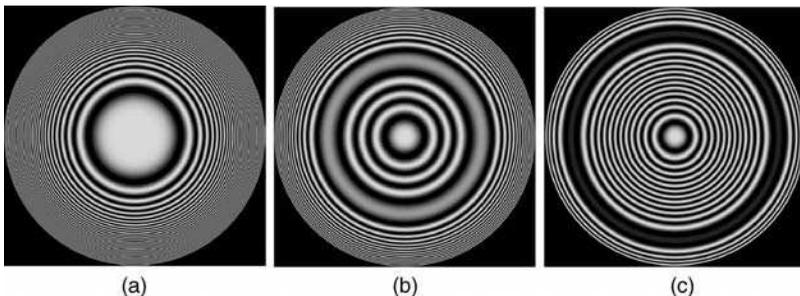


FIGURE 12.57. Interferograms of an aspheric surface with three different defocusing values. The useful ring where the fringes can be processed has a larger contrast than the rest.

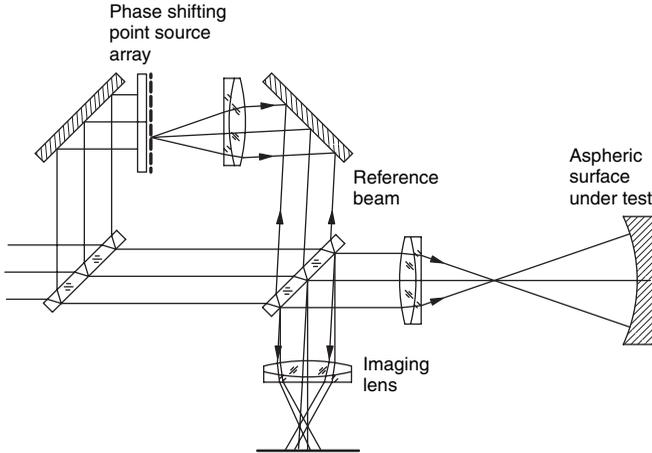


FIGURE 12.58. Schematics of the system to produce a dynamic tilt switching of the reference wavefront in a Mach-Zehnder interferometer.

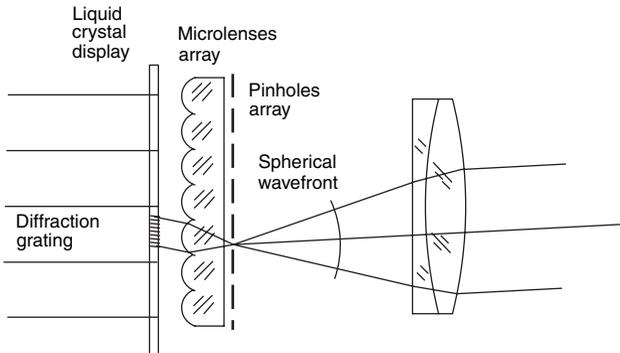


FIGURE 12.59. Detail of the device used to the dynamic tilt switching.

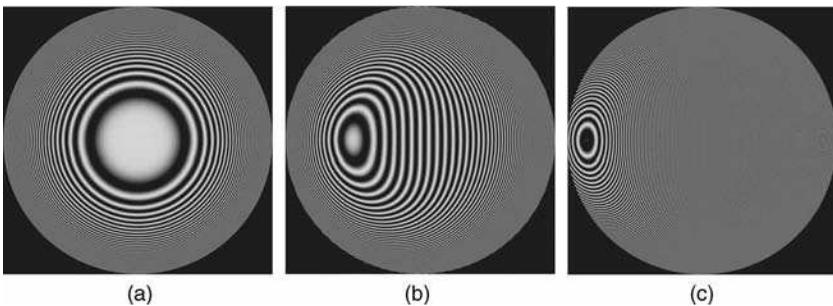


FIGURE 12.60. Interferograms with three different tilts. The zone to be processed is that with the large fringe spacing, not exceeding the Nyquist limit for the detector being used.

12.14.2. Circular Zones

Another method to measure an aspherical surface by stitching is to divide the aperture in many circular zones where the tilt as well as the defocusing is optimized to maximize the minimum fringe spacing. For example, the evaluation of an aspheric surface can be made with many small glass test plates, each one optimized for a different region (Jensen et al., 1984).

This method is also useful when measuring large optical surfaces whose size is much larger than the interferometer aperture (Sjödahl and Oreb, 2002), for example, when testing an extremely large plane (Negro, 1984). Special techniques must be used to insure the continuity of the different apertures, for example with some overlapping and polynomial fitting of the apertures to join them.

12.14.3. Dynamic Tilt Switching

Still another approach, described by Liesener and Tiziani (2004) and Liesener et al. (2004), is to change the tilt in a dynamic manner. Here, again, the useful zone has a much better contrast than the undersampled zone, where the fringes cannot be processed.

To change the tilt of the reference wavefront in a Twyman–Green interferometer in a dynamic manner, the arrangement in Figure 12.58 is used, where the tilted reference beam is generated by electronically selecting different point light sources in a rectangular array. This array, called a phase shifting point source array consists of a high resolution liquid crystal display, followed by a microlens array and an array of pinholes in front of it, at the foci of the microlenses. When a small zone in front of a pinhole is made transparent, all the rest are opaque. However, even if all zones are made transparent, the light from a microlens enter a pinhole only if the light passing thorough a diffraction grating generated on the desired zone diffracts the light so that the first order diffracted light falls on the pinhole. This is illustrated in Figure 12.59. Once a tilt has been selected, the interferograms to be processed are like those in Figure 12.60.

REFERENCES

- Arnold S. M., “How to Test an Asphere with a Computer-Generated Hologram,” *Proc. SPIE*, **1052**, 191–197 (1989).
- Arnold S. M. and R. Kestner, “Verification and Certification of CGH Aspheric Nulls,” *Proc. SPIE*, **2636**, 117–126 (1995).
- Beyerlein M., N. Lindlein, and J. Schwider, “Dual-Wave-Front Computer-Generated Holograms for Quasi-Absolute Testing of Aspherics,” *Appl. Opt.*, **41**, 2440–3106 (2002).
- Biedermann K., “Information Storage Materials for Holography and Optical Data Processing,” *Opt. Acta*, **22**, 103–124 (1975).
- Broder-Bursztyn F. and D. Malacara, “Holographic Interferometer to Test Optical Surfaces,” *Appl. Opt.*, **14**, 2280–2282 (1975).

- Bruns D. G., "Null Test for Hyperbolic Convex Mirrors," *Appl. Opt.*, **22**, 12 (1983).
- Burch C. R., "On Reflection Compensators for Testing Paraboloids," *Mon. Not. R. Astron. Soc.*, **96**, 438 (1936).
- Burch C. R., "Report of the General Meeting of the Association," *J. B. Astron. Assoc.*, **48**, 99 (1938).
- Burge J. H., "Fizeau Interferometry for Large Convex Surfaces," *Proc. SPIE*, **2536**, 127–138 (1995).
- Burge J. H., D. S. Anderson, T. D. Milser, and C. L. Vernold, "Measurement of a Secondary Mirror Using a Holographic Test Plate," *Proc. SPIE*, **2199**, 193–198 (1994).
- Caulfield H. J., P. Mueller, D. Dvornik, A. Epstein, and J. S. Loomis, "Computer Holograms for Optical Testing," *Proc. SPIE*, **306**, 154–157 (1981).
- Chang Y.-C. and J. H. Burge, "Error Analysis for CGH Optical Testing," *Proc. SPIE*, **3782**, 358–366 (1999).
- Cheng Y.-Y. and J. C. Wyant, "Two-Wavelength Phase Shifting Interferometry," *Appl. Opt.*, **23**, 4539–4543 (1984).
- Cheng Y.-Y. and J. C. Wyant, "Multiple-Wavelength Phase Shifting Interferometry," *Appl. Opt.*, **24**, 804–807 (1985).
- Couder A., "Procédé d'Examen d'un Miroir Concave Non-Sphérique," *Rec. Opt. Theor. Instrum.*, **6**, 49 (1927).
- Creath K. and J. C. Wyant, "Direct Phase Measurement of Aspheric Surface Contours," *Proc. SPIE*, **645**, 101–106 (1986).
- Creath K., Y.-Y. Cheng, and J. C. Wyant, "Contouring Aspheric Surfaces Using Two-Wavelength Phase-Shifting Interferometry," *Opt. Acta*, **32**, 1455–1464 (1985).
- Dall H. E., "A Null Test for Paraboloids," *J. Br. Astron. Assoc.*, **57**, 201 (1947).
- Dall H. E., "A Null Test for Paraboloids," in *Amateur Telescope Making*, vol. 3, A. E. Ingalls, Ed., Scientific American, New York, 1953, pp. 149–153.
- Dörband and H. J. Tiziani, "Testing Aspheric Surfaces with Computer Generated-Holograms: Analysis of Adjustment and Shape Errors," *Appl. Opt.*, **24**, 2604–2611 (1985).
- Dyson J., "Unit Magnification Optical System without Seidel Aberrations," *J. Opt. Soc. Am.*, **49**, 713 (1959).
- Emmel P. M., K. M. Leung, "A New Instrument for Routine Optical Testing of General Aspherics," *Proc. SPIE*, **171**, 93–99 (1979).
- Everhart E., "Null Test for Wright Telescope Mirrors," *Appl. Opt.*, **5**, 717–718 (1966).
- Friesem A. A., Y. Katzir, Z. Rav-Noy, and B. Sharon, "Photoconductor-Thermoplastic Devices for Holographic Nondestructive Testing," *Opt. Eng.*, **19**, 659–665 (1980).
- Geary J. M., "Fiber/Cylinder Interferometer Test: Focus and Off-Axis Response," *Opt. Eng.*, **30**, 1902–1909 (1991).
- Geary J. M. and L. J. Parker, "New Test for Cylindrical Optics," *Opt. Eng.*, **26**, 813–820 (1987).
- Greivenkamp J. E., "Sub-Nyquist Interferometry," *Appl. Opt.*, **26**, 5245–5258 (1987).
- Hansler R. L., "A Holographic Foucault Knife-Edge Test for Optical Elements of Arbitrary Design," *Appl. Opt.*, **7**, 1863–1864 (1968).
- Hariharan P., "Quasi-Heterodyne Hologram Interferometry," *Opt. Eng.*, **24**, 632–638 (1985).
- Hariharan P., B. F. Oreb, and N. Brown, "A Digital Phase-Measurement System for Real-Time Holographic Interferometry," *Opt. Commun.*, **41**, 393–396 (1982).

- Heintze L. R., H. D. Polster, and J. Vrabel, "A Multiple-Beam Interferometer for Use with Spherical Wavefronts," *Appl. Opt.*, **6**, 1924 (1967).
- Hildebrand B. P., K. A. Haines, and R. Larkin, "Holography as a Tool in the Testing of Large Apertures," *Appl. Opt.*, **6**, 1267–1269 (1967a).
- Hildebrand B. P. and K. A. Haines, "Multiple-Wavelength and Multiple-Source Holography Applied to Contour Generation," *J. Opt. Soc. Am.*, **57**, 155–162 (1967b).
- Hindle J. H., "A New Test for Cassegrainian and Gregorian Secondary Mirrors," *Mon. Not. R. Astron. Soc.*, **91**, 592 (1931).
- Holleran R. T., "Immersion Null Test for Aspherics," *Appl. Opt.*, **2**, 1336 (1963).
- Holleran R. T., "Null Testing Telescope Mirrors by Immersion," *Sky Telesc.*, **28**, 242 (1964).
- Holleran R. T., "Third-Order Wavefronts and Related Null Tests," *Appl. Opt.*, **5**, 1244 (1966).
- Holleran R. T., "An Algebraic Solution for the Small Lens Null Compensator," *Appl. Opt.*, **7**, 137 (1968).
- Houston J. B., Jr., C. J. Buccini, and P. K. O'Neill, "A Laser Unequal Path Interferometer for the Optical Shop," *Appl. Opt.*, **6**, 1237 (1967).
- Howard J. W., M. H. Beaulieu, and R. J. Zielinski, "A Nonobscuring, Easily-Callibrated Method for Testing Ellipsoids on a Fizeau Interferometer," *Proc. SPIE*, **389**, 26 (1983).
- James W. E. and M. D. Waterworth, "A Method for Testing Aspheric Surfaces," *Opt. Acta*, **12**, 223–227 (1965).
- Jensen S. C., W. W. Chow, and G. N. Lawrence, "Subaperture Testing Approaches: A Comparison," *Appl. Opt.*, **23**, 740–744 (1984).
- Kirkham A. R., "The Ronchi Test for Mirrors," in: *Amateur Telescope Making*, vol. 1, A. G. Ingalls, Ed., Scientific American, New York, 1953, p. 264.
- Lam P., J. D. Gaskill, and J. C. Wyant, "Two Wavelength Holographic Interferometer," *Appl. Opt.*, **23**, 3079–3081 (1984).
- Lamrecht J., N. Lindlein, and J. Schwider, "Null Test Measurement of High-Numerical-Aperture Cylindrical Microlenses in Transmitted Light," *Proc. SPIE*, **5180**, 253 (2003).
- Landgrave J. E. A. and J. R. Moya, "Dummy Lens for the Computer Optimization of Autostigmatic Null Correctors," *Appl. Opt.*, **26**, 2673 (1987).
- Lee W. H., "Sampled Fourier Transform Hologram Generated by Computer," *Appl. Opt.*, **9**, 639–643 (1970).
- Lee W. H., "Binary Synthetic Holograms," *Appl. Opt.*, **13**, 1677–1682 (1974).
- Leung K. M., T. C. Lee, E. Bernal G., and J. C. Wyant, "Two-Wavelength Contouring with the Automated Thermoplastic Holographic Camera," *Proc. SPIE*, **192**, 184–189 (1979).
- Leung K. M., J. C. Lindquist, and L. T. Shepherd, "E-Beam Computer Generated Holograms for Aspheric Testing," *Proc. SPIE*, **215**, 70–75 (1980).
- Liesener J. and H. Tiziani, "Interferometer with Dynamic Reference," *Proc. SPIE*, **5252**, 264–271 (2004).
- Liesener J., L. Seifert, H. J. Tiziani, and W. Osten, "Active Wavefront Sensing and Wavefront Control with SLMs," *Proc. SPIE*, **5532**, 147–158 (2004).
- Liu Y.-M., G. N. Lawrence, and C. L. Koliopoulos, "Subaperture Testing of Aspheres with Annular Zones" *Appl. Opt.*, **27**, 4504–4513 (1988).
- Loomis J. S., "Computer-Generated Holography and Optical Testing," *Opt. Eng.*, **19**, 679–685 (1980).

- Lurionov N. P., A. V. Lukin, and K. S. Mustafin, "Holographic Inspection of Shapes of Unpolished Surfaces," *Sov. J. Opt. Technol.*, **39**, 154–155 (1972).
- MacGovern A. J. and J. C. Wyant, "Computer Generated Holograms for Testing Optical Elements," *Appl. Opt.*, **10**, 619–624 (1971).
- Malacara Z. and D. Malacara, "Design of Lenses to Project the Image of a Pupil in Optical Testing Interferometers," *Appl. Opt.*, **34**, 739–742 (1995).
- Malacara D. and C. Menchaca, "Imaging of the Wavefront Under Test in Interferometry," *Proc. SPIE*, **540**, 34 (1985).
- Meinel A. B. and M. P. Meinel, "Self-Null Corrector Test for Telescope Hyperbolic Secondaries," *Appl. Opt.*, **22**, 520 (1983a).
- Meinel A. B. and M. P. Meinel, "Self-Null Corrector Test for Telescope Hyperbolic Secondaries: Comments," *Appl. Opt.*, **22**, 2405 (1983b).
- Melozzi M., L. Pezzati, and A. Mazzoni, "Testing Aspheric Surfaces Using Multiple Annular Interferograms," *Opt. Eng.*, **32**, 1073–1079 (1993).
- Mercier R., F. Bridou, B. Bonino, and M. Mullet, "Control of Aspherics by In-Line Computer-Generated Holograms," *Proc. SPIE*, **235**, 80–84 (1980).
- Moya J. R. and J. E. A. Landgrave, "Third-Order Design of Refractive Offner Compensators," *Appl. Opt.*, **26**, 2667 (1987).
- Negro J. E., "Subaperture Optical System Testing," *Appl. Opt.*, **23**, 1921–1930 (1984).
- Norman B. A., "New Test for Cassegrainian Secondaries," *Sky Telesc.*, **17**, 38 (1957).
- Offner A., "A Null Corrector for Paraboloidal Mirrors," *Appl. Opt.*, **2**, 153 (1963).
- Offner A., "Field Lenses and Secondary Axial Aberration," *Appl. Opt.*, **8**, 1735 (1969).
- Ono A. and J. C. Wyant, "Plotting Errors Measurement of CGH Using an Improved Interferometric Method," *Appl. Opt.*, **23**, 3905–3910 (1984).
- Parks R. E., "Making and Testing an f/0.15 Parabola," *Appl. Opt.*, **13**, 1987–1987 (1974).
- Parks R. E. and L. Z. Shao, "Testing Large Hyperbolic Secondary Mirrors," *Opt. Eng.*, **27**, 1057 (1988).
- Pastor J., "Hologram Interferometry and Optical Technology," *Appl. Opt.*, **8**, 525–531 (1969).
- Pruss C., S. Reichelt, H. J. Tiziani, and W. Osten, "Computer-Generated Holograms in Interferometric Testing," *Opt. Eng.*, **43**, 2534–2540 (2004).
- Puryayev D. T., "A Quality Control Technique for Convex Elliptical, Parabolic and Hyperbolic Surfaces of Simple Lenses," *Sov. J. Opt. Technol.*, **38**, 684 (1971).
- Puryayev D. T., "Compensator for Inspecting the Quality of Large-Diameter Parabolic Mirrors," *Sov. J. Opt. Technol.*, **40**, 238 (1973).
- Reichelt S., C. Pruss, and H. J. Tiziani, "Absolute Testing of Aspheric Surfaces," *Proc. SPIE*, **5252**, 252–263 (2004).
- Ritchey G. W., "On the Modern Reflecting Telescope and the Making and Testing of Optical Mirrors," *Smithson. Contrib. Knowl.*, **34**, 3 (1904).
- Robbert C. F., "Typical Error Budget for Testing High-Performance Aspheric Telescope Mirror," *Proc. SPIE*, **181**, 56 (1979).
- Rodgers J. M. and R. E. Parks, "Null Tests for Oblate Spheroids," *Appl. Opt.*, **23**, 1246–1247 (1984).
- Rodgers J. M., "A Null-Lens Design Approach for Centrally Obscured Components," *Proc. SPIE*, **679**, 17 (1986).

- Ross F. E., "Parabolizing Mirrors without a Flat," *Astrophys. J.*, **98**, 341 (1943).
- Sasian J. M., "Design of Null Correctors for the Testing of Astronomical Optics," *Opt. Eng.*, **27**, 1051 (1988).
- Sasian J. M., "Optimum Configuration of the Offner Null Corrector: Testing an F/1 Paraboloid," *Proc. SPIE*, **1164**, 8 (1989).
- Schlauch J., "Construction of a Dall Null Tester," *Sky Telesc.*, **18**, 222 (1959).
- Schupmann L., *Die Medial-Fernrohre: Eine neue Konstruktion für grosse astronomische Instrumente*, Teubner, Leipzig, 1899.
- Schwider J. and R. Burow, "The Testing of Aspherics by Means of Rotational-Symmetric Synthetic Holograms," *Opt. Appl.*, **6**, 83 (1976).
- Schwider J., "Interferometric Tests for Aspherics," *OSA TOPS*, **24**, 103–114 (1999).
- Shafer D. R., "Zoom Null Lens," *Appl. Opt.*, **18**, 3863 (1979).
- Shnurr A. D. and A. Mann, "Optical Figure Characterization for Cylindrical Mirrors and Lenses," *Opt. Eng.*, **20**, 412–416 (1981).
- Silvertooth W., "A Modification of the Hindle Test for Cassegrain Secondaries," *J. Opt. Soc Am.*, **30**, 140 (1940).
- Simpson F. A., B. H. Oland, and J. Meckel, "Testing Convex Aspheric Lens Surfaces with a Modified Hindle Arrangement," *Opt. Eng.*, **13**, 101 (1974).
- Sjödahl M. and B. F. Oreb, "Stitching Interferometric Measurement Data for Inspection of Large Optical Components," *Opt. Eng.*, **41**, 403–408 (2002).
- Snow K. and R. Vandewarker, "On Using Holograms for Test Glasses," *Appl. Opt.*, **9**, 822–827 (1970).
- Stoltzmann D. E. and M. Hatch, "Extensions of the Dall Null Test," *Sky Telesc.*, **52**, 210 (1976).
- Tiziani H. J., S. Reichlet, C. Pruss, M. Rocktächel, and U. Hofbauer, "Testing of Aspheric Surfaces," *Proc. SPIE*, **4440**, 109–119 (2001).
- Uhrich C. and L. Hesselink, "Optical Surface Inspection Using Real-Time Fourier Transform Holography in Photorefractives," *Appl. Opt.*, **27**, 4497–4503 (1988).
- Wyant J. C., "Testing Aspherics Using Two-Wavelength Holography," *Appl. Opt.*, **10**, 2113–2118 (1971).
- Wyant J. C. and K. Creath, "Two-Wavelength Phase-Shifting Interferometer and Method," U.S. Patent No. 4,832,489 (1989).
- Wyant J. C. and V. P. Bennett, "Using Computer Generated Holograms to Test Aspheric Wavefronts," *Appl. Opt.*, **11**, 2833–2839 (1972).
- Wyant J. C. and P. K. O'Neill, "Computer Generated Hologram: Null Lens Test of Aspheric Wavefronts," *Appl. Opt.*, **13**, 2762–2765 (1974).
- Wyant J. C., B. F. Oreb, and P. Hariharan, "Testing Aspherics Using Two-Wavelength Holography: Use of Digital Electronic Techniques," *Appl. Opt.*, **23**, 4020–4023 (1984).
- Wyant J. C., P. K. O'Neill, and A. J. Mac Govern, "Interferometric method of measuring plotter distortion," *Appl. Opt.*, **11**, 1549–1551 (1974).

13

Zernike Polynomials and Wavefront Fitting

Virendra N. Mahajan

13.1. INTRODUCTION

Optical imaging systems generally have an axis of rotational symmetry, and their pupil is circular or annular, as in the case of the systems with mirrors. The wave aberration function of such systems can be expanded in a power series or a complete set of orthogonal polynomials. Although introduced by Zernike (1934a,b) for testing the figure of a circular mirror by a knife edge test or his phase contrast method, Zernike circle polynomials were studied extensively by Nijboer (1943, 1947) and Nienhuis and Nijboer (1949) to describe the diffraction theory of aberrations. We show that these polynomials, which are orthogonal over a circular pupil, represent balanced aberrations that yield minimum variance. For small aberrations, a minimum of aberration variance yields a maximum of Strehl ratio (Mahajan, 1982, 1983, 1993). Relationships between the coefficients of a power-series expansion and those of a Zernike-polynomial expansion are given. In particular, the conversion of Seidel coefficients into Zernike coefficients or vice versa is discussed. Similarly, Zernike annular polynomials representing balanced aberrations with minimum variance for systems with annular pupils are discussed (Mahajan, 1981a,b; 1984; 1994). Balanced aberrations that yield minimum ray aberration variance, often called the root mean square spot radius, are different and are not discussed here (Braat, 1987). Finally, the determination of Zernike coefficients from a discrete set of wavefront error data obtained by ray tracing, as in optical design, or by measurement, as in optical testing, is discussed, regardless of the shape of the pupil. The procedure described is thus applicable, for example, to elliptical pupils obtained in off-axis imaging, odd-shaped vignetted pupils, and hexagonal pupils of a segmented mirror with hexagonal segments.

13.2. ABERRATIONS OF A ROTATIONALLY SYMMETRIC SYSTEM WITH A CIRCULAR PUPIL

13.2.1. Power Series Expansion

Consider a rotationally symmetric optical system imaging a point object P as illustrated in Figure 13.1. The axis of the rotational symmetry, namely, the optical axis, lies along the z axis. Let the position vector of the object point be \vec{h} with Cartesian coordinates (x_o, y_o) in a plane orthogonal to the optical axis. Similarly, let \vec{r} be the position vector of a point Q with Cartesian coordinates (x, y) in the plane of the exit pupil of the system, which is also orthogonal to the optical axis. The origins of (x_o, y_o) and (x, y) lie on the optical axis and we assume, for example, that the x_o and x axes are coplanar. The polar coordinates (h, θ_o) and (r, θ) of the object and pupil points are related to their Cartesian coordinates according to

$$(x_o, y_o) = h(\cos \theta_o, \sin \theta_o), \quad (13.1a)$$

and

$$(x, y) = (r \cos \theta, \sin \theta). \quad (13.1b)$$

It should be evident from Figure 13.2 that we are using the standard convention for a polar angle as the angle of a position vector with the x axis.

It is quite common in the optics literature to consider a point object lying along the y axis. When using polar coordinates of a point in the plane of the exit pupil, the polar angle θ in that case is defined as the angle made by the position vector of the point with the y axis, contrary to the standard convention as the angle with the x axis. We choose a point object along the x axis so that, for example, coma aberration is expressed as $x(x^2 + y^2)$ and not as $y(x^2 + y^2)$. A positive value of our coma aberration yields a diffraction point spread function that is symmetric *about* the x axis (or symmetric *in* y) with its peak and centroid shifted to positive values of x with respect to the Gaussian image point.

The aberration function of a rotationally symmetric system depends on the four coordinates (x_o, y_o) and (x, y) through three rotational invariants h^2 , r^2 , and $hr \cos(\theta - \theta_o)$. Without loss of generality, we assume that the point object lies along the x_o axis, so that $\theta_o = 0$. A power-series expansion of the aberration function $W(h'; r, \theta)$ in that case may be written (Mahajan, 2001) as

$$\begin{aligned} W(h'; r, \theta) &= \sum_{l=0}^{\infty} \sum_{p=0}^{\infty} \sum_{m=0}^{\infty} C_{lpm} (h^2)^l (r^2)^p (hr \cos \theta)^m \\ &= \sum_{l=0}^{\infty} \sum_{p=0}^{\infty} \sum_{m=0}^{\infty} C_{lpm} h^{2l+m} r^{2p+m} \cos^m \theta \\ &= \sum_{l=0}^{\infty} \sum_{n=0}^{\infty} \sum_{m=0}^n 2_{l+m} a_{nm} h^{2l+m} r^n \cos^m \theta \end{aligned} \quad (13.2)$$

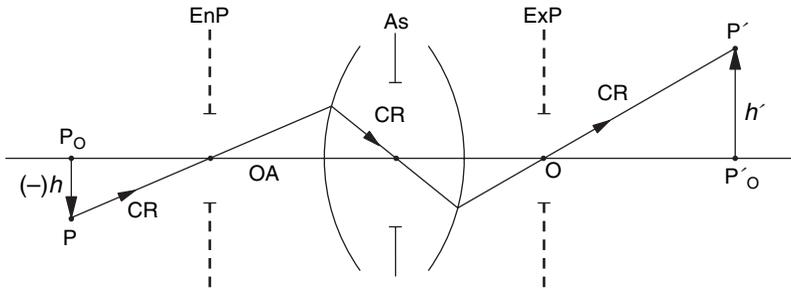


FIGURE 13.1. Schematic diagram of an optical imaging system. A point object P at a height h from the optical axis OA is imaged at P' at a height h' . AS is the aperture stop, EnP is the entrance pupil, ExP is the exit pupil, and CR is the chief ray.

where C_{lpm} and a_{nm} are the expansion coefficients; and l , p , and m are positive integers, including zero, and $n = 2p + m$. Note that $n - m \geq 0$ and even. A term with $n = 0 = m$, called piston aberration, is also included although such a term does not constitute an aberration (since it corresponds to the chief ray, which by definition has a zero aberration associated with it). The aberration function is defined with respect to a Gaussian reference sphere of radius of curvature R that passes through the center of the exit pupil and whose center of curvature lies at the Gaussian image point P' at a height h' from the optical axis. The magnification of the image is $M = h'/h$. The wave aberration $W(x, y)$ of a ray passing through a point $Q(x, y)$ in the plane of the exit pupil represents the difference of its optical path length from that of the chief ray, which passes through the centers of the entrance and exit pupils. It is positive or negative depending on whether it travels a longer or a shorter optical path length compared to the chief ray in reaching the reference sphere.

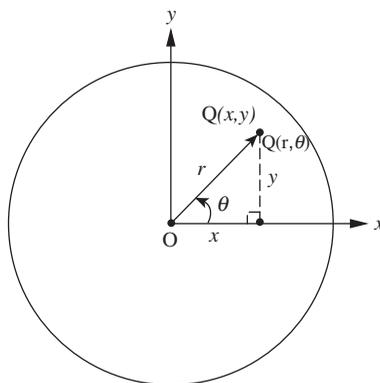


FIGURE 13.2. Cartesian and polar coordinates (x, y) and (r, θ) , respectively, of a pupil point Q in the plane of the exit pupil.

The order i of an aberration term, which is equal to its degree in the object and pupil coordinates, is given by $i = 2l + m + n$. The order i is always even, and the number of terms N_i of this order is given by

$$N_i = (i + 2)(i + 4)/8 \quad (13.3)$$

The aberration terms are referred to as the classical aberrations. Since the ray aberration of a ray, representing the coordinates of the point of its intersection with the image plane with respect to the Gaussian image point, is related to its wave aberration by a spatial derivative according to (Welford, 1974; Malacara and Malacara, 1994; Mouroulis and Macdonald, 1997; Mahajan, 2001)

$$(x_i, y_i) = \frac{R}{n_i} \left(\frac{\partial W}{\partial x}, \frac{\partial W}{\partial y} \right) \quad (13.4)$$

where $W(x, y)$ is the wave aberration in the Cartesian coordinates and n_i is the refractive index of the image space (which is generally unity in practice), its order is one less than that of the wave aberration. For example, a fourth-order wave aberration yields a third-order ray aberration.

If the aberration terms having different dependencies on the object coordinates but the same dependence on pupil coordinates are combined so that there is only one term for each pair of (n, m) values, Eq. (13.2) for the power-series expansion of the aberration function may be written as

$$W(\rho, \theta) = \sum_{n=0}^{\infty} \sum_{m=0}^n a_{nm} \rho^n \cos^m \theta \quad (13.5a)$$

where $\rho = r/a$ is a radial variable normalized by the radius a of the exit pupil, and an expansion coefficient

$$a_{nm} = a^n \sum_{l=0}^{\infty} 2^{l+m} a_{nm} h^{2l+m} \quad (13.5b)$$

represents the peak value of an aberration term. The number of terms through a certain order i in the reduced power-series expansion of the aberration function is also given by Eq. (13.3). The terms of Eq. (13.2) through a certain order i correspond to those terms of Eq. (13.4) for which $n + m \leq i$. It should be evident that when a system is tested interferometrically for its aberrations, the explicit dependence of an aberration term on h' is lost.

13.2.2. Primary or Seidel Aberration Function

The primary or Seidel aberration function corresponds to $i = 4$ and consists of five terms given by

$$W_P(h'; r, \theta) = {}_0a_{40}r^4 + {}_1a_{31}h'r^3 \cos \theta + {}_2a_{22}h'^2r^2 \cos^2 \theta + {}_2a_{20}h'^2r^2 + {}_3a_{11}h'^3r \cos \theta \quad (13.6)$$

TABLE 13.1 Primary aberrations in a simplified form; $i = 4$, $n + m \leq 4$.

n	m	Aberration term $a_{nm}\rho^n \cos^m \theta$	Aberration name
1	1	$a_{11}\rho \cos \theta$	Distortion
2	0	$a_{20}\rho^2$	Field curvature
2	2	$a_{22}\rho^2 \cos^2 \theta$	Astigmatism
3	1	$a_{31}\rho^3 \cos \theta$	Coma
4	0	$a_{40}\rho^4$	Spherical

where ${}_0a_{40}$, ${}_1a_{31}$, ${}_2a_{22}$, ${}_2a_{20}$, and ${}_3a_{11}$ represent the coefficients of spherical aberration, coma, astigmatism, field curvature, and distortion. The primary aberration function ($n + m \leq 4$) in the reduced form is given by

$$W_P(\rho, \theta) = a_{40}\rho^4 + a_{31}\rho^3 \cos \theta + a_{22}\rho^2 \cos^2 \theta + a_{20}\rho^2 + a_{11}\rho \cos \theta \quad (13.7)$$

The primary aberrations are listed in Table 13.1. The distortion term has the same dependence on the pupil coordinates (ρ, θ) as a wavefront tilt aberration, but its coefficient a_{11} varies with h' as h'^3 . Similarly, the field curvature term has the same dependence on the pupil coordinates as a wavefront defocus aberration, but its coefficient a_{20} varies with h' as h'^2 .

To simplify the notation further, we write the primary aberration function in the form

$$W(r, \theta; h') = a_s r^4 + a_c h' r^3 \cos \theta + a_a h'^2 r^2 \cos^2 \theta + a_d h'^2 r^2 + a_t h'^3 r \cos \theta \quad (13.8)$$

or

$$W(\rho, \theta) = A_s \rho^4 + A_c \rho^3 \cos \theta + A_a \rho^2 \cos^2 \theta + A_d \rho^2 + A_t \rho \cos \theta \quad (13.9)$$

It should be evident that an aberration coefficient A_i represents the peak value of the corresponding aberration term. For spherical aberration A_s , the ray geometrical spot radius in the Gaussian image plane is equal $8FA_s$, and the corresponding longitudinal spherical aberration is $16F^2A_s$, where F is the focal ratio of the image-forming light cone (Mahajan, 2001). For coma A_c , the tangential and sagittal coma, representing the length and half width of the coma spot pattern, are given by $6FA_c$ and $2FA_c$, respectively. For astigmatism A_a , the sagittal and tangential line images have a length of $8FA_a$ separated by a distance $8F^2A_a$, called the longitudinal astigmatism. The effect of distortion is to displace the image along the axis by $2FA_t$.

The field curvature coefficient generally consists of two terms according to

$$A_d = \frac{1}{2}A_a + A_p \quad (13.10)$$

where A_p is the coefficient of Petzval field curvature. Whereas the coefficient A_a of astigmatism depends on the object location, A_p is independent of it. The astigmatism

and field curvature terms can be combined as follows:

$$\begin{aligned} W_{a+d}(\rho, \theta) &= A_a \rho^2 \cos^2 \theta + A_d \rho^2 \\ &= \frac{1}{2} A_a (3x_n^2 + y_n^2) + A_p (x_n^2 + y_n^2) \\ &= A_{ta} x_n^2 + A_{sa} y_n^2 \end{aligned} \quad (13.11)$$

where $(x_n, y_n) = (x, y)/a$ are the normalized Cartesian coordinates, and

$$A_{ta} = \left(\frac{3}{2} A_a + A_p \right) = A_a + A_d \quad (13.12a)$$

and

$$A_{sa} = \left(\frac{1}{2} A_a + A_p \right) = A_d \quad (13.12b)$$

are called the tangential and sagittal astigmatism wave aberrations, respectively. The corresponding ray aberrations are given by $4FA_{ta}$ and $4FA_{sa}$. The sagittal, tangential, and Petzval images lie at distances $8F^2 A_d$, $8F^2 (A_a + A_d)$, and $8F^2 A_p$, respectively, from the Gaussian image point. The radii of curvature of the corresponding image surfaces are given by $R_s = h^2/16F^2 A_d$, $R_t = h^2/16F^2 (A_a + A_d)$, and $R_p = h^2/16F^2 A_p$. They are related to each other according to (Malacara and Malacara, 1994; Mahajan, 2001)

$$\frac{3}{R_s} - \frac{1}{R_t} = \frac{2}{R_p} \quad (13.13)$$

It has the consequence that the Petzval surface is three times as far from the tangential surface as it is from the sagittal surface. Moreover, the sagittal surface always lies between the tangential and Petzval surfaces. It should be evident that a field curvature of A_p implies a spot radius of $4FA_p$ in the Gaussian image plane or a point image on the Petzval surface.

The aberration function can also be written in terms of Seidel sums for a point object with maximum image height h'_{max} according to (Welford, 1974; Malacara and Malacara, 1994; Mouroulis and Macdonald, 1997; Mahajan, 2001)

$$\begin{aligned} W(\rho, \theta; h') &= \frac{1}{8} S_I \rho^4 + \frac{1}{2} S_{II} \frac{h'}{h'_{max}} \rho^3 \cos \theta + \frac{1}{2} S_{III} \left(\frac{h'}{h'_{max}} \right)^2 \rho^2 \cos^2 \theta \\ &\quad + \frac{1}{4} (S_{III} + S_{IV}) \left(\frac{h'}{h'_{max}} \right)^2 \rho^2 + \frac{1}{2} S_V \left(\frac{h'}{h'_{max}} \right)^3 \rho \cos \theta \end{aligned} \quad (13.14)$$

where S_I , S_{II} , S_{III} , S_{IV} , and S_V are called the *Seidel sums* for spherical aberration, coma, astigmatism, Petzval field curvature (also called simply the Petzval sum), and

distortion, respectively. For $h' = h'_{max}$, the multiplication of a Seidel sum by F gives the spot radius in the Gaussian image plane in the case of S_I , sagittal coma in the case of S_{II} , radius of the circle of least confusion in the case of S_{III} (lying half way between the sagittal and tangential line images), spot radius in the case of $S_{III} + S_{IV}$, and the image displacement in the case of S_V .

The various aberration coefficients are related to each other according to

$$\begin{aligned} A_s \equiv a_{40} &= a_s a^4 = \frac{1}{8} S_I, \quad A_c \equiv a_{31} = a_c h' a^3 = \frac{1}{2} \frac{h'}{h'_{max}} S_{II} \\ A_a \equiv a_{22} &\equiv a_d h'^2 a^2 = \frac{1}{2} \left(\frac{h'}{h'_{max}} \right)^2 S_{III} \\ A_d \equiv a_{20} &= a_d h'^2 a^2 = \frac{1}{4} \left(\frac{h'}{h'_{max}} \right)^2 (S_{III} + S_{IV}), \quad A_t \equiv a_{11} = a_t h'^3 a = \frac{1}{2} \left(\frac{h'}{h'_{max}} \right)^3 S_V \end{aligned} \quad (13.15)$$

New aberrations arise when the elements of a rotationally symmetric system are misaligned with respect to each other. The misalignment of an element may be the decentering of its vertex and/or tilting of its optical axis (Mahajan, 2001). A misalignment introduces axial coma, which depends on the pupil coordinates as does the primary coma. However, unlike primary coma, it does not depend on the image height h' , that is, it is constant across the entire image of an extended object. A misalignment also introduces astigmatism and field curvature which vary linearly with h' , and distortion varies as h'^2 . Thus, the degree or the power with which each aberration term introduced by a misalignment varies with h' is one less than that for the corresponding terms for an aligned system. Accordingly, the degree of each new aberration term in the image (or object) and pupil coordinates is three, that is, one less than the nominal four for a primary aberration. When one or more surfaces of a system is warped so that it is curved more in one plane than in an orthogonal plane, as a toric surface in the case of an astigmatic human eye, it yields (axial) astigmatism independent of h' , that is, even for an axial point object.

13.2.3. Secondary or Schwarzschild Aberration Function

The secondary or Schwarzschild aberration function corresponding to $i = 6$ consists of nine terms given by

$$\begin{aligned} W_S(h'; r, \theta) &= {}_0a_{60}r^6 + {}_1a_{51}h'r^5 \cos \theta + {}_2a_{42}h'^2r^4 \cos^2 \theta + {}_2a_{40}h'^2r^4 + {}_3a_{33}h'^3r^3 \cos^3 \theta \\ &\quad + {}_3a_{31}h'^3r^3 \cos \theta + {}_4a_{22}h'^4r^2 \cos^2 \theta + {}_4a_{20}h'^4r^2 + {}_5a_{11}h'^5r \cos \theta \end{aligned} \quad (13.16)$$

where the first four terms with $l = 0$ are referred to as the secondary aberrations [spherical, coma, astigmatism (also called wings), and arrows], and the last five are called the lateral aberrations [spherical, coma, astigmatism, field curvature, and distortion]. The lateral aberrations are similar to the primary aberrations except for their dependence on the image height h' . The lateral spherical aberration ${}_2a_{40}h'^2r^4$ is

also called the oblique spherical aberration. The aberration function through the sixth order ($n + m \leq 6$) according to Eq. (13.4) consists of nine terms given by

$$W(\rho, \theta) = a_{11}\rho \cos \theta + a_{20}\rho^2 + a_{22}\rho^2 \cos^2 \theta + a_{31}\rho^3 \cos \theta + a_{33}\rho^3 \cos^3 \theta + a_{40}\rho^4 + a_{42}\rho^4 \cos^2 \theta + a_{51}\rho^5 \cos \theta + a_{60}\rho^6 \quad (13.17)$$

Since the dependence of a term on the image height h' is contained in the coefficient a_{nm} , the coefficients in Eqs. (13.7) and (13.17) that seem to be the same are actually different from each other. For example, a_{11} is equal to ${}_3a_{11}h'^3a$ in Eq. (13.7) but equal to $({}_3a_{11}h'^3 + {}_5a_{11}h'^5)a$ in Eq. (13.17). Similarly, the coefficient a_{40} now consists of both the Seidel and lateral spherical aberrations.

13.2.4. Zernike Circle Polynomial Expansion

The aberration function of a system with a circular exit pupil can be expanded in terms of a complete set of Zernike circle polynomials $R_n^m(\rho) \cos m\theta$ that are orthogonal over a unit circle in the form (Born and Wolf, 1999; Mahajan, 2004)

$$W(\rho, \theta) = \sum_{n=0}^{\infty} \sum_{m=0}^n c_{nm} Z_n^m(\rho, \theta) \quad (13.18)$$

where c_{nm} are the expansion coefficients that depend on the location of the point object, n and m are positive integers including zero, $n - m \geq 0$ and even, and

$$Z_n^m(\rho, \theta) = [2(n+1)/(1+\delta_{m0})]^{1/2} R_n^m(\rho) \cos m\theta \quad (13.19a)$$

is an orthonormal Zernike polynomial. Here, δ_{ij} is a Kronecker delta, and

$$R_n^m(\rho) = \sum_{s=0}^{(n-m)/2} \frac{(-1)^s (n-s)!}{s! (\frac{n+m}{2} - s)! (\frac{n-m}{2} - s)!} \rho^{n-2s} \quad (13.19b)$$

is a polynomial of degree n in ρ containing terms in $\rho^n, \rho^{n-2}, \dots$, and ρ^m . The radial circle polynomials $R_n^m(\rho)$ are even or odd in ρ , depending on whether n (or m) is even or odd. Also, $R_n^n(1) = 1$, $R_n^n(\rho) = \rho^n$, and $R_n^m(0) = \delta_{m0}$ for even $n/2$ and $-\delta_{m0}$ for odd $n/2$. The Zernike polynomials are orthogonal according to

$$\frac{1}{\pi} \int_0^1 \int_0^{2\pi} Z_n^m(\rho, \theta) Z_{n'}^{m'}(\rho, \theta) \rho d\rho d\theta = \delta_{nm} \delta_{n'm'} \quad (13.20a)$$

$$\int_0^{2\pi} \cos m\theta \cos m'\theta d\theta = \pi(1 + \delta_{m0}) \delta_{mm'} \quad (13.20b)$$

and

$$\int_0^1 R_n^m(\rho)R_{n'}^m(\rho)\rho d\rho = \frac{1}{2(n+1)}\delta_{nn'} \tag{13.20c}$$

The Zernike expansion coefficients are given by

$$c_{nm} = \frac{1}{\pi} \int_0^1 \int_0^{2\pi} W(\rho, \theta)Z_n^m(\rho, \theta)\rho d\rho d\theta \tag{13.21}$$

as may be seen by substituting Eq. (13.18) and utilizing the orthogonality relations.

The Zernike circle polynomials are unique in that they are the only polynomials in two variables ρ and θ , which (a) are orthogonal over a unit circle, (b) are invariant in

TABLE 13.2. Orthonormal Zernike circle polynomials and balanced aberrations.

n	m	Orthonormal Zernike polynomial		Aberration Name ^a
		$Z_n^m(\rho, \theta) = \left[\frac{2(n+1)}{1+\delta_{m0}} \right]^{1/2}$	$R_n^m(\rho) \cos m\theta$	
0	0	1		Piston
1	1	$2\rho \cos \theta$		Distortion (tilt)
2	0	$\sqrt{3}(2\rho^2 - 1)$		Field curvature (defocus)
2	2	$\sqrt{6}\rho^2 \cos 2\theta$		Primary astigmatism
3	1	$\sqrt{8}(3\rho^3 - 2\rho) \cos \theta$		Primary coma
3	3	$\sqrt{8}\rho^3 \cos 3\theta$		
4	0	$\sqrt{5}(6\rho^4 - 6\rho^2 + 1)$		Primary spherical
4	2	$\sqrt{10}(4\rho^4 - 3\rho^2) \cos 2\theta$		Secondary astigmatism
4	4	$\sqrt{10}\rho^4 \cos 4\theta$		
5	1	$\sqrt{12}(10\rho^5 - 12\rho^3 + 3\rho) \cos \theta$		Secondary coma
5	3	$\sqrt{12}(5\rho^5 - 4\rho^3) \cos 3\theta$		
5	5	$\sqrt{12}\rho^5 \cos 5\theta$		
6	0	$\sqrt{7}(20\rho^6 - 30\rho^4 + 12\rho^2 - 1)$		Secondary spherical
6	2	$\sqrt{14}(15\rho^6 - 20\rho^4 + 6\rho^2) \cos 2\theta$		Tertiary astigmatism
6	4	$\sqrt{14}(6\rho^6 - 5\rho^4) \cos 4\theta$		
6	6	$\sqrt{14}\rho^6 \cos 6\theta$		
7	1	$4(35\rho^7 - 60\rho^5 + 30\rho^3 - 4\rho) \cos \theta$		Tertiary coma
7	3	$4(21\rho^7 - 30\rho^5 + 10\rho^3) \cos 3\theta$		
7	5	$4(7\rho^7 - 6\rho^5) \cos 5\theta$		
7	7	$4\rho^7 \cos 7\theta$		
8	0	$3(70\rho^8 - 140\rho^6 + 90\rho^4 - 20\rho^2 + 1)$		Tertiary spherical

^aThe words “orthonormal Zernike” are to be associated with these names, for example, *orthonormal Zernike primary astigmatism*.

form with respect to rotation of the coordinate axes about the origin, and (c) include a polynomial for each permissible pair of n and m values (Bhatia and Wolf, 1954; Born and Wolf, 1999).

The orthonormal Zernike polynomials and the names associated with some of them when identified with aberrations are listed in Table 13.2 for $n \leq 8$. The number of Zernike (or orthogonal) aberration terms in the expansion of an aberration function through a certain order n is given by

$$N_n = \left(\frac{n}{2} + 1\right)^2 \quad \text{for even } n \quad (13.22a)$$

$$= (n + 1)(n + 3)/4 \quad \text{for odd } n \quad (13.22b)$$

Consider a typical Zernike aberration term in Eq. (13.18):

$$W_n^m(\rho, \theta) = c_{nm} Z_n^m(\rho, \theta) \quad (13.23)$$

Unless $n = m = 0$ its mean value is zero; i.e.,

$$\begin{aligned} \langle W_n^m(\rho, \theta) \rangle &= \frac{1}{\pi} \int_0^1 \int_0^{2\pi} W_n^m(\rho, \theta) \rho d\rho d\theta \\ &= 0, \quad n \neq 0, m \neq 0 \end{aligned} \quad (13.24)$$

For $m = 0$, this may be seen with the help of Eq. (13.20b) and the fact that $R_0^0(\rho) = 1$ is a member of the polynomial set. The orthogonality Eq. (13.20c) yields the result that the mean value of $R_n^0(\rho)$ is zero. The presence of the constant term in $R_n^0(\rho)$ yields a zero mean value. When $m \neq 0$, the average value of $\cos m\theta$ is zero. Similarly, the mean square value of the aberration is given by

$$\langle [W_n^m(\rho, \theta)]^2 \rangle = \frac{1}{\pi} \int_0^1 \int_0^{2\pi} [W_n^m(\rho, \theta)]^2 \rho d\rho d\theta = c_{nm}^2 \quad (13.25)$$

Hence, its variance is given by

$$\begin{aligned} \sigma_{nm}^2 &= \langle (W_n^m)^2 \rangle - \langle W_n^m \rangle^2 \\ &= c_{nm}^2, \quad n \neq 0, m \neq 0 \end{aligned} \quad (13.26)$$

Thus, each expansion coefficient, with the exception of c_{00} , represents the standard deviation of the corresponding aberration term. The variance of the aberration function is accordingly given by

$$\sigma_W^2 = \langle W^2(\rho, \theta) \rangle - \langle W(\rho, \theta) \rangle^2 = \sum_{n=1}^{\infty} \sum_{m=0}^n c_{nm}^2 \quad (13.27)$$

Unless the mean value of the aberration $\langle W \rangle = c_{00} = 0$, $\sigma_W \neq W_{\text{RMS}}$, where $W_{\text{RMS}} = \langle W^2 \rangle^{1/2}$ is the root-mean-square (rms) value of the aberration.

The utility of Zernike polynomials comes from the fact that they are orthogonal over a circular pupil and represent balanced aberrations yielding minimum variance (Born and Wolf, 1999; Mahajan, 2004). The advantage of using orthonormal

polynomials is that the standard deviation of an aberration term $c_{nm}Z_n^m(\rho, \theta)$ is simply equal to the Zernike coefficient c_{nm} . Moreover, because of the orthogonality, the value of a coefficient c_{nm} given by Eq. (13.21) is independent of the number of terms used to expand an aberration function. Thus, the coefficients are not affected if one or more orthogonal terms are added or subtracted from the aberration function. Accordingly, the addition of tilt or defocus aberration to the aberration function of a system, as is done often in optical testing, does not change the other aberration coefficients.

Figure 13.3 shows how a radial circle polynomial $R_n^m(\rho)$ varies with ρ . As stated earlier, its value at the edge of the pupil $\rho = 1$ is unity, that is, $R_n^m(1) = 1$. For defocus and spherical aberrations, that is, for $m = 0$, its value at the origin $\rho = 0$ is 1 or -1 depending on whether $n/2$ is even or odd. Its value at the origin in the case of tilt, coma, and astigmatism aberrations is zero.

13.2.5. Zernike Circle Polynomials as Balanced Aberrations for Minimum Wave Aberration Variance

Zernike polynomials represent balanced aberrations such that an aberration of a certain order in pupil coordinates in the power series expansion is balanced with aberrations of lower or equal order to minimize its variance (Born and Wolf, 1999; Mahajan, 2001; Mahajan, 2004). For example, consider spherical aberration $A_s\rho^4$ combined with defocus aberration $B_d\rho^2$:

$$W(\rho) = A_s\rho^4 + B_d\rho^2 \quad (13.28)$$

The defocus aberration is obtained by observing the image in a defocused image plane. If R is the radius of curvature of the Gaussian reference sphere with respect to which the aberration is defined, then an observation in a defocused image plane at a distance z implies a defocus wave aberration of $B_d\rho^2$, where the defocus coefficient is given by $B_d = (z^{-1} - R^{-1})a^2/2$. The variance of the combined aberration is given by

$$\sigma_W^2 = \frac{4A_s^2}{45} + \frac{B_d^2}{12} + \frac{A_sB_d}{6} \quad (13.29)$$

Letting $\partial\sigma_W^2/\partial B_d = 0$, we find that the aberration variance is minimum when $B_d = -A_s$, that is, equal and opposite amount of defocus. The standard deviation σ_W is reduced by a factor of 4 from $2A_s/3\sqrt{5}$ to $A_s/6\sqrt{5}$. The balanced aberration $A_s(\rho^4 - \rho^2)$ is similar to the Zernike polynomial $Z_4^0(\rho)$. The constant (independent of ρ and θ) or the piston term represented by 1 (and multiplied by $\sqrt{5}$) in the expression for $Z_4^0(\rho)$ does not change the standard deviation of the balanced aberration. However, it makes the mean value of the aberration zero. Similarly, coma can be combined with tilt

$$W(\rho, \theta) = A_c\rho^3 \cos \theta + B_t\rho \cos \theta \quad (13.30)$$

where B_t is the coefficient of tilt wave aberration. The variance is minimized when $B_t = -2A_c/3$ and the standard deviation is reduced by a factor of 3 from $A_c/2\sqrt{2}$ to

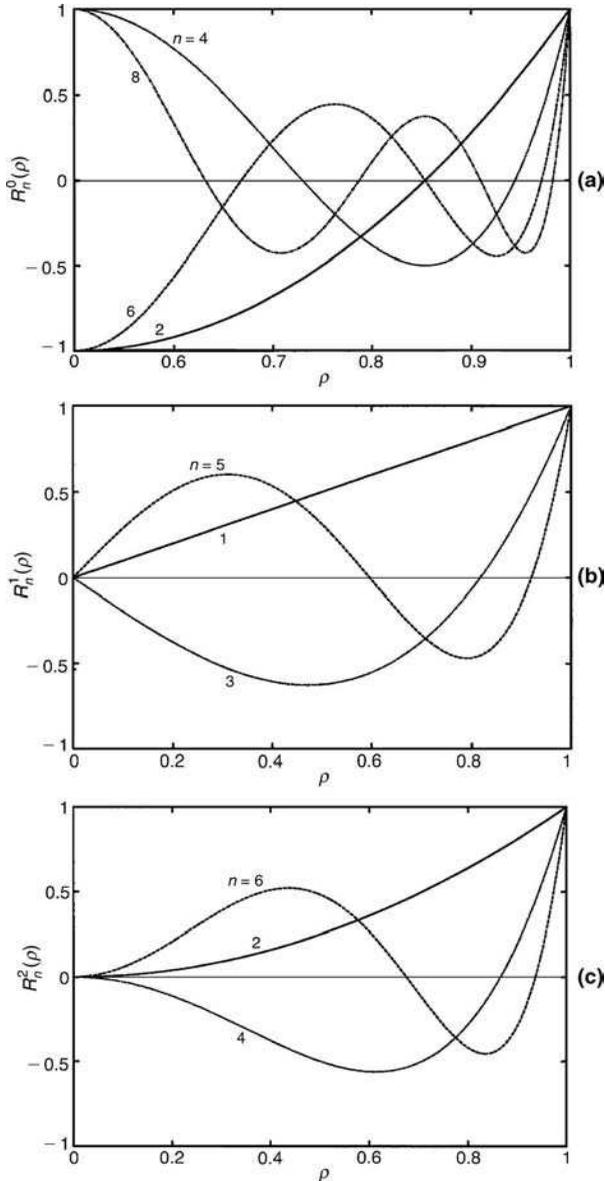


FIGURE 13.3. Variation of a radial circle polynomial $R_n^m(\rho)$ with ρ . (a) Defocus and spherical aberrations. (b) Tilt and coma. (c) Astigmatism.

$A_c/6\sqrt{2}$. It implies that the aberration variance is minimum when the aberration is measured with respect to the point $(4FA_c/3, 0)$. The balanced aberration $A_c[\rho^3 - (2/3)\rho]\cos\theta$ is similar to the Zernike polynomial $Z_3^1(\rho, \theta)$. When we balance astigmatism $A_a\rho^2\cos^2\theta$ with defocus, we find that its standard deviation is reduced by a factor of 1.225 from $A_a/4$ to $A_a/2\sqrt{6}$ when combined with a defocus

aberration of $B_d = -A_a/2$. The balanced aberration $(1/2)A_a\rho^2 \cos 2\theta$ is similar to the Zernike polynomial $Z_2^2(\rho, \theta)$. Moreover, the field curvature (or defocus) and distortion (or wavefront tilt) aberrations can be identified with $Z_2^0(\rho)$ and $Z_1^1(\rho, \theta)$, respectively. Again, there is a piston term in $Z_2^0(\rho)$, which does not affect the variance. The piston polynomial $Z_0^0(\rho, \theta)$ is simply equal to unity.

The higher-order aberrations can be balanced in a similar manner. For example, when secondary spherical aberration varying as ρ^6 is combined with primary spherical aberration and defocus to minimize its variance, the balanced aberration obtained has the same form as the Zernike polynomial $Z_6^0(\rho)$ (except for a constant term). Similarly, when secondary coma varying as $\rho^5 \cos \theta$ is combined with primary coma and tilt to minimize its variance, the balanced aberration obtained is the same as the Zernike polynomial $Z_5^1(\rho, \theta)$. In the case of secondary astigmatism $\rho^4 \cos^2 \theta$, the balancing aberrations are primary spherical aberration, primary astigmatism, and defocus, as may be seen from the identity

$$\rho^4 \cos^2 \theta - \frac{1}{2}\rho^4 - \frac{3}{4}\rho^2 \cos^2 \theta + \frac{3}{8}\rho^2 = \frac{1}{2}\left(\rho^4 - \frac{3}{4}\rho^2\right) \cos 2\theta \quad (13.31)$$

The balanced aberration has the form of the Zernike polynomial $Z_4^2(\rho, \theta)$. Thus, Zernike polynomials can be identified with balanced aberrations; that, in fact, is their advantage over another complete set that is orthogonal over a circular pupil. Moreover, the fundamental aberrations of wavefront tilt and defocus are members of the polynomial set; without their inclusion, the expansion would be incomplete.

13.2.6. Relationships Between Coefficients of Power-Series and Zernike-Polynomial Expansions

A classical aberration term of Eq. (13.5a) has the form $\rho^n \cos^m \theta$ and an orthogonal aberration of Eq. (13.18) has the form $R_n^m(\rho) \cos m\theta$. Thus, the radial variation ρ^n is replaced by $R_n^m(\rho)$ and the angular variation $\cos^m \theta$ is replaced by $\cos m\theta$. To relate the power-series coefficients a_{nm} of Eq. (13.5a) to the Zernike coefficients c_{nm} of Eq. (13.18), we first change the indices n and m of coefficients a_{nm} to k and l [not to be confused with the l used earlier in Eq. (13.2)], respectively, and write the power-series expansion of the aberration function in the form

$$W(\rho, \theta) = \sum_{k=0}^{\infty} \sum_{l=0}^k a_{kl} \rho^k \cos^l \theta \quad (13.32)$$

where k and l are positive integers including zero, and $k - l \geq 0$ and even. The coefficients a_{kl} and c_{nm} can be related to each other by comparing Eqs. (13.18) and (13.32) and using the identities

$$\cos m\theta = (1 + d_{m0})2^{m-1} \cos^m \theta + m \sum_{q=1}^{m-1} \frac{(-1)^q (m-q-1)! 2^{m-2q-1}}{q(q-1)!(m-2q)!} (\cos \theta)^{m-2q} \quad (13.33)$$

and

$$\cos^l \theta = \frac{1}{2l} \sum_{q=0}^l \frac{l!}{q!(l-q)!} \cos(l-2q)\theta \quad (13.34)$$

If the Zernike coefficients c_{nm} are known, the power-series coefficients a_{kl} can be obtained from them as follows (Tyson, 1982; Mahajan, 2001). Substituting Eqs. (13.19b) and (13.33) into Eq. (13.18) and equating the coefficient of the term containing the factor $\rho^k \cos^l \theta$ in the equation thus obtained to the corresponding coefficient in Eq. (13.32), we find that

$$a_{00} = c_{00} \quad (13.35)$$

and

$$a_{kl} = \sum_{n=0}^{\infty} \sum_{m=0}^n b_{klnm} c_{nm} \quad (13.36)$$

where

$$b_{klnm} = \begin{cases} [2(n+1)(1+\delta_{m0})]^{1/2} \frac{(-1)^{\frac{n-k}{2}} \left(\frac{n+k}{2}\right)! 2^{m-1}}{\left(\frac{n-k}{2}\right)! \left(\frac{k+m}{2}\right)! \left(\frac{k-m}{2}\right)!}, & m = l \\ \left[\frac{2(n+1)}{1+\delta_{m0}} \right]^{1/2} \frac{(-1)^{\frac{n-k}{2}} \left(\frac{n+k}{2}\right)! m \ 2^l (-1)^{\frac{m-l}{2}} \left(\frac{m+l}{2}-1\right)!}{\left(\frac{n-k}{2}\right)! \left(\frac{k+m}{2}\right)! \left(\frac{k-m}{2}\right)! l!(m-l) \left(\frac{m-l}{2}-1\right)!}, & m \neq l \end{cases} \quad (13.37a)$$

The values of b_{klnm} for $k \leq 8$ and $n \leq 8$ are given in Table 13.3. Note that only those coefficients exist, for which $k-l$, $n-m$, $n-k$, $k-m$, and $m-l$ are all positive even integers including zero.

If the power-series coefficients a_{kl} are known, the Zernike coefficients c_{nm} can be obtained from them as follows (Conforti, 1983; Mahajan, 2001). Substituting Eq. (13.34) into Eq. (13.32) and then substituting the equation thus obtained and Eq. (13.19b) into Eq. (13.21), we obtain

$$c_{nm} = \sum_{k=0}^{\infty} \sum_{l=0}^{\infty} d_{nmkl} a_{kl} \quad (13.38)$$

where

$$d_{nmkl} = \left[\frac{2(n+1)}{1+\delta_{m0}} \right]^{1/2} \frac{l!}{2^l \left(\frac{l-m}{2}\right)! \left(\frac{l+m}{2}\right)!} \sum_{s=0}^{(n-m)/2} \frac{(-1)^s (n-s)!}{s! \left(\frac{n+m}{2}-s\right)! \left(\frac{n-m}{2}-s\right)! (n-2s+k+2)} \quad (13.39)$$

TABLE 13.3. Values of b_{klm} used for obtaining power-series aberration coefficients a_{kl} from Zernike aberration coefficients c_{nm} .

n	m	k	0	1	2	2	3	3	4	4	4	4	5	5	5	6	6	6	6	7	7	7	7	8
		l	0	1	0	2	1	3	0	2	4	1	3	5	0	2	4	6	1	3	5	7	0	
0	0		1																					
1	1			2																				
2	0		$-\sqrt{3}$		$2\sqrt{3}$																			
2	2				$-\sqrt{6}$	$2\sqrt{6}$																		
3	1			$-4\sqrt{2}$				$6\sqrt{2}$																
3	3					$-6\sqrt{2}$	$8\sqrt{2}$																	
4	0		$\sqrt{5}$		$-6\sqrt{5}$			$6\sqrt{5}$																
4	2				$3\sqrt{10}$	$-6\sqrt{10}$		$-4\sqrt{10}$	$8\sqrt{10}$															
4	4							$\sqrt{10}$	$-8\sqrt{10}$	$8\sqrt{10}$														
5	1			$6\sqrt{3}$		$-24\sqrt{3}$				$20\sqrt{3}$														
5	3					$24\sqrt{3}$	$-32\sqrt{3}$			$-30\sqrt{3}$	$40\sqrt{3}$													
5	5									$10\sqrt{3}$	$-40\sqrt{3}$	$32\sqrt{3}$												
6	0		$-\sqrt{7}$		$6\sqrt{7}$			$-30\sqrt{7}$							$20\sqrt{7}$									
6	2				$-6\sqrt{14}$	$12\sqrt{14}$		$20\sqrt{14}$	$-40\sqrt{14}$						$-15\sqrt{14}$	$30\sqrt{14}$								
6	4							$-5\sqrt{14}$	$40\sqrt{14}$	$-40\sqrt{14}$					$6\sqrt{14}$	$-48\sqrt{14}$	$48\sqrt{14}$							
6	6														$-\sqrt{14}$	$18\sqrt{14}$	$-48\sqrt{14}$	$32\sqrt{14}$						
7	1			-16			120			-240										140				
7	3						-120	160		360	-480									-252	336			
7	5									-120	480	-384								140	-560	448		
7	7																			-28	224	-448	256	
8	0		3		-60			270							-420									210

The values of d_{nmkl} for $n \leq 8$ and $k \leq 8$ are given in Table 13.4. Note that only those coefficients exist for which $n - m$, $k - l$, and $l - m$ are all positive even integers including zero.

13.2.7. Conversion of Seidel Aberrations into Zernike Aberrations

From Table 13.4 or by inspection, each Seidel aberration term in Eq. (13.7) can be written in terms of the orthonormal Zernike polynomials $Z_n^m(\rho, \theta)$ as follows:

$$a_{11}\rho \cos \theta = \frac{a_{11}}{2} Z_1^1(\rho, \theta) \quad (13.40a)$$

$$a_{20}\rho^2 = a_{20} \left(\frac{Z_2^0}{2\sqrt{3}} + \frac{Z_0^0}{2} \right) \quad (13.40b)$$

$$a_{22}\rho^2 \cos^2 \theta = a_{22} \left(\frac{Z_2^2}{2\sqrt{6}} + \frac{Z_2^0}{4\sqrt{3}} + \frac{Z_0^0}{4} \right) \quad (13.40c)$$

$$a_{31}\rho^3 \cos \theta = a_{31} \left(\frac{Z_3^1}{6\sqrt{2}} + \frac{Z_1^1}{3} \right) \quad (13.40d)$$

and

$$a_{40}\rho^4 = a_{40} \left(\frac{Z_4^0}{6\sqrt{5}} + \frac{Z_2^0}{2\sqrt{3}} + \frac{Z_0^0}{3} \right) \quad (13.40e)$$

Accordingly, the Seidel aberration function of Eq. (13.7) can be written in terms of Zernike polynomials

$$W_P(\rho, \theta) = c_{00}Z_0^0 + c_{11}Z_1^1 + c_{20}Z_2^0 + c_{22}Z_2^2 + c_{31}Z_3^1 + c_{40}Z_4^0 \quad (13.41)$$

where

$$c_{00} = \frac{a_{20}}{2} + \frac{a_{22}}{4} + \frac{a_{40}}{3}, \quad c_{11} = \frac{a_{11}}{2} + \frac{a_{31}}{3}, \quad c_{20} = \frac{a_{20}}{2\sqrt{3}} + \frac{a_{22}}{4\sqrt{3}} + \frac{a_{40}}{2\sqrt{3}} \quad (13.42a)$$

$$c_{22} = \frac{a_{22}}{2\sqrt{6}}, \quad c_{31} = \frac{a_{31}}{6\sqrt{2}}, \quad \text{and} \quad c_{40} = \frac{a_{40}}{6\sqrt{5}} \quad (13.42b)$$

Its mean and mean square values are given by

$$\langle W_P \rangle = c_{00} = \frac{a_{20}}{2} + \frac{a_{22}}{4} + \frac{a_{40}}{3} \quad (13.43a)$$

and

$$\langle W_P^2 \rangle = \sum_{n=0}^4 \sum_{m=0}^2 c_{nm}^2 \quad (13.43b)$$

$$= \langle W_P \rangle^2 + \left(\frac{a_{11}}{2} + \frac{a_{31}}{3} \right)^2 + \left(\frac{a_{20}}{2\sqrt{3}} + \frac{a_{22}}{4\sqrt{3}} + \frac{a_{40}}{2\sqrt{3}} \right)^2 + \frac{a_{22}^2}{24} + \frac{a_{31}^2}{72} + \frac{a_{40}^2}{180} \quad (13.43c)$$

TABLE 13.4. Values of d_{nmkl} for obtaining Zernike aberration coefficients c_{nm} from the power-series aberration coefficients a_{kl} .

n	m	$k \backslash l$	0	1	2	2	3	3	4	4	4	5	5	5	6	6	6	7	7	7	7	8
			0	1	0	2	1	3	0	2	4	1	3	5	0	2	4	6	1	3	5	7
0	0		1		$\frac{1}{2}$	$\frac{1}{4}$		$\frac{1}{3}$	$\frac{1}{6}$	$\frac{1}{8}$		$\frac{1}{4}$		$\frac{1}{8}$	$\frac{3}{32}$	$\frac{5}{64}$					$\frac{1}{5}$	
1	1			$\frac{1}{2}$			$\frac{1}{3}$	$\frac{1}{4}$				$\frac{1}{4}$	$\frac{3}{16}$	$\frac{5}{32}$			$\frac{1}{5}$	$\frac{3}{20}$	$\frac{1}{8}$	$\frac{7}{64}$		
2	0				$\frac{1}{2\sqrt{3}}$	$\frac{1}{4\sqrt{3}}$			$\frac{1}{2\sqrt{3}}$	$\frac{1}{4\sqrt{3}}$	$\frac{\sqrt{3}}{16}$			$\frac{3\sqrt{3}}{20}$	$\frac{3\sqrt{3}}{40}$	$\frac{9\sqrt{3}}{160}$	$\frac{3\sqrt{3}}{64}$				$\frac{2}{5\sqrt{3}}$	
2	2					$\frac{1}{2\sqrt{6}}$			$\frac{1}{8}\sqrt{\frac{3}{2}}$	$\frac{1}{8}\sqrt{\frac{3}{2}}$				$\frac{1}{10}\sqrt{\frac{3}{2}}$	$\frac{1}{10}\sqrt{\frac{3}{2}}$	$\frac{3}{32}\sqrt{\frac{3}{2}}$						
3	1						$\frac{1}{6\sqrt{2}}$	$\frac{1}{8\sqrt{2}}$				$\frac{1}{5\sqrt{2}}$	$\frac{3}{20\sqrt{2}}$	$\frac{1}{8\sqrt{2}}$				$\frac{1}{5\sqrt{2}}$	$\frac{3}{20\sqrt{2}}$	$\frac{1}{8\sqrt{2}}$	$\frac{7}{64\sqrt{2}}$	
3	3							$\frac{1}{8\sqrt{2}}$					$\frac{1}{10\sqrt{2}}$	$\frac{1}{8\sqrt{2}}$					$\frac{1}{12\sqrt{2}}$	$\frac{5}{48\sqrt{2}}$	$\frac{7}{64\sqrt{2}}$	
4	0							$\frac{1}{6\sqrt{5}}$	$\frac{1}{12\sqrt{5}}$	$\frac{1}{16\sqrt{5}}$				$\frac{1}{4\sqrt{5}}$	$\frac{1}{8\sqrt{5}}$	$\frac{3}{32\sqrt{5}}$	$\frac{\sqrt{5}}{64}$				$\frac{2}{7\sqrt{5}}$	
4	2								$\frac{1}{8\sqrt{10}}$	$\frac{1}{8\sqrt{10}}$					$\frac{1}{6\sqrt{10}}$	$\frac{1}{6\sqrt{10}}$	$\frac{1}{32}\sqrt{\frac{5}{2}}$					
4	4									$\frac{1}{8\sqrt{5}}$					$\frac{\sqrt{5}}{48}$	$\frac{\sqrt{5}}{32}$						
5	1										$\frac{1}{20\sqrt{3}}$	$\frac{\sqrt{3}}{80}$	$\frac{1}{32\sqrt{3}}$					$\frac{\sqrt{5}}{35}$	$\frac{3\sqrt{3}}{140}$	$\frac{\sqrt{3}}{56}$	$\frac{\sqrt{3}}{64}$	
5	3											$\frac{1}{40\sqrt{3}}$	$\frac{1}{32\sqrt{3}}$						$\frac{5}{140\sqrt{3}}$	$\frac{5}{112\sqrt{3}}$	$\frac{\sqrt{3}}{64}$	
5	5												$\frac{1}{32\sqrt{3}}$							$\frac{\sqrt{3}}{112}$	$\frac{\sqrt{3}}{64}$	
6	0													$\frac{1}{20\sqrt{7}}$	$\frac{1}{40\sqrt{7}}$	$\frac{1}{160\sqrt{7}}$	$\frac{1}{64\sqrt{7}}$					$\frac{1}{10\sqrt{7}}$
6	2														$\frac{1}{30\sqrt{14}}$	$\frac{1}{30\sqrt{14}}$	$\frac{1}{32\sqrt{14}}$					
6	4															$\frac{1}{40\sqrt{14}}$	$\frac{1}{32\sqrt{14}}$					
6	6																$\frac{1}{32\sqrt{14}}$					
7	1																	$\frac{1}{140}$	$\frac{3}{560}$	$\frac{1}{224}$	$\frac{1}{256}$	
7	3																		$\frac{1}{336}$	$\frac{5}{1344}$	$\frac{1}{256}$	
7	5																			$\frac{1}{448}$	$\frac{1}{256}$	
7	7																				$\frac{1}{256}$	
8	0																					$\frac{1}{210}$

Hence, its variance is given by

$$\begin{aligned}\sigma_W^2 &= \langle W_P^2 \rangle - \langle W_P \rangle^2 = \sum_{n=1}^4 \sum_{m=0}^2 c_{nm}^2 \\ &= \frac{a_{11}^2}{4} + \frac{a_{20}^2}{12} + \frac{a_{22}^2}{16} + \frac{a_{31}^2}{8} + \frac{4a_{40}^2}{45} + \frac{a_{11}a_{31}}{3} + \frac{a_{20}a_{40}}{6} + \frac{a_{22}a_{40}}{12}\end{aligned}\quad (13.44)$$

13.2.8. Conversion of Zernike Aberrations into Seidel Aberrations

Just as a Seidel aberration is made up of more than one Zernike aberrations [see Eqs. (13.40)], similarly, a given Zernike aberration is made up of more than one Seidel aberration. For example, Zernike astigmatism $c_{22}Z_2^2(\rho, \theta)$ consists of Seidel astigmatism $\rho^2 \cos^2 \theta$ and defocus ρ^2 . Similarly, Zernike coma $c_{31}Z_3^1(\rho, \theta)$ consists of Seidel coma $\rho^3 \cos \theta$ and tilt $\rho \cos \theta$, and Zernike spherical aberration $c_{40}Z_4^0(\rho)$ consists of Seidel aberration ρ^4 , defocus, and piston. Of course, higher-order Zernike aberrations also consist of Seidel aberrations. For example, secondary Zernike spherical aberration $c_{60}Z_6^0(\rho)$ consists of secondary spherical aberration ρ^6 , Seidel spherical aberration ρ^4 , defocus, and piston. Since a Seidel term is contained in several Zernike terms, the value of a Seidel coefficient as obtained from the Zernike coefficients depends on the number of Zernike terms used to represent an aberration function. For example, the Seidel spherical aberration, which varies as ρ^4 , is contained in Zernike polynomials $Z_4^0(\rho)$, $Z_4^2(\rho, \theta)$, $Z_4^4(\rho, \theta)$, $Z_6^0(\rho)$, etc. Given a set of Zernike coefficients c_{nm} of an aberration function, the corresponding Seidel coefficients can be written easily using Table 13.2

$$a_{00} = c_{00} - \sqrt{3}c_{20} + \sqrt{5}c_{40} - \sqrt{7}c_{60} + 3c_{80} + \dots \quad (13.45a)$$

$$a_{11} = 2c_{11} - 4\sqrt{2}c_{31} + 6\sqrt{3}c_{51} - 16c_{71} + \dots \quad (13.45b)$$

$$a_{20} = 2\sqrt{3}c_{20} - \sqrt{6}c_{22} - 6\sqrt{5}c_{40} + 3\sqrt{10}c_{42} + 6\sqrt{7}c_{60} - 6\sqrt{14}c_{62} - 60c_{80} + \dots \quad (13.45c)$$

$$a_{22} = 2\sqrt{6}c_{22} - 6\sqrt{10}c_{42} + 12\sqrt{14}c_{62} + \dots \quad (13.45d)$$

$$a_{31} = 6\sqrt{2}c_{31} - 6\sqrt{2}c_{33} - 24\sqrt{3}c_{51} + 24\sqrt{3}c_{53} + 120c_{71} - 120c_{73} + \dots \quad (13.45e)$$

and

$$a_{40} = 6\sqrt{5}c_{40} - 4\sqrt{10}c_{42} + \sqrt{10}c_{44} - 30\sqrt{7}c_{60} + \dots \quad (13.45f)$$

Once again, the tilt component in the coefficient a_{11} cannot be distinguished from distortion, the defocus component in the coefficient a_{20} cannot be distinguished from the field curvature, and the spherical aberration component in the coefficient a_{40} cannot be distinguished from the lateral spherical aberration.

13.3. ABERRATION FUNCTION OF A SYSTEM WITH A CIRCULAR PUPIL, BUT WITHOUT AN AXIS OF ROTATIONAL SYMMETRY

13.3.1. Zernike Circle Polynomial Expansion

The aberration function for a system without an axis of rotational symmetry will consist of terms not only in $\cos m\theta$ but in $\sin m\theta$ as well. This is true also of aberrations resulting from fabrication errors as well as those introduced by atmospheric turbulence (Noll, 1976). In such cases the aberration function can be expanded in terms of orthonormal Zernike circle polynomials $Z_j(\rho, \theta)$ in the form

$$W(\rho, \theta) = \sum_j a_j Z_j(\rho, \theta) \quad (13.46)$$

where a_j are the *expansion coefficients*. The polynomials may be written

$$Z_{\text{even}j}(\rho, \theta) = \sqrt{2(n+1)} R_n^m(\rho) \cos m\theta, \quad m \neq 0 \quad (13.47a)$$

$$Z_{\text{odd}j}(\rho, \theta) = \sqrt{2(n+1)} R_n^m(\rho) \sin m\theta, \quad m \neq 0 \quad (13.47b)$$

$$Z_j(\rho, \theta) = \sqrt{n+1} R_n^0(\rho), \quad m = 0 \quad (13.47c)$$

where $R_n^m(\rho)$ are the radial polynomials given by Eq. (13.19b). As stated earlier, n and m are positive integers (including zero), and $n - m \geq 0$ and even. The index n represents the radial degree or the order of the polynomial since it represents the highest power of ρ in the polynomial, and m may be called the *azimuthal frequency*.

The index j is a *polynomial-ordering number* and is a function of n and m . The orthonormal polynomials and the relationship among the indices j , n , and m are given in Table 13.5. The polynomials are ordered such that an even j corresponds to a symmetric polynomial varying as $\cos m\theta$, while an odd j corresponds to an antisymmetric polynomial varying as $\sin m\theta$. A polynomial with a lower value of n is ordered first, while for a given value of n , a polynomial with a lower value of m is ordered first. The number of terms for a given value of n is $n + 1$. The number of terms up to and including a certain order n is given by

$$N_n = (n+1)(n+2)/2 \quad (13.48)$$

The orthonormality of Zernike polynomials implies that

$$\int_0^1 \int_0^{2\pi} Z_j(\rho, \theta) Z_{j'}(\rho, \theta) \rho d\rho d\theta \bigg/ \int_0^1 \int_0^{2\pi} \rho d\rho d\theta = \delta_{jj'} \quad (13.49)$$

TABLE 13.5. Orthonormal Zernike circle polynomials $Z_j(\rho, \theta)$. The indices j , n , and m are called the polynomial number, radial degree, and azimuthal frequency, respectively. The polynomials Z_j are ordered such that an even j corresponds to a symmetric polynomial varying as $\cos m\theta$, while an odd j corresponds to an antisymmetric polynomial varying as $\sin m\theta$. A polynomial with a lower value of n is ordered first, and for a given value of n , a polynomial with a lower value of m is ordered first.

j	n	m	$Z_j(\rho, \theta)$	Aberration name
1	0	0	1	Piston
2	1	1	$2\rho \cos \theta$	x tilt
3	1	1	$2\rho \sin \theta$	y tilt
4	2	0	$\sqrt{3}(2\rho^2 - 1)$	Defocus
5	2	2	$\sqrt{6}\rho^2 \sin 2\theta$	Primary astigmatism at 45°
6	2	2	$\sqrt{6}\rho^2 \cos 2\theta$	Primary astigmatism at 0°
7	3	1	$\sqrt{8}(3\rho^3 - 2\rho) \sin \theta$	Primary y coma
8	3	1	$\sqrt{8}(3\rho^3 - 2\rho) \cos \theta$	Primary x coma
9	3	3	$\sqrt{8}\rho^3 \sin 3\theta$	
10	3	3	$\sqrt{8}\rho^3 \cos 3\theta$	
11	4	0	$\sqrt{5}(6\rho^4 - 6\rho^2 + 1)$	Primary spherical
12	4	2	$\sqrt{10}(4\rho^4 - 3\rho^2) \cos 2\theta$	Secondary astigmatism at 0°
13	4	2	$\sqrt{10}(4\rho^4 - 3\rho^2) \sin 2\theta$	Secondary astigmatism at 45°
14	4	4	$\sqrt{10}\rho^4 \cos 4\theta$	
15	4	4	$\sqrt{10}\rho^4 \sin 4\theta$	
16	5	1	$\sqrt{12}(10\rho^5 - 12\rho^3 + 3\rho) \cos \theta$	Secondary x coma
17	5	1	$\sqrt{12}(10\rho^5 - 12\rho^3 + 3\rho) \sin \theta$	Secondary y coma
18	5	3	$\sqrt{12}(5\rho^5 - 4\rho^3) \cos 3\theta$	
19	5	3	$\sqrt{12}(5\rho^5 - 4\rho^3) \sin 3\theta$	
20	5	5	$\sqrt{12}\rho^5 \cos 5\theta$	
21	5	5	$\sqrt{12}\rho^5 \sin 5\theta$	
22	6	0	$\sqrt{7}(20\rho^6 - 30\rho^4 + 12\rho^2 - 1)$	Secondary spherical
23	6	2	$\sqrt{14}(15\rho^6 - 20\rho^4 + 6\rho^2) \sin 2\theta$	Tertiary astigmatism at 45°
24	6	2	$\sqrt{14}(15\rho^6 - 20\rho^4 + 6\rho^2) \cos 2\theta$	Tertiary astigmatism at 0°
25	6	4	$\sqrt{14}(6\rho^6 - 5\rho^4) \sin 4\theta$	
26	6	4	$\sqrt{14}(6\rho^6 - 5\rho^4) \cos 4\theta$	
27	6	6	$\sqrt{14}\rho^6 \sin 6\theta$	
28	6	6	$\sqrt{14}\rho^6 \cos 6\theta$	
29	7	1	$4(35\rho^7 - 60\rho^5 + 30\rho^3 - 4\rho) \sin \theta$	Tertiary y coma
30	7	1	$4(35\rho^7 - 60\rho^5 + 30\rho^3 - 4\rho) \cos \theta$	Tertiary x coma
31	7	3	$4(21\rho^7 - 30\rho^5 + 10\rho^3) \sin 3\theta$	
32	7	3	$4(21\rho^7 - 30\rho^5 + 10\rho^3) \cos 3\theta$	
33	7	5	$4(7\rho^7 - 6\rho^5) \sin 5\theta$	
34	7	5	$4(7\rho^7 - 6\rho^5) \cos 5\theta$	
35	7	7	$4\rho^7 \sin 7\theta$	
36	7	7	$4\rho^7 \cos 7\theta$	
37	8	0	$3(70\rho^8 - 140\rho^6 + 90\rho^4 - 20\rho^2 + 1)$	Tertiary spherical

The orthogonality of the angular functions is described by

$$\int_0^{2\pi} d\phi \begin{cases} \cos m\phi \cos m'\phi, & j \text{ and } j' \text{ are both even} \\ \cos m\phi \sin m'\phi, & j \text{ is even and } j' \text{ is odd} \\ \sin m\phi \cos m'\phi, & j \text{ is odd and } j' \text{ is even} \\ \sin m\phi \sin m'\phi, & j \text{ and } j' \text{ are both odd} \end{cases} = \begin{cases} \pi(1 + \delta_{m0})\delta_{mm'}, & j \text{ and } j' \text{ are both even} \\ \pi\delta_{mm'}, & j \text{ and } j' \text{ are both odd} \\ 0, & \text{otherwise} \end{cases} \quad (13.50)$$

The expansion coefficients are given by

$$a_j = \frac{1}{\pi} \int_0^1 \int_0^{2\pi} W(\rho, \theta) Z_j(\rho, \theta) \rho d\rho d\theta \quad (13.51)$$

as may be seen by substituting Eq. (13.46) and utilizing the orthogonality relations. The aberration variance is given by

$$\sigma_W^2 = \langle W^2(\rho, \theta) \rangle - \langle W(\rho, \theta) \rangle^2 = \sum_{j=2} a_j^2 \quad (13.52)$$

Because of the orthonormality of the Zernike polynomials, an expansion coefficient a_j of an aberration term $a_j Z_j(\rho, \theta)$, with the exception of a_1 , represents its standard deviation. Moreover, its value is independent of the values of the other coefficients or the number of polynomials used to represent the aberration function. Accordingly, if one or more aberration terms are added or subtracted from the aberration function, the values of these coefficients do not change.

The orthonormal Zernike circle polynomials $Z_j(x, y)$ in Cartesian coordinates (x, y) , which are normalized by the pupil radius a , are given in Table 13.6. The subscript n on the coordinates is omitted for simplicity.

13.3.2. Relationships Among the Indices n , m , and j

For a given value of n , the number of polynomials N_n represents the largest value of j . Since the number of polynomials with the same value of n but different values of m is equal to $n + 1$, the smallest value of j for a given value of n is $N_n - n$. For a given value of n and m , there are two j values, $N_n - n + m - 1$ and $N_n - n + m$. The even value of j represents the $\cos m\theta$ polynomials and the odd value of j represents the $\sin m\theta$ polynomials. The value of j with $m = 0$ is $N_n - n$. For example, for $n = 5$, $N_n = 21$ and $j = 21$ represents the $\sin 5\theta$ polynomials. The number of the corresponding $\cos 5\theta$ polynomials is $j = 20$. The two polynomials with $m = 3$, for example, have j values of 18 and 19, representing the $\cos 3\theta$ and the $\sin 3\theta$ polynomials, respectively.

TABLE 13.6. Orthonormal Zernike circle polynomials $Z_j(x, y)$ in Cartesian coordinates (x, y) , which are normalized by the pupil radius a . The subscript n on the coordinates indicating the normalization is omitted for simplicity. Note that $x = \rho \cos \theta$, $y = \rho \sin \theta$, and $0 \leq \rho = \sqrt{x^2 + y^2} \leq 1$.

Polynomial	$Z_j(x, y)$
Z_1	1
Z_2	$2x$
Z_3	$2y$
Z_4	$\sqrt{3}(2\rho^2 - 1)$
Z_6	$2\sqrt{6}xy$
Z_6	$\sqrt{6}(x^2 - y^2)$
Z_7	$\sqrt{8}y(3\rho^2 - 2)$
Z_8	$\sqrt{8}x(3\rho^2 - 2)$
Z_9	$\sqrt{8}y(3x^2 - y^2)$
Z_{10}	$\sqrt{8}x(x^2 - 3y^2)$
Z_{11}	$\sqrt{5}(6\rho^4 - 6\rho^2 + 1)$
Z_{12}	$\sqrt{10}(x^2 - y^2)(4\rho^2 - 3)$
Z_{13}	$2\sqrt{10}xy(4\rho^2 - 3)$
Z_{14}	$\sqrt{10}(\rho^4 - 8x^2y^2)$
Z_{15}	$4\sqrt{10}xy(x^2 - y^2)$
Z_{16}	$\sqrt{12}x(10\rho^4 - 12\rho^2 + 3)$
Z_{17}	$\sqrt{12}y(10\rho^4 - 12\rho^2 + 3)$
Z_{18}	$\sqrt{12}x(x^2 - 3y^2)(5\rho^2 - 4)$
Z_{19}	$\sqrt{12}y(3x^2 - y^2)(5\rho^2 - 4)$
Z_{20}	$\sqrt{12}x(16x^4 - 20x^2\rho^2 + 5\rho^4)$
Z_{21}	$\sqrt{12}y(16y^4 - 20y^2\rho^2 + 5\rho^4)$
Z_{22}	$\sqrt{7}(20\rho^6 - 30\rho^4 + 12\rho^2 - 1)$
Z_{23}	$2\sqrt{14}xy(15\rho^4 - 20\rho^2 + 6)$
Z_{24}	$\sqrt{14}(x^2 - y^2)(15\rho^4 - 20\rho^2 + 6)$
Z_{25}	$4\sqrt{14}xy(x^2 - y^2)(6\rho^2 - 5)$
Z_{26}	$\sqrt{14}(8x^4 - 8x^2\rho^2 + \rho^4)(6\rho^2 - 5)$
Z_{27}	$\sqrt{14}xy(32x^4 - 32x^2\rho^2 + 6\rho^4)$
Z_{28}	$\sqrt{14}(32x^6 - 48x^4\rho^2 + 18x^2\rho^4 - \rho^6)$

For a given value of j , n is given by

$$n = [(2j - 1)^{1/2} + 0.5]_{\text{integer}} - 1 \quad (13.53)$$

where the subscript integer implies the integer value of the number in brackets. Once n is known, the value of m is given by

$$m = \begin{cases} 2\{[2j + 1 - n(n + 1)]/4\}_{\text{integer}} & \text{when } n \text{ is even} \\ 2\{[2(j + 1) - n(n + 1)]/4\}_{\text{integer}} - 1 & \text{when } n \text{ is odd} \end{cases} \quad (13.54a)$$

$$(13.54b)$$

For example, suppose we want to know the values of n and m for the polynomial $j = 10$. From Eq. (13.53), $n = 3$ and from Eq. (13.54b), $m = 3$. Hence, it is a $\cos 3\theta$ polynomial. Similarly, for example, for $j = 25$, we find that $n = 6$ and $m = 3$, representing the $\sin 3\theta$ polynomial.

13.3.3. Isometric, Interferometric, and PSF Plots for a Zernike Circle Polynomial Aberration

The Zernike circle polynomials for $n \leq 6$ are shown in the form of a pyramid in Figure 13.4. If we number the rows in the pyramid from top to bottom, the row number represents the value of n for the polynomials in that row. The number of polynomials in a certain row is $n + 1$, representing different values of m and the fact that, for a given value of m , there is a cosine polynomial with an even j number and a sine polynomial with an odd j number. Each polynomial is illustrated in three different but equivalent ways. For each polynomial, the isometric plot at the top illustrates its shape as produced, for example, in a deformable mirror. The standard deviation of each polynomial aberration in the figure is one wave. An interferogram, as in optical testing, is shown on the left. The number of fringes, which is equal to the

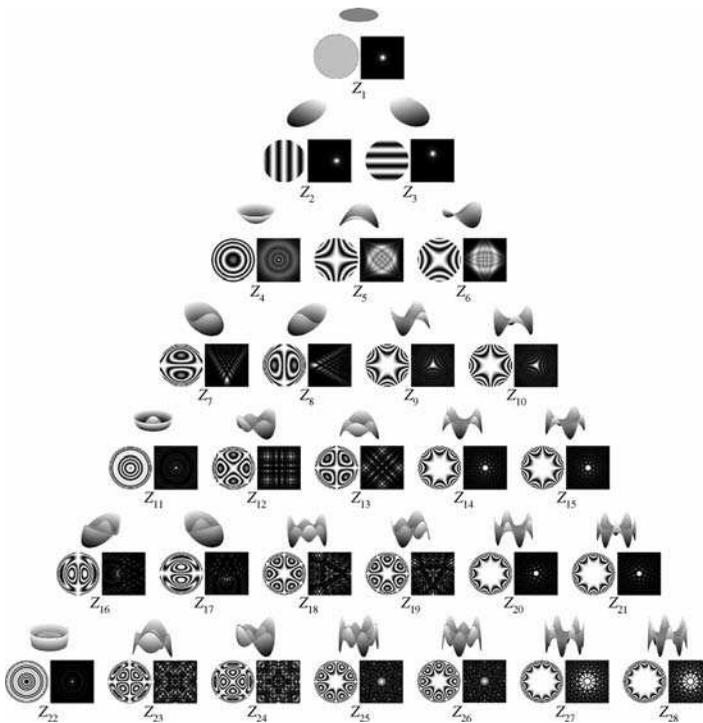


FIGURE 13.4. Zernike circle polynomials pyramid showing isometric plot on the top, interferogram on the left, and PSF on the right for each polynomial. The standard deviation of each polynomial aberration is one wave.

number of times the aberration changes by one wave as we move from the center to the edge of the pupil, is different for the different polynomials. Each fringe represents a contour of constant phase or aberration. The fringe is dark when the phase is an odd multiple of π or the aberration is an odd multiple of $\lambda/2$. In the case of tilts, for example, the aberration changes by one four times, which is the same as the peak-to-valley value of 4 waves. Hence, 4 straight line fringes symmetric about the center are obtained. The x -tilt polynomial Z_2 yields vertical fringes and y -tilt polynomial Z_3 yields horizontal fringes. Similarly, defocus aberration Z_4 yields about 3.5 fringes. In the case of spherical aberration Z_{11} , the aberration starts at a value of $\sqrt{5}$, decreases to zero, reaches a negative value of $-\sqrt{5}/2$, and then increases to $\sqrt{5}$. Hence, the total number of times the aberration changes by unity is equal to 6.7, and approximately seven circular fringes are obtained.

On the right for each polynomial are shown the PSFs (point-spread function), which represent the images of a point object in the presence of a polynomial aberration. Thus, for example, piston yields the aberration-free PSF (since it has no effect on the PSF), and the x - and y -tilts displace the aberration-free PSF along the x and y axes, respectively. More light has been introduced into some of the PSFs to display their details.

13.3.4. Primary Zernike Aberrations and Their Relationships with Seidel Aberrations

The names of some of the Zernike aberrations are given in Table 13.5. The Zernike tilt aberration

$$a_2 Z_2(\rho, \theta) = 2a_2 \rho \cos \theta \quad (13.55)$$

represents x tilt and corresponds to a displacement of the center of the reference sphere along the x axis by $4Fa_2$ or a tilt of the wavefront about the y axis by an angle $\beta = 2a_2/a$. Similarly, the Zernike tilt aberration

$$a_3 Z_3(\rho, \theta) = 2a_3 \rho \sin \theta \quad (13.56)$$

represents y tilt and corresponds to a displacement of the center of the reference sphere along the y axis by $4Fa_3$ or a tilt of the wavefront about the x axis by an angle $\alpha = 2a_3/a$. The Zernike primary astigmatism

$$a_6 Z_6(\rho, \theta) = \sqrt{6}a_6 \rho^2 \cos 2\theta \quad (13.57)$$

is referred to as the x or 0° astigmatism. It consists of Seidel astigmatism $\rho^2 \cos^2 \theta$ balanced with defocus aberration ρ^2 to yield minimum variance. It yields a uniform circular spot diagram, but a line sagittal image along the x axis (i.e., in a plane that zeros out the defocus part). The Zernike primary astigmatism

$$a_5 Z_5(\rho, \theta) = \sqrt{6}a_5 \rho^2 \sin 2\theta \quad (13.58)$$

can be written

$$a_5 Z_5(\rho, \theta) = \sqrt{6} a_5 \rho^2 \cos[2(\theta + \pi/4)] \quad (13.59)$$

Comparing with Eq. (13.57), it is equivalent to changing θ to $\theta + \pi/4$. Accordingly, it is called the 45° astigmatism or astigmatism at 45° . The secondary Zernike astigmatism given by

$$a_{12} Z_{12}(\rho, \theta) = \sqrt{10}(4\rho^4 - 3\rho^2) \cos 2\theta$$

does not yield a line image in any plane. However, it is referred to as the x or the 0° astigmatism in conformance with the primary x astigmatism, because of its variation with θ as $\cos 2\theta$. Similarly, the other forms of astigmatism are referred to as indicated in Table 13.2. The Zernike coma terms $a_8 Z_8(\rho, \theta)$ and $a_7 Z_7(\rho, \theta)$ are called the x and y Zernike comas. They yield PSFs that are symmetric about the x and y axes, respectively. Similarly, the names for the secondary and tertiary coma can be explained.

It should be evident that when the cosine and sine terms of a certain aberration are present simultaneously, as in optical testing, their combination represents the aberration whose orientation depends on the value of the component terms (Wyant and Creath, 1992; Mahajan and Swantner, 2006). For example, if both x and y Zernike tilts are present in the form

$$W(\rho, \theta) = a_2 Z_2(\rho, \theta) + a_3 Z_3(\rho, \theta) \quad (13.60a)$$

$$= 2a_2 \rho \cos \theta + 2a_3 \rho \sin \theta \quad (13.60b)$$

it can be written as

$$W(\rho, \theta) = 2(a_2^2 + a_3^2)^{1/2} \rho \cos[\theta - \tan^{-1}(a_3/a_2)] \quad (13.61)$$

Thus, it represents a wavefront tilt of $2(a_2^2 + a_3^2)^{1/2}$ about an axis that is orthogonal to a line making an angle of $\tan^{-1}(a_3/a_2)$ with the x axis.

It is easy to see that when both a_2 and a_3 are negative, $(a_2^2 + a_3^2)^{1/2}$ in Eq. (13.61) must be replaced by $-(a_2^2 + a_3^2)^{1/2}$. However, when one of the coefficients is positive and the other is negative, then $\tan^{-1}(a_3/a_2)$ of a negative argument has two solutions: a negative acute angle or its complementary angle. The choice is made depending on whether a_2 or a_3 is negative according to

$$\tan^{-1}(a_3/a_2) = \begin{cases} -\tan^{-1}(|a_3/a_2|) & \text{for positive } a_2 \text{ and negative } a_3 \\ \pi - \tan^{-1}(|a_3/a_2|) & \text{for negative } a_2 \text{ and positive } a_3 \end{cases} \quad (13.62a)$$

$$(13.62b)$$

An alternative when a_2 is negative is to let the angle be $-\tan^{-1}(|a_3/a_2|)$, as when a_2 is positive, but also replace $(a_2^2 + a_3^2)^{1/2}$ by $-(a_2^2 + a_3^2)^{1/2}$.

When both x and y Zernike comas are present, the aberration may be written as

$$W(\rho, \theta) = a_8 Z_8(\rho, \theta) + a_7 Z_7(\rho, \theta) \quad (13.63a)$$

$$= \sqrt{8} a_8 (3\rho^3 - 2\rho) \cos \theta + \sqrt{8} a_7 (3\rho^3 - 2\rho) \sin \theta \quad (13.63b)$$

$$= (a_7^2 + a_8^2)^{1/2} \sqrt{8} (3\rho^3 - 2\rho) \cos[\theta - \tan^{-1}(a_7/a_8)] \quad (13.63c)$$

which is equivalent to a Zernike coma of magnitude $(a_7^2 + a_8^2)^{1/2}$ inclined at an angle of $\tan^{-1}(a_7/a_8)$ with the x axis. Similarly, if both x and 45° astigmatisms are present so that

$$W(\rho, \theta) = a_6 Z_6(\rho, \theta) + a_5 Z_5(\rho, \theta) \quad (13.64a)$$

$$= \sqrt{6} a_6 \rho^2 \cos 2\theta + \sqrt{6} a_5 \rho^2 \sin 2\theta \quad (13.64b)$$

we may write it in the form

$$W(\rho, \theta) = (a_5^2 + a_6^2)^{1/2} \sqrt{6} \rho^2 \cos \left\{ 2 \left[\theta - \frac{1}{2} \tan^{-1}(a_5/a_6) \right] \right\} \quad (13.65)$$

showing that it is Zernike astigmatism of magnitude $(a_5^2 + a_6^2)^{1/2}$ at an angle of $(1/2) \tan^{-1}(a_5/a_6)$.

It should be noted that the wavefront tilt aberration given by given by Eq. (13.61) represents the tilt aberration obtained from Zernike tilt aberrations. However, there are other Zernike aberrations that also contain tilt aberration built into them, for example, Zernike primary, secondary, or tertiary coma. Similarly, the Seidel coma $3\sqrt{8}(a_7^2 + a_8^2)^{1/2}$ in Eq. (13.63c) at an angle of $\tan^{-1}(a_7/a_8)$ is only from the primary Zernike comas. But the secondary and tertiary Zernike comas also contain Seidel coma. Hence, only if the higher-order Zernike comas are zero or negligible, the PSF aberrated by primary Zernike coma will be symmetric about a line making angle of $\tan^{-1}(a_7/a_8)$ with the x axis. Similarly, only if the secondary and tertiary astigmatisms are zero or negligible, the Seidel astigmatism is $2\sqrt{6}(a_5^2 + a_6^2)^{1/2}$, as in Eq. (13.65). It yields an aberrated PSF which is symmetric about two orthogonal axes one of which is along a line that makes an angle of $(1/2) \tan^{-1}(a_5/a_6)$ with the x axis. In the case of coma and astigmatism, the same rules apply for determining the angle $\tan^{-1}(\cdot)$, as in the case of a wavefront tilt aberration.

It should be evident that there is ambiguity in determining astigmatism, since it can be written in different but equivalent forms by separating defocus aberration from it. For example, a 0° astigmatism can be written as

$$a_6 Z_6(\rho, \theta) = a_6 (\sqrt{6} \rho^2 \cos 2\theta) \quad (13.66a)$$

$$= a_6 \sqrt{6} (2\rho^2 \cos^2 \theta - \rho^2) \quad (13.66b)$$

$$= a_6 \sqrt{6} (-2\rho^2 \sin^2 \theta + \rho^2) \quad (13.66c)$$

It is clear that a 0° Zernike astigmatism given by Eq. (13.66a) can be written as a combination of 0° positive Seidel astigmatism and a negative defocus, as in Eq. (13.66b), or a 90° negative Seidel astigmatism and a positive defocus, as in Eq. (13.66c).

To illustrate how a wrong Seidel coefficient can be inferred unless it is obtained from all of the significant Zernike terms, we consider an axial image aberrated by one wave of secondary spherical aberration ρ^6 . In terms of Zernike polynomials it will be written as

$$W(\rho) = a_{22}Z_{22}(\rho) + a_{11}Z_{11}(\rho) + a_4Z_4(\rho) + a_1Z_1(\rho) \quad (13.67a)$$

where

$$a_{22} = 1/20\sqrt{7}, \quad a_{11} = 1/4\sqrt{5}, \quad a_4 = (9/20\sqrt{3}), \quad a_1 = 1/4 \quad (13.67b)$$

If we infer the Seidel spherical aberration from only the primary Zernike aberration $a_{11}Z_{11}(\rho)$, its amount would be 1.5 waves. Such a conclusion is obviously incorrect, since in reality the amount of Seidel spherical aberration is zero. Needless to say, if we expand the aberration function up to the first, say, as many as 21 terms, we will in fact incorrectly conclude that the amount of Seidel spherical aberration is 1.5 waves. However, the Seidel spherical aberration will correctly reduce to zero when at least the first 22 terms are included in the expansion. For an off-axis image, there are angle-dependent aberrations, for example, Z_{14} , that also contain Seidel aberrations. Hence, it is important that the expansion be carried out up to a certain number of terms such that any additional terms do not change significantly the mean square difference between the function and its estimate. Otherwise, the inferred Seidel aberrations will be erroneous.

If we approximate a certain aberration function by the primary Zernike aberrations only, we may write

$$W(\rho, \theta) = \sum_{j=1}^{j=8} a_j Z_j(\rho, \theta) + a_{11} Z_{11}(\rho) \quad (13.68a)$$

$$\begin{aligned} &= A_p + A_t \rho \cos(\theta - \beta_t) + A_d \rho^2 + A_a \rho^2 \cos^2(\theta - \beta_a) \\ &\quad + A_c \rho \cos(\theta - \beta_c) + A_s \rho^4 \end{aligned} \quad (13.68b)$$

where A_p is the piston aberration, other coefficients A_i represent the peak value of the corresponding Seidel aberration term, and β_i is the orientation angle of the Seidel aberration. They are given by

$$A_p = a_1 - \sqrt{3}a_4 + \sqrt{5}a_{11} \quad (13.69a)$$

$$A_t = 2[(a_2 - \sqrt{8}a_8)^2 + (a_3 - \sqrt{8}a_7)^2]^{1/2}, \quad \beta_t = \tan^{-1} \left(\frac{a_3 - \sqrt{8}a_7}{a_2 - \sqrt{8}a_8} \right) \quad (13.69b)$$

$$A_d = 2(\sqrt{3}a_4 - 3\sqrt{5}a_{11} - A_a) \quad (13.69c)$$

$$A_a = 2\sqrt{6}(a_5^2 + a_6^2)^{1/2}, \quad \beta_a = \frac{1}{2} \tan^{-1}(a_5/a_6) \quad (13.69d)$$

$$A_c = 6\sqrt{2}(a_7^2 + a_8^2)^{1/2}, \quad \beta_c = \tan^{-1}(a_7/a_8) \quad (13.69e)$$

and

$$A_s = 6\sqrt{5}a_{11} \quad (13.69f)$$

The approximation is good only when the higher-order aberrations are negligible, since they also contain Seidel aberration terms.

13.4. ZERNIKE ANNULAR POLYNOMIALS AS BALANCED ABERRATIONS FOR SYSTEMS WITH ANNULAR PUPILS

13.4.1. Balanced Aberrations

For an annular pupil with an obscuration ratio of ϵ , the radial variable $\epsilon \leq \rho \leq 1$. Accordingly, the mean and mean square values of an aberration $W(\rho, \theta; \epsilon)$ are given by (Mahajan, 1981a,b, 1984, 1994, 2003, 2004)

$$\langle W^k \rangle = \frac{1}{\pi(1-\epsilon^2)} \int_{\epsilon}^1 \int_0^{2\pi} W^k(\rho, \theta; \epsilon) \rho d\rho d\theta \quad (13.70)$$

with $k = 1$ and 2 , respectively. Table 13.7 lists the standard deviation of primary aberrations. Table 13.8 lists the corresponding balanced aberrations, their standard deviation, and the diffraction focus (i.e., the point of maximum central irradiance in the case of a small aberration). We note that the balancing defocus in the case of spherical aberration or the balancing wavefront tilt in the case of coma depends on the value of ϵ . They both increase as ϵ increases. The amount of balancing defocus is larger by a factor of $1 + \epsilon^2$ compared to that for a circular pupil. The amount of balancing defocus in the case of astigmatism, is independent of ϵ . These results reduce to those for circular pupils as $\epsilon \rightarrow 0$.

13.4.2. Zernike Annular Polynomials

The aberration function of a rotationally symmetric system with an annular pupil can be expanded in terms of a complete set of Zernike annular polynomials

TABLE 13.7. Primary aberrations and their standard deviation for systems with annular pupils.

Aberration	$W(\rho, \theta)$	σ_w
Spherical	$A_s \rho^4$	$(4 - \epsilon^2 - 6\epsilon^4 - \epsilon^6 + 4\epsilon^8)^{1/2} A_s / 3\sqrt{5}$
Coma	$A_c \rho^3 \cos \theta$	$(1 + \epsilon^2 + \epsilon^4 + \epsilon^6)^{1/2} A_c / 2\sqrt{2}$
Astigmatism	$A_a \rho^2 \cos^2 \theta$	$(1 + \epsilon^2)^{1/2} A_a / 4$
Field curvature	$A_d \rho^2$	$(1 - \epsilon^2) A_d / 2\sqrt{3}$
Distortion	$A_t \rho \cos \theta$	$(1 + \epsilon^2)^{1/2} A_t / 2$

TABLE 13.8. Balanced primary aberrations, their standard deviation, and diffraction focus for systems with annular pupils.

Balanced aberration	$W(\rho, \theta)$	σ_w	Diffraction focus
Spherical	$A_s[\rho^2 - (1 + \epsilon^2)\rho^2]$	$\frac{1}{6\sqrt{5}}(1 - \epsilon^2)^2 A_s$	$[0, 0, 8(1 + \epsilon^2)F^2 A_s]$
Coma	$A_c \left(\rho^3 - \frac{2}{3} \frac{1 + \epsilon^2 + \epsilon^4}{1 + \epsilon^2} \rho \right) \cos \theta$	$\frac{(1 - \epsilon^2)(1 + 4\epsilon^2 + \epsilon^4)^{1/2}}{6\sqrt{2}(1 + \epsilon^2)^{1/2}} A_c$	$\left[\frac{4(1 + \epsilon^2 + \epsilon^4)}{3(1 + \epsilon^2)} F A_c, 0, 0 \right]$
Astigmatism	$A_a \rho^2 (\cos^2 \theta - 1/2)$	$\frac{1}{2\sqrt{6}}(1 + \epsilon^2 + \epsilon^4)^{1/2} A_a$	$(0, 0, 4F^2 A_a)$

$R_n^m(\rho; \epsilon) \cos m\theta$ that are orthogonal over the annulus in the form (Mahajan, 1981a,b, 1984, 1994, 2003, 2004)

$$W(\rho, \theta; \epsilon) = \sum_{n=0}^{\infty} \sum_{m=0}^n c_{nm} Z_n^m(\rho, \theta, \epsilon) \quad (13.71)$$

where c_{nm} are the expansion coefficients, n and m are positive integers, $n - m \geq 0$ and even, and

$$Z_n^m(\rho, \theta, \epsilon) = [2(n+1)/(1 + \delta_{m0})]^{1/2} R_n^m(\rho; \epsilon) \cos m\theta \quad (13.72)$$

is an orthonormal Zernike annular polynomial obeying the orthonormality relation

$$\frac{1}{\pi(1 - \epsilon^2)} \int_{\epsilon}^1 \int_0^{2\pi} Z_n^m(\rho, \theta, \epsilon) Z_{n'}^{m'}(\rho, \theta, \epsilon) \rho d\rho d\theta = \delta_{nm} \delta_{n'm'} \quad (13.73)$$

Thus, the radial polynomials obey the orthogonality relation

$$\int_{\epsilon}^1 R_n^m(\rho; \epsilon) R_{n'}^m(\rho; \epsilon) \rho d\rho = \frac{1 - \epsilon^2}{2(n+1)} \delta_{nn'} \quad (13.74)$$

The Zernike annular polynomials can be obtained from the corresponding circle polynomials by the Gram–Schmidt orthogonalization process according to (Korn and Korn, 1968)

$$R_n^m(\rho; \epsilon) = N_n^m \left[R_n^m(\rho) - \sum_{i \geq 1}^{(n-m)/2} (n - 2i + 1) \langle R_n^m(\rho) R_{n-2i}^m(\rho; \epsilon) \rangle R_{n-2i}^m(\rho, \epsilon) \right] \quad (13.75)$$

where

$$\langle R_n^m(\rho)R_n^m(\rho; \epsilon) \rangle = \frac{2}{1 - \epsilon^2} \int_{\epsilon}^1 R_n^m(\rho)R_n^m(\rho; \epsilon)\rho d\rho \quad (13.76)$$

and N_n^m is a normalization constant such that the radial polynomials satisfy the orthogonality Eq. (13.74). For $m = 0$, the radial polynomials are equal to the Legendre polynomials $P_n(\cdot)$ according to

$$R_{2n}^0(\rho; \epsilon) = P_n \left[\frac{2(\rho^2 - \epsilon^2)}{1 - \epsilon^2} - 1 \right] \quad (13.77)$$

Thus, they can be obtained from the circle radial polynomials $R_{2n}^0(\rho)$ by replacing ρ by $[(\rho^2 - \epsilon^2)/(1 - \epsilon^2)]^{1/2}$; that is,

$$R_{2n}^0(\rho; \epsilon) = R_{2n}^0 \left[\left(\frac{\rho^2 - \epsilon^2}{1 - \epsilon^2} \right)^{1/2} \right] \quad (13.78)$$

It can be seen from Eqs. (13.74) and (13.75) that

$$R_n^n(\rho; \epsilon) = \rho^n / \left(\sum_{i=0}^n \epsilon^{2i} \right)^{1/2} \quad (13.79a)$$

$$= \rho^n \{ (1 - \epsilon^2) / [1 - \epsilon^{2(n+1)}] \}^{1/2} \quad (13.79b)$$

Moreover,

$$R_n^{n-2}(\rho; \epsilon) = \frac{n\rho^n - (n-1)[(1 - \epsilon^{2n})/(1 - \epsilon^{2(n-1)})]\rho^{n-2}}{\{(1 - \epsilon^2)^{-1}[n^2(1 - \epsilon^{2(n+1)}) - (n^2 - 1)(1 - \epsilon^{2n})^2/(1 - \epsilon^{2(n-1)})]\}^{1/2}} \quad (13.80)$$

It is evident that the annular radial polynomial $R_n^n(\rho; \epsilon)$ differs from the corresponding circle radial polynomial $R_n^n(\rho)$ only in its normalization. We also note that

$$R_n^m(1; \epsilon) = 1, \quad m = 0 \quad (13.81a)$$

$$\neq 1, \quad m \neq 0 \quad (13.81b)$$

The annular polynomials are similar to the circle polynomials, except that they are orthogonal over an annular pupil. Thus, $R_n^m(\rho; \epsilon)$ is a radial polynomial of degree n in ρ containing terms in $\rho^n, \rho^{n-2}, \dots$, and ρ^m with coefficients that depend on ϵ . The annular polynomials are unique in the same manner as the circle polynomials. The radial Zernike annular polynomials for $n \leq 6$ are listed in Table 13.9.

TABLE 13.9. Radial Zernike annular polynomials $R_n^m(\rho; \epsilon)$, where $\epsilon \leq \rho \leq 1$.

n	m	$R_n^m(\rho; \epsilon)$
0	0	1
1	1	$\rho/(1 + \epsilon^2)^{1/2}$
2	0	$(2\rho^2 - 1 - \epsilon^2)/(1 - \epsilon^2)$
2	2	$\rho^2/(1 + \epsilon^2 + \epsilon^4)^{1/2}$
3	1	$\frac{3(1 + \epsilon^2)\rho^3 - 2(1 + \epsilon^2 + \epsilon^4)\rho}{(1 - \epsilon^2)[(1 + \epsilon^2)(1 + 4\epsilon^2 + \epsilon^4)]^{1/2}}$
3	3	$\rho^3/(1 + \epsilon^2 + \epsilon^4 + \epsilon^6)^{1/2}$
4	0	$[6\rho^4 - 6(1 + \epsilon^2)\rho^2 + 1 + 4\epsilon^2 + \epsilon^4]/(1 - \epsilon^2)^2$
4	2	$\frac{4\rho^4 - 3[(1 - \epsilon^8)/(1 - \epsilon^6)]\rho^2}{\{(1 - \epsilon^2)^{-1}[16(1 - \epsilon^{10}) - 15(1 - \epsilon^8)^2/(1 - \epsilon^6)]\}^{1/2}}$
4	4	$\rho^4/(1 + \epsilon^2 + \epsilon^4 + \epsilon^6 + \epsilon^8)^{1/2}$
5	1	$\frac{10(1 + 4\epsilon^2 + \epsilon^4)\rho^5 - 12(1 + 4\epsilon^2 + 4\epsilon^4 + \epsilon^6)\rho^3 + 3(1 + 4\epsilon^2 + 10\epsilon^4 + 4\epsilon^6 + \epsilon^8)\rho}{(1 - \epsilon^2)^2[(1 + 4\epsilon^2 + \epsilon^4)(1 + 9\epsilon^2 + 9\epsilon^4 + \epsilon^6)]^{1/2}}$
5	3	$\frac{5\rho^5 - 4[(1 - \epsilon^{10})/(1 - \epsilon^8)]\rho^3}{\{(1 - \epsilon^2)^{-1}[25(1 - \epsilon^{1/2}) - 24(1 - \epsilon^{10})^2/(1 - \epsilon^8)]\}^{1/2}}$
5	5	$\rho^5/(1 + \epsilon^2 + \epsilon^4 + \epsilon^6 + \epsilon^8 + \epsilon^{10})^{1/2}$
6	0	$[20\rho^6 - 30(1 + \epsilon^2)\rho^4 + 12(1 + 3\epsilon^2 + \epsilon^4)\rho^2 - (1 + 9\epsilon^2 + 9\epsilon^4 + \epsilon^6)]/(1 - \epsilon^2)^3$ $15(1 + 4\epsilon^2 + 10\epsilon^4 + 4\epsilon^6 + \epsilon^8)\rho^6 - 20(1 + 4\epsilon^2 + 10\epsilon^4 + 10\epsilon^6 + 4\epsilon^8 + \epsilon^{10})\rho^4$
6	2	$\frac{+6(1 + 4\epsilon^2 + 10\epsilon^4 + 20\epsilon^6 + 10\epsilon^8 + 4\epsilon^{10} + \epsilon^{12})\rho^2}{(1 - \epsilon^2)^2[(1 + 4\epsilon^2 + 10\epsilon^4 + 4\epsilon^6 + \epsilon^8)(1 + 9\epsilon^2 + 45\epsilon^4 + 65\epsilon^6 + 45\epsilon^8 + 9\epsilon^{10} + \epsilon^{12})]^{1/2}}$
6	4	$\frac{6\rho^6 - 5[(1 - \epsilon^{12})/(1 - \epsilon^{10})]\rho^4}{\{(1 - \epsilon^2)^{-1}[36(1 - \epsilon^{14}) - 35(1 - \epsilon^{12})^2/(1 - \epsilon^{10})]\}^{1/2}}$
6	6	$\rho^6/(1 + \epsilon^2 + \epsilon^4 + \epsilon^6 + \epsilon^8 + \epsilon^{10} + \epsilon^{12})^{1/2}$

The expansion coefficients are given by

$$c_{nm} = \frac{1}{\pi(1 - \epsilon^2)} \int_{\epsilon}^1 \int_0^{2\pi} W(\rho, \theta, \epsilon) Z_n^m(\rho, \theta, \epsilon) \rho d\rho d\theta \quad (13.82)$$

as may be seen by substituting Eq. (13.71) and utilizing the orthogonality Eq. (13.73). The aberration variance is given by

$$\sigma_W^2 = \langle W^2(\rho, \theta, \epsilon) \rangle - \langle W(\rho, \theta, \epsilon) \rangle^2 = \sum_{n=1}^{\infty} \sum_{m=0}^n c_{nm}^2 \quad (13.83)$$

Once again, the balanced primary aberrations can be identified with the corresponding Zernike annular polynomials. For example, balanced spherical aberration is similar to the Zernike annular polynomial $Z_4^0(\rho, \theta; \epsilon)$. Similarly, balanced coma and balanced astigmatism are similar to the polynomials $Z_3^1(\rho, \theta; \epsilon)$ and $Z_2^2(\rho, \theta; \epsilon)$, respectively. Moreover, the field curvature (and defocus) and distortion (and wavefront tilt) aberrations can be identified with $Z_2^0(\rho, \theta; \epsilon)$ and $Z_1^1(\rho, \theta; \epsilon)$, respectively. The piston polynomial $Z_0^0(\rho, \theta; \epsilon)$ is simply equal to unity.

Figure 13.5 shows how a radial annular polynomial $R_n^m(\rho; \epsilon)$ varies with ρ for $\epsilon = 0.5$, that is, for $0.5 \leq \rho \leq 1$. Compared to a circle polynomial (see Figure 13.3), the value of an annular polynomial at the outer edge of the pupil $\rho = 1$ is unity only when $m = 0$. Its value in the case of tilt, coma, and astigmatism aberrations at the inner edge of the pupil $\rho = 0.5$ is, of course, not zero.

If a system has an annular pupil without an axis of rotational symmetry, for example, an arbitrarily deformed annular mirror, then the polynomials with $\sin m\theta$ terms are also included in the expansion of the aberration function. The orthonormal Zernike annular polynomials $Z_j(\rho, \theta; \epsilon)$ in polar coordinates and their numbering are listed in Table 13.10. Their numbering or sequencing is exactly as for the Zernike circle polynomials given in Table 13.5. The corresponding polynomials $Z_j(x, y; \epsilon)$ in Cartesian coordinates (x, y) , which are normalized by the outer radius a of the pupil, are given in Table 13.11. The subscript n on the coordinates is omitted for simplicity.

13.4.3. Isometric, Interferometric, and PSF Plots for a Zernike Annular Polynomial Aberration

The Zernike annular polynomials for $n \leq 6$ are shown in the form of a pyramid in Figure 13.6. As in Figure 13.4 for Zernike circle polynomials, the row number represents the value of n and the number of terms in a certain row is $n + 1$. Each annular polynomial is illustrated in three different but equivalent ways. The value of ϵ in this figure is 0.5. For each polynomial, the isometric plot at the top illustrates its shape as produced, for example, in a deformable mirror. The standard deviation of each polynomial aberration in the figure is one wave. An interferogram, as in optical testing, is shown on the left. Because of the obscuration, the number of fringes is different from the corresponding number in the interferograms for a circular pupil. The number of fringes in this case is equal to the number of times the aberration changes by one wave as we move from the inner edge of the annular pupil to its outer edge. The PSF (point-spread function) on the right represents the image of a point object in the presence of the polynomial aberration. As in the case of PSFs for circular pupils, more light has been introduced into some PSFs for annular pupils also to display their details.

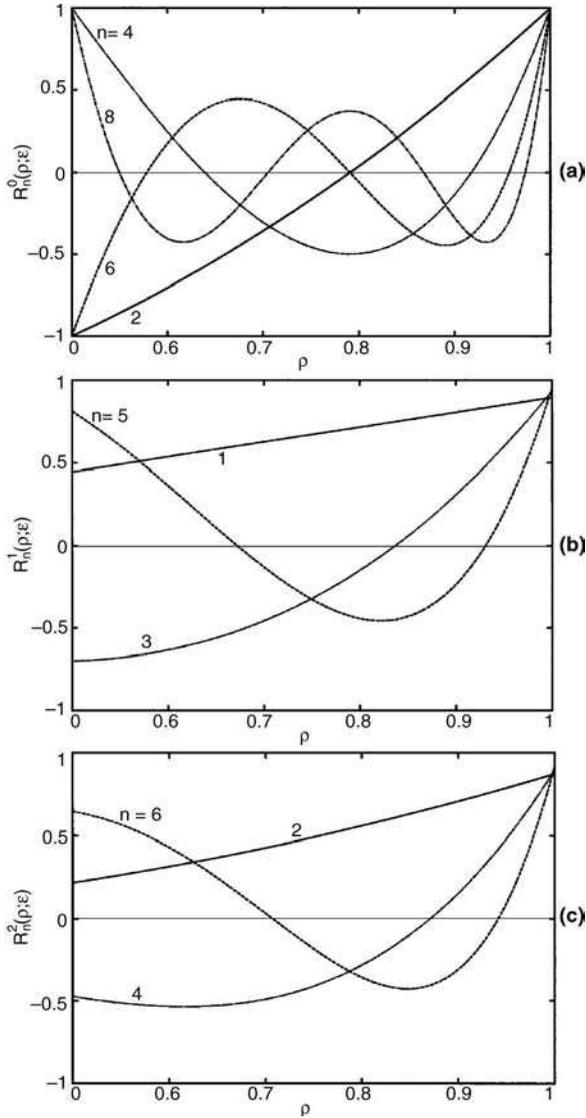


FIGURE 13.5. Variation of a radial annular polynomial $R_n^m(\rho; \epsilon)$ with ρ for $\epsilon = 0.5$. (a) Defocus and spherical aberrations. (b) Tilt and coma. (c) Astigmatism.

13.5. DETERMINATION OF ZERNIKE COEFFICIENTS FROM DISCRETE WAVEFRONT ERROR DATA

13.5.1. Introduction

We have seen that when the aberration function is known analytically, its aberration coefficients can be determined by calculating the integral in Eqs. (13.21), (13.51), or

TABLE 13.10. Orthonormal Zernike annular polynomials $Z_j(\rho, \theta; \epsilon)$. The indices j , n , and m are called the polynomial number, radial degree, and azimuthal frequency, respectively. The polynomials Z_j are ordered such that an even j corresponds to a symmetric polynomial varying as $\cos m\theta$, while an odd j corresponds to an antisymmetric polynomial varying as $\sin m\theta$. A polynomial with a lower value of n is ordered first, and for a given value of n , a polynomial with a lower value of m is ordered first.

j	n	m	$Z_j(\rho, \theta; \epsilon)$	Zernike aberration name
1	0	0	$R_0^0(\rho; \epsilon)$	Piston
2	1	1	$2R_1^1(\rho; \epsilon) \cos \theta$	x tilt
3	1	1	$2R_1^1(\rho; \epsilon) \sin \theta$	y tilt
4	2	0	$\sqrt{3}R_2^0(\rho; \epsilon)$	Defocus
5	2	2	$\sqrt{6}R_2^2(\rho; \epsilon) \sin 2\theta$	Primary astigmatism at 45°
6	2	2	$\sqrt{6}R_2^2(\rho; \epsilon) \cos 2\theta$	Primary astigmatism at 0°
7	3	1	$\sqrt{8}R_3^1(\rho; \epsilon) \sin \theta$	Primary y coma
8	3	1	$\sqrt{8}R_3^1(\rho; \epsilon) \cos \theta$	Primary x coma
9	3	3	$\sqrt{8}R_3^3(\rho; \epsilon) \sin 3\theta$	
10	3	3	$\sqrt{8}R_3^3(\rho; \epsilon) \cos 3\theta$	
11	4	0	$\sqrt{5}R_4^0(\rho; \epsilon)$	Primary spherical
12	4	2	$\sqrt{10}R_4^2(\rho; \epsilon) \cos 2\theta$	Secondary astigmatism at 0°
13	4	2	$\sqrt{10}R_4^2(\rho; \epsilon) \sin 2\theta$	Secondary astigmatism at 45°
14	4	4	$\sqrt{10}R_4^4(\rho; \epsilon) \cos 4\theta$	
15	4	4	$\sqrt{10}R_4^4(\rho; \epsilon) \sin 4\theta$	
16	5	1	$\sqrt{12}R_5^1(\rho; \epsilon) \cos \theta$	Secondary x coma
17	5	1	$\sqrt{12}R_5^1(\rho; \epsilon) \sin \theta$	Secondary y coma
18	5	3	$\sqrt{12}R_5^3(\rho; \epsilon) \cos 3\theta$	
19	5	3	$\sqrt{12}R_5^3(\rho; \epsilon) \sin 3\theta$	
20	5	5	$\sqrt{12}R_5^5(\rho; \epsilon) \cos 5\theta$	
21	5	5	$\sqrt{12}R_5^5(\rho; \epsilon) \sin 5\theta$	
22	6	0	$\sqrt{7}R_6^0(\rho; \epsilon)$	Secondary spherical
23	6	2	$\sqrt{14}R_6^2(\rho; \epsilon) \sin 2\theta$	Tertiary astigmatism at 45°
24	6	2	$\sqrt{14}R_6^2(\rho; \epsilon) \cos 2\theta$	Tertiary astigmatism at 0°
25	6	4	$\sqrt{14}R_6^4(\rho; \epsilon) \cos 4\theta$	
26	6	4	$\sqrt{14}R_6^4(\rho; \epsilon) \sin 4\theta$	
27	6	6	$\sqrt{14}R_6^6(\rho; \epsilon) \sin 6\theta$	
28	6	6	$\sqrt{14}R_6^6(\rho; \epsilon) \cos 6\theta$	

(13.82), depending on whether the system is rotationally symmetric or not, or if the pupil is circular or annular. Numerically efficient algorithms to evaluate these integrals can be written that overcome the computational problems associated with the oscillatory behavior of the Zernike polynomials (Wang and Silva, 1980; Prata and

TABLE 13.11. Orthonormal Zernike circle polynomials $Z_j(x, y, \epsilon)$ in Cartesian coordinates (x, y) , which are normalized by the pupil radius a . The subscript n on the coordinates indicating the normalization is omitted for simplicity. Note that $x = \rho \cos \theta$, $y = \rho \sin \theta$, and $\epsilon \leq \rho = \sqrt{x^2 + y^2} \leq 1$.

Polynomial	$Z_f(x, y; \epsilon)$
Z_1	1
Z_2	$2x/(1 + \epsilon^2)^{1/2}$
Z_3	$2y/(1 + \epsilon^2)^{1/2}$
Z_4	$\sqrt{3}(2\rho^2 - 1 - \epsilon^2)/(1 - \epsilon^2)$
Z_5	$2\sqrt{6}xy/(1 + \epsilon^2 + \epsilon^4)^{1/2}$
Z_6	$\sqrt{6}(x^2 - y^2)/(1 + \epsilon^2 + \epsilon^4)^{1/2}$
Z_7	$\frac{\sqrt{8}y[3(1 + \epsilon^2)\rho^2 - 2(1 + \epsilon^2 + \epsilon^4)]}{(1 - \epsilon^2)[1 + \epsilon^2](1 + 4\epsilon^2 + \epsilon^4)^{1/2}}$
Z_8	$\frac{\sqrt{8}x[3(1 + \epsilon^2)\rho^2 - 2(1 + \epsilon^2 + \epsilon^4)]}{(1 - \epsilon^2)[1 + \epsilon^2](1 + 4\epsilon^2 + \epsilon^4)^{1/2}}$
Z_9	$\sqrt{8}y(3x^2 - y^2)/(1 + \epsilon^2 + \epsilon^4 + \epsilon^6)^{1/2}$
Z_{10}	$\sqrt{8}x(x^2 - 3y^2)/(1 + \epsilon^2 + \epsilon^4 + \epsilon^6)^{1/2}$
Z_{11}	$\sqrt{5}[6\rho^4 - 6(1 + \epsilon^2)\rho^2 + (1 + 4\epsilon^2 + \epsilon^4)]/(1 - \epsilon^2)^2$
Z_{12}	$\frac{\sqrt{10}(x^2 - y^2)[4\rho^2 - 3(1 - \epsilon^8)/(1 - \epsilon^6)]}{\{(1 - \epsilon^2)^{-1}[16(1 - \epsilon^{10}) - 15(1 - \epsilon^8)^2/(1 - \epsilon^6)]\}^{1/2}}$
Z_{13}	$\frac{2\sqrt{10}xy[4\rho^2 - 3(1 - \epsilon^8)/(1 - \epsilon^6)]}{\{(1 - \epsilon^2)^{-1}[16(1 - \epsilon^{10}) - 15(1 - \epsilon^8)^2/(1 - \epsilon^6)]\}^{1/2}}$
Z_{14}	$\sqrt{10}(\rho^4 - 8x^2y^2)/(1 + \epsilon^2 + \epsilon^4 + \epsilon^6 + \epsilon^8)^{1/2}$
Z_{15}	$4\sqrt{10}xy(x^2 - y^2)/(1 + \epsilon^2 + \epsilon^4 + \epsilon^6 + \epsilon^8)^{1/2}$
Z_{16}	$\frac{\sqrt{12}x[10(1 + 4\epsilon^2 + \epsilon^4)\rho^4 - 12(1 + 4\epsilon^2 + 4\epsilon^4 + \epsilon^6)\rho^2 + 3(1 + 4\epsilon^2 + 10\epsilon^4 + 4\epsilon^6 + \epsilon^8)]}{(1 - \epsilon^2)^2[(1 + 4\epsilon^2 + \epsilon^4)(1 + 9\epsilon^2 + 9\epsilon^4 + \epsilon^6)]^{1/2}}$
Z_{17}	$\frac{\sqrt{12}y[10(1 + 4\epsilon^2 + \epsilon^4)\rho^4 - 12(1 + 4\epsilon^2 + 4\epsilon^4 + \epsilon^6)\rho^2 + 3(1 + 4\epsilon^2 + 10\epsilon^4 + 4\epsilon^6 + \epsilon^8)]}{(1 - \epsilon^2)^2[(1 + 4\epsilon^2 + \epsilon^4)(1 + 9\epsilon^2 + 9\epsilon^4 + \epsilon^6)]^{1/2}}$
Z_{18}	$\frac{\sqrt{12}x(x^2 - 3y^2)[5\rho^2 - 4(1 - \epsilon^{10})/(1 - \epsilon^8)]}{\{(1 - \epsilon^2)^{-1}[25(1 - \epsilon^{12}) - 24(1 - \epsilon^{10})^2/(1 - \epsilon^8)]\}^{1/2}}$
Z_{19}	$\frac{\sqrt{12}y(3x^2 - y^2)[5\rho^2 - 4(1 - \epsilon^{10})/(1 - \epsilon^8)]}{\{(1 - \epsilon^2)^{-1}[25(1 - \epsilon^{12}) - 24(1 - \epsilon^{10})^2/(1 - \epsilon^8)]\}^{1/2}}$

TABLE 13.11. (Continued)

Polynomial	$Z_f(x, y; \epsilon)$
Z_{20}	$\sqrt{12}x(16x^4 - 20x^2\rho^2 + 5\rho^4)/[1 + \epsilon^2 + \epsilon^4 + \epsilon^6 + \epsilon^8 + \epsilon^{10}]^{1/2}$
Z_{21}	$\sqrt{12}y(16y^4 - 20y^2\rho^2 + 5\rho^4)/[1 + \epsilon^2 + \epsilon^4 + \epsilon^6 + \epsilon^8 + \epsilon^{10}]^{1/2}$
Z_{22}	$\sqrt{7}[20\rho^6 - 30(1 + \epsilon^2)\rho^4 + 12(1 + 3\epsilon^2 + \epsilon^4)\rho^2 - (1 + 9\epsilon^2 + 9\epsilon^4 + \epsilon^6)]/(1 - \epsilon^2)^3$
Z_{23}	$\frac{2\sqrt{14}xy[15(1 + 4\epsilon^2 + 10\epsilon^4 + 4\epsilon^6 + \epsilon^8)\rho^4 - 20(1 + 4\epsilon^2 + 10\epsilon^4 + 10\epsilon^6 + 4\epsilon^8 + \epsilon^{10})\rho^2 + 6(1 + 4\epsilon^2 + 10\epsilon^4 + 20\epsilon^6 + 10\epsilon^8 + 4\epsilon^{10} + \epsilon^{12})]}{(1 - \epsilon^2)^2[1 + 4\epsilon^2 + 10\epsilon^4 + 4\epsilon^6 + \epsilon^8](1 + 9\epsilon^2 + 45\epsilon^4 + 65\epsilon^6 + 45\epsilon^8 + 9\epsilon^{10} + \epsilon^{12})}^{1/2}$
Z_{24}	$\frac{\sqrt{14}(x^2 - y^2)[15(1 + 4\epsilon^2 + 10\epsilon^4 + 4\epsilon^6 + \epsilon^8)\rho^4 - 20(1 + 4\epsilon^2 + 10\epsilon^4 + 10\epsilon^6 + 4\epsilon^8 + \epsilon^{10})\rho^2 + 6(1 + 4\epsilon^2 + 10\epsilon^4 + 20\epsilon^6 + 10\epsilon^8 + 4\epsilon^{10} + \epsilon^{12})]}{(1 - \epsilon^2)^2[1 + 4\epsilon^2 + 10\epsilon^4 + 4\epsilon^6 + \epsilon^8](1 + 9\epsilon^2 + 45\epsilon^4 + 65\epsilon^6 + 45\epsilon^8 + 9\epsilon^{10} + \epsilon^{12})}^{1/2}$
Z_{25}	$\frac{4\sqrt{14}xy(x^2 - y^2)[6\rho^2 - 5(1 - \epsilon^{12})/(1 - \epsilon^{10})]}{\{(1 - \epsilon^2)^{-1}[36(1 - \epsilon^{14}) - 35(1 - \epsilon^{12})^2/(1 - \epsilon^{10})]\}^{1/2}}$
Z_{26}	$\frac{\sqrt{14}(8x^4 - 8x^2\rho^2 + \rho^4)[6\rho^2 - 5(1 - \epsilon^{12})/(1 - \epsilon^{10})]}{\{(1 - \epsilon^2)^{-1}[36(1 - \epsilon^{14}) - 35(1 - \epsilon^{12})^2/(1 - \epsilon^{10})]\}^{1/2}}$
Z_{27}	$\sqrt{14}xy(32x^4 - 32x^2\rho^2 + 6\rho^4)/(1 + \epsilon^2 + \epsilon^4 + \epsilon^6 + \epsilon^8 + \epsilon^{10} + \epsilon^{12})^{1/2}$
Z_{28}	$\sqrt{14}(32x^6 - 48x^4\rho^2 + 18x^2\rho^4 - \rho^6)/(1 + \epsilon^2 + \epsilon^4 + \epsilon^6 + \epsilon^8 + \epsilon^{10} + \epsilon^{12})^{1/2}$

Rusch, 1989). The variance of the aberration is given by the sum of the squares of the aberration coefficients, with the exception of a_1 which represents the mean value of the aberration. Although, in principle, the number of polynomials representing the aberration function is infinite, in practice, the number will be finite. This number is determined by comparing the variance obtained by summing the squares of the finite number of coefficients with the actual variance. The number of polynomials is increased until the difference between the two variances is below a prechosen value, which depends on the tolerance to which the aberration is being determined.

When the aberration function is known only at a discrete set of points, as in a digitized interferogram, the integral for determining the aberration coefficients reduces to a sum and the coefficients thus obtained may be in error, since the Zernike polynomials $Z_i(x, y)$, which are orthogonal over the full circular region, are not orthogonal over the discrete points of the aberration data set. The magnitude of the error decreases as the number of points increases. This is not a serious problem when the wavefront errors are determined by, say, phase-shifting interferometry, since the number of points can be very large. However, when the pupil is neither circular nor annular, then the corresponding Zernike polynomials are not orthogonal over the data set even if the number of points is large. Such cases are real and quite common. For example, a system with a circular pupil for an axial point object yields an approximately elliptical pupil for an off-axis point object. The pupil may even be

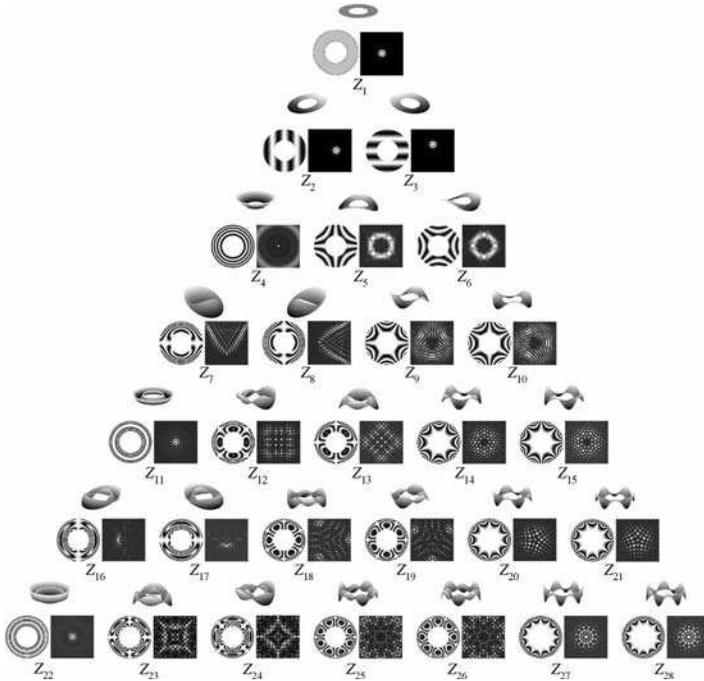


FIGURE 13.6. Zernike annular polynomials pyramid showing the isometric plot on the top, interferogram on the left, and PSF on the right for each polynomial. The value of ϵ is 0.5, and the standard deviation of each polynomial aberration is one wave.

irregular in shape due to vignetting. Hence, ray tracing or testing of the system yields wavefront error data at an array of points across a region that may or may not be circular or annular. Similarly, in the case of a segmented mirror, the segment under test may be hexagonal in shape, as in the Keck telescope. When we test such systems, we measure wavefront error data at an array of points, presumed regularly spaced, either in reality or by interpolation. For interpolation schemes, see, for example, Korn and Korn (1968). But invariably we want to know how much of the aberration is a focus error, spherical aberration, or astigmatism, etc. In such cases, we first determine the coefficients of an expansion in terms of polynomials that are orthogonal over the data set and obtained by the Gram–Schmidt orthogonalization process, and then determine the Zernike coefficients of the aberration function regardless of the shape of the pupil (Wang and Silva, 1980; Malacara et al., 1990; Swantner and Chow, 1994; Upton and Ellerbroek, 2004; Mahajan and Dai, 2006). This process is outlined with a numerical example where a hexagonal pupil is considered (Mahajan and Dai, 2006; Dai and Mahajan, 2007). Since analytical results for the orthonormal polynomials for such a pupil are available, it is easy to compare the numerical results with the corresponding analytical ones and thereby assess the validity of a software program.

It should be noted that the wavefront must be relatively smooth in order to consider its Zernike aberration coefficients. It does not make sense, for example, to consider the

Zernike coefficients of a localized region of a pupil, such as the one resulting from a narrow bump of a mirror. In principle, Zernike coefficients in such cases can also be determined, but it may take a large number of polynomials to represent such wavefronts. The physical significance of these coefficients may not be of much use, however. The use of Gaussian functions has been suggested for such cases (Montoya-Hernández et. al., 1999). It should be evident that in the case of a segmented mirror, the segments must be phased (Chanan and Ohara, 1998; Chanan and Troy, 1998), the subwavefronts of the segments stitched to determine the overall wavefront of the mirror (Thunen and Kwon, 1982; Jenson, Chow, and Lawrence, 1984; Fleig, Dumas, Murphy, and Forbes, 2003) and then determine the Zernike coefficients of the mirror.

13.5.2. Orthonormal Coefficients and Aberration Variance

Let N be the number of points at which the wavefront error $W_n \equiv W(x_n, y_n)$, $n = 1, 2, 3, \dots, N$, is known by way of an interferogram. The variance of the aberration is given by

$$\begin{aligned}\sigma_N^2 &= \langle W^2 \rangle - \langle W \rangle^2 \\ &= \frac{1}{N} \sum_{n=1}^N W_n^2 - \left(\frac{1}{N} \sum_{n=1}^N W_n \right)^2\end{aligned}\quad (13.84)$$

Let the aberration data be approximated by an aberration function $\hat{W}(x_n, y_n)$ in terms of L polynomials $F_l(x_n, y_n)$, $l = 1, 2, 3, \dots, L$, in the form

$$\hat{W}(x_n, y_n) = \sum_{l=1}^L a_l F_l(x_n, y_n) \quad (13.85)$$

where a_l are the expansion or the aberration coefficients. The mean square difference between the actual and the estimated wavefront errors is given by

$$\begin{aligned}E &= \frac{1}{N} \sum_{n=1}^N (W_n - \hat{W}_n)^2 \\ &= \frac{1}{N} \sum_{n=1}^N \left(W_n - \sum_{l=1}^L a_l F_{ln} \right)^2\end{aligned}\quad (13.86)$$

The mean square difference is minimum if the coefficients a_l are such that

$$0 = \frac{\partial E}{\partial a_l} = \frac{2}{N} \sum_{n=1}^N \left(\sum_{k=1}^L a_k F_{kn} - W_n \right) F_{ln}$$

or

$$\sum_{k=1}^L a_k \sum_{n=1}^N F_{kn} F_{ln} = \sum_{n=1}^N W_n F_{ln} \quad (13.87)$$

If the polynomials F_l are orthonormal over the data set, that is, if

$$\frac{1}{N} \sum_{n=1}^N F_{kn} F_{ln} = \delta_{kl} \quad (13.88)$$

then Eq. (13.87) yields the coefficients

$$a_l = \frac{1}{N} \sum_{n=1}^N W_n F_{ln} \quad (13.89)$$

Note that the value of an orthonormal coefficient a_l is independent of the number L of terms used to estimate the aberration data.

The aberration variance in terms of the aberration coefficients can be obtained from the mean $\langle \hat{W} \rangle$ and the mean square $\langle \hat{W}^2 \rangle$ values as follows:

$$\begin{aligned} \langle \hat{W} \rangle &= \frac{1}{N} \sum_{n=1}^N \sum_{l=1}^L a_l F_{ln} \\ &= \sum_{l=1}^L a_l \frac{1}{N} \sum_{n=1}^N F_{ln} \\ &= a_1 \end{aligned} \quad (13.90)$$

$$\begin{aligned} \langle \hat{W}^2 \rangle &= \frac{1}{N} \sum_{n=1}^N \sum_{k=1}^L a_k F_{kn} \sum_{l=1}^L a_l F_{ln} \\ &= \sum_{k=1}^L a_k \sum_{l=1}^L a_l \frac{1}{N} \sum_{n=1}^N F_{kn} F_{ln} \\ &= \sum_{l=1}^L a_l^2 \end{aligned} \quad (13.91)$$

and

$$\hat{\sigma}^2 = \langle \hat{W}^2 \rangle - \langle \hat{W} \rangle^2 = \sum_{l=2}^L a_l^2 \quad (13.92)$$

In obtaining Eq. (13.90), we have used the orthonormality Eq. (13.88) subject to the usual condition $F_1 = 1$. Thus, the variance of the estimated aberration is simply the sum of the squares of the aberration coefficients a_l , with the exception of a_1 which represents the mean value of the aberration. The coefficient of each term of Eq. (13.85), except $l = 1$, represents its standard deviation. The addition or subtraction of one or more orthonormal terms does not affect the value of the coefficients a_l . The number of polynomials L is again dictated by the tolerance on the difference between the estimated variance $\hat{\sigma}^2$ and its actual value σ_N^2 . For example, we may want that $\hat{\sigma}$ be equal to σ_N within $\lambda/30$.

13.5.3. Orthonormal Polynomials

The orthonormal polynomials $F_l(x_n, y_n)$ may be obtained from the Zernike polynomials $Z_f(x_n, y_n)$ by the Gram–Schmidt orthogonalization process (Korn and Korn, 1968):

$$G_1 = Z_1 = 1 \tag{13.93a}$$

$$G_{l+1} = \sum_{k=1}^l c_{l+1,k} F_k + Z_{l+1}, \quad l = 1, 2, \dots, L - 1 \tag{13.93b}$$

$$F_{l+1} = \frac{G_{l+1}}{\|G_{l+1}\|} = \frac{G_{l+1}}{\left(\frac{1}{N} \sum_{n=1}^N G_{l+1,n}^2\right)^{1/2}} \tag{13.93c}$$

where

$$c_{l+1,k} = -\frac{1}{N} \sum_{n=1}^N Z_{l+1,n} F_{kn} \tag{13.93d}$$

Thus, the F-polynomials can be obtained recursively as follows:

$$G_1 = Z_1 = 1 \tag{13.94a}$$

$$G_2 = c_{21} F_1 + Z_2 = c_{21} Z_1 + Z_2 \tag{13.94b}$$

$$\begin{aligned} G_3 &= c_{31} F_1 + c_{32} F_2 + Z_3 \\ &= \left(c_{31} + \frac{c_{32} c_{21}}{\|G_2\|}\right) Z_1 + \frac{c_{32}}{\|G_2\|} Z_2 + Z_3 \end{aligned} \tag{13.94c}$$

$$\begin{aligned} G_4 &= c_{41} F_1 + c_{42} F_2 + c_{43} F_3 + Z_4 \\ &= \left[c_{41} + \frac{c_{42} c_{21}}{\|G_2\|} + \frac{c_{43}}{\|G_3\|} \left(c_{31} + \frac{c_{32} c_{21}}{\|G_2\|} \right) \right] Z_1 \\ &\quad + \frac{1}{\|G_2\|} \left(c_{42} + \frac{c_{43} c_{32}}{\|G_3\|} \right) Z_2 + \frac{c_{43}}{\|G_3\|} Z_3 + Z_4 \end{aligned} \tag{13.94d}$$

etc. A general expression for a G-polynomial may be written in the form

$$G_l = \sum_{i=1}^{l-1} \sum_{j=1}^{l-i} c_{l,l-j} M_{l-j,i} Z_i + Z_l \tag{13.95a}$$

where

$$M_{li} = \frac{1}{\|G_l\|} \sum_{j=1}^{l-i} c_{l,l-j} M_{l-j,i} \tag{13.95b}$$

Each G- and, therefore, F-polynomial is thus a linear combination of the Zernike polynomials. The orthonormal F-polynomials represent the unit vectors of an

L -dimensional space for expressing the aberration data. They can be written in a matrix form according to

$$F_l(x_n, y_n) = \sum_{i=1}^l M_{li} Z_i(x_n, y_n) \quad \text{with} \quad M_{ll} = \frac{1}{\|G_l\|} \quad (13.96)$$

While the diagonal elements of the M -matrix are simply equal to the normalization constants of the G -polynomials, there are no matrix elements above the diagonal. The matrix is triangular and the missing elements may be given a value of zero. The orthonormal polynomials can be obtained nonrecursively with a matrix approach (Dai and Mahajan, 2007).

13.5.4. Zernike Coefficients

Substituting Eq. (13.96) into Eq. (13.85), we obtain

$$\begin{aligned} \hat{W}(x_n, y_n) &= \sum_{l=1}^L a_l \sum_{i=1}^l M_{li} Z_i(x_n, y_n) \\ &= \sum_{l=1}^L \sum_{i=l}^L a_i M_{il} Z_l(x_n, y_n) \end{aligned} \quad (13.97)$$

Writing the aberration function in terms of Zernike polynomials in the form

$$\hat{W}(x_n, y_n) = \sum_{l=1}^L b_l Z_l(x_n, y_n) \quad (13.98)$$

and comparing with Eq. (13.97) yields the Zernike coefficients

$$b_l = \sum_{i=l}^L a_i M_{il} \quad (13.99)$$

Since the Zernike polynomials are not orthogonal over the data set, the aberration coefficients b_l are not independent of each other, the mean value of a polynomial is not zero, b_1 is not equal to the mean value a_1 of $\hat{W}(x_n, y_n)$, the variance of the aberration cannot be obtained by summing the squares of the aberration coefficients b_l with $l = 2, 3, \dots, L$, and the value of a coefficient changes as the number L of the aberration terms changes. Hence, when one or more aberration terms are added or subtracted from the aberration function, the coefficients b_l change.

In practice, the F -polynomials are not known. What we know are the wavefront errors at an array of points within a pupil of arbitrary shape, either by ray tracing, as in optical design, or by measurement, as in optical testing, and we want to determine the Zernike coefficients that match these errors in a least squares sense. A summary of the procedure for determining these coefficients is as follows.

1. Using the Zernike circle or annular polynomials, depending on the whether the shape of the actual pupil is close to being circular or annular, determine the

F-polynomials and the M-matrix recursively using Eqs. (13.93)–(13.95). In this step, only the location of the data points is needed. The F-polynomials are orthogonal over the data set, and they change as the data set changes.

2. The aberration coefficients a_l are determined from Eq. (13.89) by substituting the F-polynomials and the wavefront error data.
3. Finally, the Zernike coefficients b_l are obtained from Eq. (13.99).

The number L of the F- (or G-) and, therefore, the Zernike polynomials is increased until the least square error is below a prechosen value. It should be evident that, as the pupil approaches a circular (or an annular) shape and the number of data points is sufficiently large, the F-polynomials approach the orthonormal Zernike circle (or annular) polynomials, the matrix M approaches a unity matrix (i.e., one with unity diagonal elements and zero others), and the expansion coefficients $a_l \rightarrow b_l$.

13.5.5 Numerical Example

As a numerical example, we consider a hexagonal pupil since analytical expressions are available for the orthonormal polynomials and the M-matrix (Mahajan and Dai, 2006; Dai and Mahajan, 2007). Hence, numerical results can be compared with the corresponding analytical ones. First we consider the analytical results. The first eleven polynomials that are orthonormal over the hexagonal pupil are given in Table 13.12. They can be obtained from the Zernike circle polynomials by multiplying them with the matrix M given in Table 13.13 according to Eq. (13.96).

TABLE 13.12. Orthonormal hexagonal polynomials in polar and Cartesian coordinates, where $(x, y) = \rho(\cos \theta, \sin \theta)$ and (x, y) and ρ lie inside the hexagon.

j	$H_j(\rho, \theta) = H_j(x, y)$
1	1
2	$2\sqrt{6/5}\rho \cos \theta = 2\sqrt{6/5}x$
3	$2\sqrt{6/5}\rho \sin \theta = 2\sqrt{6/5}y$
4	$\sqrt{5/43}(12\rho^2 - 5)$
5	$2\sqrt{15/7}\rho^2 \sin 2\theta = 4\sqrt{15/7}xy$
6	$2\sqrt{15/7}\rho^2 \cos 2\theta = 2\sqrt{15/7}(x^2 - y^2)$
7	$24\sqrt{42/3685}(25\rho^3 - 14\rho) \sin \theta = 24\sqrt{42/3685}y(25\rho^2 - 14)$
8	$24\sqrt{42/3685}(25\rho^3 - 14\rho) \cos \theta = 24\sqrt{42/3685}x(25\rho^2 - 14)$
9	$(4/3)\sqrt{10}\rho^3 \sin 3\theta = (4/3)\sqrt{10}y(3x^2 - y^2)$
10	$4\sqrt{70/103}\rho^3 \cos 3\theta = 4\sqrt{70/103}x(x^2 - 3y^2)$
11	$(3/\sqrt{1072205})(6020\rho^4 - 5140\rho^2 + 737)$

TABLE 13.13. Analytical matrix M for obtaining the Zernike coefficients b_l from the orthonormal hexagonal coefficients a_l .

1	0	0	0	0	0	0	0	0	0	0
0	$\sqrt{6/5}$	0	0	0	0	0	0	0	0	0
0	0	$\sqrt{6/5}$	0	0	0	0	0	0	0	0
$\sqrt{5/43}$	0	0	$2\sqrt{15/43}$	0	0	0	0	0	0	0
0	0	0	0	$\sqrt{10/7}$	0	0	0	0	0	0
0	0	0	0	0	$\sqrt{10/7}$	0	0	0	0	0
0	0	$16\sqrt{\frac{14}{11055}}$	0	0	0	$10\sqrt{\frac{35}{2211}}$	0	0	0	0
0	$16\sqrt{\frac{14}{11055}}$	0	0	0	0	0	$10\sqrt{\frac{35}{2211}}$	0	0	0
0	0	0	0	0	0	0	0	$\frac{2}{3}\sqrt{5}$	0	0
0	0	0	0	0	0	0	0	0	$2\sqrt{\frac{35}{103}}$	0
$\frac{521}{\sqrt{1072205}}$	0	0	$88\sqrt{\frac{15}{214441}}$	0	0	0	0	0	0	$14\sqrt{\frac{43}{4987}}$

TABLE 13.14. Analytical inverse matrix M^{-1} for obtaining the orthonormal hexagonal coefficients a_l from the Zernike coefficients b_l .

1	0	0	0	0	0	0	0	0	0	0
0	$\sqrt{5/6}$	0	0	0	0	0	0	0	0	0
0	0	$\sqrt{5/6}$	0	0	0	0	0	0	0	0
$-1/2\sqrt{3}$	0	0	$\sqrt{43/15}/2$	0	0	0	0	0	0	0
0	0	0	0	$\sqrt{7/10}$	0	0	0	0	0	0
0	0	0	0	0	$\sqrt{7/10}$	0	0	0	0	0
0	0	$-8/5\sqrt{15}$	0	0	0	$\sqrt{\frac{2211}{35}}/10$	0	0	0	0
0	$-8/5\sqrt{15}$	0	0	0	0	0	$\sqrt{\frac{2211}{35}}/10$	0	0	0
0	0	0	0	0	0	0	0	$3/2\sqrt{5}$	0	0
0	0	0	0	0	0	0	0	0	$\sqrt{\frac{103}{35}}/2$	0
$-1/2\sqrt{5}$	0	0	$-22/7\sqrt{43}$	0	0	0	0	0	0	$\sqrt{\frac{4987}{43}}/14$

Since only 16 of the 121 matrix elements are zero, we obtain simple expressions for the Zernike coefficients b_l in terms of the orthonormal coefficients a_l :

$$\begin{aligned} b_1 &= a_1 M_{11} + a_4 M_{41} + a_{11} M_{11,1} \\ &= a_1 + \sqrt{5/53} a_4 + (521/\sqrt{1072205}) a_{11} \end{aligned} \quad (13.100a)$$

$$\begin{aligned} b_2 &= a_2 M_{22} + a_8 M_{82} \\ &= \sqrt{6/5} a_2 + 16\sqrt{14/11055} a_8 \end{aligned} \quad (13.100b)$$

$$\begin{aligned} b_3 &= a_3 M_{33} + a_7 M_{73} \\ &= \sqrt{6/5} a_3 + 16\sqrt{14/11055} a_7 \end{aligned} \quad (13.100c)$$

$$\begin{aligned} b_4 &= a_4 M_{44} + a_{11} M_{11,4} \\ &= 2\sqrt{15/43} a_4 + 88\sqrt{15/214441} a_{11} \end{aligned} \quad (13.100d)$$

and

$$b_l = a_l M_{ll}, \quad l = 5, 6, \dots, 11 \quad (13.100e)$$

Thus, for example, $b_{11} = a_{11} M_{11,11} = 14\sqrt{43/4987} a_{11} = 1.3a_{11}$.

The orthonormal coefficients a_l can be obtained from Eq. (13.99) by inverting it. Thus,

$$a_l = \sum_{i=1}^L b_i M_{il}^{-1} \quad (13.101)$$

where M_{il}^{-1} is the inverse matrix given in Table 13.14. Hence, the orthonormal coefficients are given by

$$\begin{aligned} a_1 &= b_1 M_{11} + b_4 M_{41}^{-1} + b_{11} M_{11,1}^{-1} \\ &= b_1 - (1/2\sqrt{3}) b_4 - (1/2\sqrt{5}) b_{11} \end{aligned} \quad (13.102a)$$

$$\begin{aligned} a_2 &= b_2 M_{22}^{-1} + b_8 M_{82}^{-1} \\ &= \sqrt{5/6} b_2 - (8/5\sqrt{15}) b_8 \end{aligned} \quad (13.102b)$$

$$\begin{aligned} a_3 &= b_3 M_{33}^{-1} + b_7 M_{73}^{-1} \\ &= \sqrt{5/6} b_3 - (8/5\sqrt{15}) b_7 \end{aligned} \quad (13.102c)$$

$$\begin{aligned} a_4 &= b_4 M_{44}^{-1} + b_{11} M_{11,4}^{-1} \\ &= (1/2)\sqrt{43/15} b_4 - (22/7\sqrt{43}) b_{11} \end{aligned} \quad (13.102d)$$

and

$$a_l = b_l M_{ll}^{-1}, \quad l = 5, 6, \dots, 11 \quad (13.102e)$$

Equations (100 e) and (102 e) hold when the aberration function is expanded in terms of the first 11 polynomials. However, if say, the first 15 polynomials are used in the expansion, then b_5 , for example, also depends on a_{13} and a_{15} (Dai and Mahajan, 2007).

It is evident from Eq. (13.51) that the orthonormal hexagonal coefficients a_l are independent of the number L of the polynomials used in the expansion of an aberration function. Moreover, a_1 represents the mean value and the sum of the squares of the coefficients a_l with $l > 1$ represents the variance of the aberration. This is not true of the Zernike coefficients b_l , since the Zernike circle polynomials are not orthogonal over the hexagonal pupil. Thus, the values of the coefficients change as the number of polynomials changes, the mean value of a Zernike polynomial is not zero, b_1 does not represent the mean value of the aberration, and the variance of the aberration is not equal to the sum of the squares of the coefficients b_l with $l > 1$. Hence, these coefficients are not as useful as the orthonormal coefficients a_l .

Now we outline the procedure for simulating a numerical example. Consider an interferogram represented by an aberration function consisting of some aberration terms with coefficients b_l . We calculate the corresponding orthonormal coefficients a_l from Eqs. (13.102). The mean value of the aberration is given by a_1 and the variance by

$$\sigma_\infty^2 = \sum_{l=2} a_l^2 \quad (13.103)$$

where the subscript ∞ indicates an infinite number of points since the variance has been determined analytically. In practice, the calculation process is of course the reverse of the above. We are given the aberration data at N points, we determine the polynomials F_l that are orthonormal over the data, determine the estimated matrix \hat{M} and the orthonormal coefficients \hat{a}_l , and then determine the Zernike coefficients b_l .

Given the simulated aberration function, we calculate the aberration values W_n at an array of N points distributed uniformly within a unit hexagon. The standard deviation σ_N of the aberration is calculated using Eq. (13.84). This value approaches σ_∞ as N increases. The polynomials F_l that are orthonormal over the chosen array of points are determined by use of Eqs. (13.93)–(13.94). Suppose we approximate the aberration data by 11 F-polynomials, then $L = 11$, and we can determine the value of each polynomial at the array of points, and compare with those obtained from the analytical expressions for the F-polynomials given in Table 13.12.

The quantities that are calculated in the process of determining the orthonormal F-polynomials also yield the elements of the M-matrix. For example, since $F_l = G_l / \|G_l\|$, from Eq. (13.94c)

$$M_{31} = \frac{1}{\|G_3\|} \left(c_{31} + \frac{c_{32}c_{21}}{\|G_2\|} \right), \quad M_{32} = \frac{c_{32}}{\|G_3\| \|G_2\|}, \quad M_{33} = \frac{1}{\|G_3\|} \quad (13.104)$$

Of course, $M_{3l} = 0$ for $l = 4, 5, \dots, 11$. Because of the hexagonal symmetry of the pupil, c_{21} and c_{31} are also equal to zero (Mahajan and Dai, 2006). Hence, even M_{31} and M_{32} are equal to zero. The full \hat{M} -matrix thus obtained can be compared with the corresponding analytical matrix given in Table 13.13.

Substituting the matrix elements and the orthogonal coefficients into Eq. (13.99), we obtain the eleven estimated Zernike coefficients \hat{b}_l and compare them with the coefficients b_l used in simulating the interferogram. Substituting the aberration values W_n and a polynomial F_l into Eq. (13.89), we calculate the estimated orthonormal coefficient \hat{a}_l . The 11 coefficients thus obtained are substituted into Eq. (13.92) to obtain the estimated standard deviation $\hat{\sigma}_N$ of the aberration and compare it with the actual value σ_N .

If, for better estimation, the number of the data points is increased without changing the number of polynomials L , then the values of the matrix elements change somewhat. They become more nearly the same as those obtained analytically. Accordingly, the calculated orthonormal polynomials match more closely their analytical counterparts, and the precision of the orthonormal and Zernike coefficients improves. If the number of polynomials is increased for a better fit of the data, say, from L_1 to L_2 , then the matrix elements of L_1 are a subset of those of L_2 , that is, the matrix elements for a given array of points do not change as new ones are added when L is increased.

13.6. SUMMARY

Starting with a power-series expansion, the primary or Seidel, and secondary or Schwarzschild aberrations are described. The Zernike circle and annular polynomials representing balanced aberrations for minimum variance that are orthogonal over a circular or an annular pupil, respectively, are discussed. These polynomials are written in an orthonormal form, both in polar as well as in Cartesian coordinates, so that when an aberration function is expanded in terms of them, each expansion coefficient represents the standard deviation of the corresponding aberration term evaluated across the pupil. Accordingly, the variance of the aberration function is given by the sum of the squares of the aberration coefficients, with the exception of the piston coefficient. The number of terms required to adequately represent an aberration function is increased until the variance thus obtained equals the actual variance within a prechosen tolerance. The value of an aberration coefficient, however, as determined by the integral in Eq. (13.21), (13.51), or (13.82), is independent of the number of terms. Similarly, the aberration coefficients are independent of each other. Equations for the conversion of Seidel aberrations into Zernike aberrations and vice versa are given. It is emphasized that the Seidel coefficients can be obtained from the primary Zernike aberrations only if the higher-order Zernike terms are negligible.

When the aberration function is known only at a discrete array of points, which is often the case in practice, for example, by ray tracing, as in optical design, or by measurement, as in optical testing, then the integral reduces to a sum. The sum approaches the integral as the number of points increases. However, the Zernike

polynomials are not orthogonal over the data set, and, therefore, the coefficients are not independent of each other or the number of terms used in the expansion. Similarly, Zernike polynomials are not orthogonal when the pupil is neither circular nor annular, regardless of the number of points. In such cases, a procedure is described for determining the Zernike coefficients that match the data values in a least squares sense. In this procedure, we first determine polynomials that are orthogonal over the data set by the Gram–Schmidt orthogonalization process. The Zernike coefficients of the expansion are then determined from the coefficients of these orthogonal polynomials. These coefficients are not independent of each other and their values change as the number of aberration terms changes. Moreover, the variance of the aberration is not equal to the sum of the squares of the aberration coefficients. A numerical example of a hexagonal pupil is outlined to illustrate these concepts. Since analytical results are available for such a pupil, it is expected that it will help verify the software program written by a user.

ACKNOWLEDGMENT

The author gratefully acknowledges helpful discussions with Drs. G. Dai, W. H. Swantner, P. Mouroulis, and Professor J. C. Wyant. The Zernike pyramids shown in Figures 13.4 and 13.6 were prepared by Dr. Dai.

REFERENCES

- Bhatia A. B. and E. Wolf, “On the Circle Polynomials of Zernike and Related Orthogonal Sets,” *Proc. Cambridge Philos. Soc.* **50**, 40–48 (1954).
- Born M. and E. Wolf, *Principles of Optics*, 7th ed., Oxford, New York, 1999.
- Braat, J. “Polynomial Expansion of Severely Aberrated Wave Fronts,” *J. Opt. Soc. Am.* **A4**, 643–650 (1987).
- Chanan, G. M. Troy, F. Dekenés, S. Michaels, J. Nelson, T. Mast, and D. Kirkman, “Phasing the Mirror Segments of the Keck Telescopes: the Broadband Phasing Algorithm,” *Appl. Opt.* **37**, 140–155 (1998).
- Chanan, G. C. Ohara, M. Troy, F. Dekenés, S. Michaels, J. Nelson, T. Mast, and D. Kirkman, “Phasing the Mirror Segments of the Keck Telescopes II: the Narrow-band Phasing Algorithm,” *Appl. Opt.*, **39**, 4706–4714 (1998).
- Conforti, G. “Zernike Aberration Coefficients From Seidel and Higher-order Power-series Coefficients,” *Opt. Lett.*, **8**, 407–408 (1983).
- Dai G. -m. and V. N. Mahajan, “Nonrecursive Determination of Orthonormal Polynomials with Matrix Formulation,” *Opt. Lett.* **32**, 74–76 (2007).
- Fleig, J. P. Dumas, P. Murphy, and G. W. Forbes, “An Automated Subaperture Stitching Interferometer Workstation for Spherical and Aspherical Surfaces,” *SPIE Proc.*, **5188**, 296–307 (2003).
- Jenson, S. C. W. W. Chow, and G. N. Lawrence, “Subaperture Testing Approaches: a Comparison,” *Appl. Opt.*, **23**, 740–745 (1984).

- Korn G. A. and T. M. Korn, *Mathematical Handbook for Scientists and Engineers*, McGraw-Hill, New York, 1968.
- Mahajan, V. N. "Strehl Ratio for Primary Aberrations: Some Analytical Results for Circular and Annular Pupils," *J. Opt. Soc. Am.* **72**, 1258–1266 (1982), Errata, **10**, 2092 (1993); "Strehl Ratio for Primary Aberrations in Terms of Their Aberration Variance," *J. Opt. Soc. Am.* **73**, 860–861 (1983).
- Mahajan, V. N. "Zernike Annular Polynomials for Imaging Systems with Annular Pupils," *J. Opt. Soc. Am.* **71**, 75–85 (1981a); **71**, 1408 (1981b); **1**, 685 (1984); "Zernike Annular Polynomials and Optical Aberrations of Systems With Annular Pupils," *Appl. Opt.* **33**, 8125–8127 (1994).
- Mahajan, V. N. *Optical Imaging and Aberrations*, Part I: *Ray Geometrical Optics*, SPIE Press, Bellingham, Washington, second printing 2001.
- Mahajan, V. N. "Zernike Polynomials and Aberration Balancing," *SPIE Proc.* **5173**, 1–17 (2003).
- Mahajan, V. N. *Optical Imaging and Aberrations*, Part II: *Wave Diffraction Optics*, SPIE Press, Bellingham, Washington, Second Printing 2004.
- Mahajan V. N. and Dai, G.-m. "Orthonormal Polynomials for Hexagonal Pupils," *Opt. Lett.* **31**, 2462–2464 (2006).
- Mahajan V. N. and Swantner W. H., "Seidel Coefficients in Optical Testing," *Asian J. Phys.* **15**, 203–209 (2006).
- Malacara, D. J. M. Carpio-Valdéz, and J. Javier Sánchez-Mondragón, "Wavefront Fitting With Discrete Orthogonal Polynomials in a Unit Radius Circle," *Opt. Eng.*, **29**, 672–675 (1990).
- Malacara D. and Z. Malacara, *Handbook of Lens Design*, Dekkar, New York, 1994.
- Montoya-Hernández, M. M. Servín, D. Malacara-Hernández, and G. Paez, "Wavefront Fitting Using Gaussian Functions," *Opt. Comm.* **163**, 259–269 (1999).
- Mouroulis P. and J. Macdonald, *Geometrical Optics and Optical Design*, Oxford, New York, 1997.
- Nijboer, B. R. A. "The Diffraction Theory of Optical Aberrations. Part I: General Discussion of the Geometrical Aberrations," *Physica*, **10**, 679–692 (1943); "The Diffraction Theory of Optical Aberrations. Part II: Diffraction Pattern in the Presence of Small Aberrations," *Physica*, **13**, 605–620 (1947); and Nienhuis K. and B. R. A. Nijboer, "The Diffraction Theory of Optical Aberrations. Part III: General Formulae for Small Aberrations: Experimental Verification of the Theoretical Results," *Physica*, **14**, 590–608 (1949).
- Noll, R. J. "Zernike Polynomials and Atmospheric Turbulence," *J. Opt. Soc. Am.* **66**, 207–211 (1976).
- Prata A. and W. V. T. Rusch, "Algorithm for Computation of Zernike Polynomials Expansion Coefficients," *Appl. Opt.* **28**, 749–753 (1989).
- Swantner W. and W. W. Chow, "Gram-Schmidt orthonormalization of Zernike polynomials for general aperture shape," *Appl. Opt.* **33**, 1832–1837 (1994).
- Tyson, R. K. "Conversion of Zernike aberration coefficients to Seidel and Higher-order Power-series Aberration Coefficients," *Opt. Lett.*, **7**, 262–264 (1982).
- Thunen J. and Kwon O., "Full Aperture Testing With Subaperture Test Optics," *Proc. SPIE*, **351**, 19–27 (1982).
- Upton R. and B. Ellerbroek "Gram-Schmidt orthogonalization of the Zernike polynomials on apertures of arbitrary shape," *Opt. Lett.* **29**, 2840–2842 (2004). As pointed out by Mahajan

and Dai (2006), the matrix for obtaining the orthonormal hexagonal polynomials from the Zernike circle polynomials has numerical errors.

Wang J. Y. and D. E. Silva, "Wave-front Interpretation With Zernike Polynomials," *Appl. Opt.*, **19**, 1510–1518 (1980).

Welford, W. T. *Aberrations of the Symmetrical Optical System*, Academic Press, New York, 1974.

Wyant J. C. and K. Creath, "Basic Wavefront Aberration Theory for Optical Metrology," *Appl. Opt. Optical Eng.* **XI**, 1–53 (1992).

Zernike, F. "Diffraction Theory of Knife-edge Test and its Improved Form, the Phase Contrast Method," *Mon. Not. R. Astron. Soc.* **94**, 377–384 (1934a); "Beugungstheorie Des Schneidverfahrens und Seiner Verbesserten Form, der Phasenkontrastmethode," *Physica* **1**, 689–794 (1934b).

14

Phase Shifting Interferometry

Horst Schreiber and John H. Bruning

14.1. INTRODUCTION

The single biggest change in all types of instrumentation over the past 30 years or so has been the integration of computers into the measurement system. Interferometry is no exception, and the prime manifestation of this situation has been the development of phase shifting interferometry (PSI). Unlike many of the techniques discussed elsewhere in this book, PSI is not a specific optical hardware configuration, but rather a data collection and analysis method that can be applied to a great variety of testing situations.

While the analysis of static interferograms has benefited from computerization, it suffers from the need to find fringe centers, and the resulting trade-off between precision and number of data points. Data is collected only along the fringe centers, and most analyses need the data on a regular grid, therefore requiring interpolation. This low spatial sampling along the fringe centers also limits the lateral resolution of the method. With a static interferogram, an additional piece of information is also required to determine the polarity of the wavefront.

PSI electronically records a series of interferograms while the reference phase of the interferometer is changed. The wavefront phase is encoded in the variations in the intensity pattern of the recorded interferograms, and a simple point-by-point calculation recovers the phase. The need to locate the fringe centers is eliminated, as are the associated problems. While the earliest reference to this technique dates back to 1966 (Carré, 1966), the development of PSI techniques really began in the early 1970's (Crane, 1969; Bruning et al. 1974; Wyant, 1975; Johnson and Moore, 1977; Hardy et al. 1977; Stump, 1979; Bruning, 1978; Moore, 1973; Massie, et al. 1979). The applications for this early work included optical testing, real-time wavefront sensing for active optics, and microscopy. With the advent of optical lithography in the 80's, PSI development received a push for higher accuracies, and therefore a theory around algorithms less sensitive to error sources was

developed (Schwider et al., 1983; Creath, 1988; Malacara, 1990; Larkin and Oreb, 1992; Schwider et al. 1993). In the 90s a more rigorous treatment of the PSI approach resulted in several theories on how to develop custom PSI algorithms tailored to the applications need (Surrel 1993, 1996, 1997, 1998, Phillon 1997, Hibino et al. 1995, 1997, 1998, 2004 Malacara-Doblado et al. 2000, 2001, Zhu and Gemma 2001).

Over the years, this general technique has been known by several names including phase-measuring interferometry, fringe scanning interferometry, real-time interferometry, AC interferometry and heterodyne interferometry. All describe the same basic technique. In this chapter, we will review the fundamentals of PSI, show some of the various algorithms that are used, consider repeatability issues and error sources, examine some of the alternate implementations, and look at the future trends for this technology.

14.2. FUNDAMENTAL CONCEPTS

The relatively simple (in retrospect, anyway) concept behind phase shifting interferometry is that a time-varying phase shift is introduced between the reference wavefront and the test or sample wavefront in the interferometer. A time-varying signal is then produced at each measurement point in the interferogram, and the relative phase between the two wavefronts at that location is encoded in these signals.

For simplicity in our discussions, we will use the wavefront phase for our analysis. This choice allows the analysis to be independent of the specific hardware configuration so that the conversion from phase to surface errors or optical path difference (OPD) is straightforward. For example, a surface with height errors $h(x, y)$ tested in reflection will produce a wavefront error $\phi(x, y)$:

$$\phi(x, y) = 4\pi h(x, y)/\lambda \quad (14.1)$$

where x and y are the spatial coordinates and λ is the wavelength. This expression is for normal incidence, and the obliquity factors must be added in other situations.

General expressions for the reference and test wavefronts in the interferometer are

$$w_r(x, y, t) = a_r(x, y)e^{i[\phi_r(x, y) - \delta(t)]} \quad (14.2)$$

and

$$w_t(x, y) = a_t(x, y)e_t^{i\phi(x, y)} \quad (14.3)$$

respectively, where $a_r(x, y)$ and $a_t(x, y)$ are the wavefront amplitudes, $\phi_r(x, y)$ and $\phi_t(x, y)$ are the wavefront phases, and $\delta(t)$ is a time-varying phase shift introduced into the reference beam. To be more precise, $\delta(t)$ is the relative phase shift between

the two beams; it can physically result from changes in either the reference or test beams. The resulting intensity pattern is

$$I(x, y, t) = |w_r(x, y, t) + w_t(x, y)|^2 \tag{14.4}$$

or

$$I(x, y, t) = I'(x, y) + I''(x, y) \cos[\phi_t(x, y) - \phi_r(x, y) + \delta(t)] \tag{14.5}$$

where $I'(x, y) = a_r^2(x, y) + a_t^2(x, y)$ is the average intensity, and $I''(x, y) = 2a_r(x, y)a_t(x, y)$ is the fringe or intensity modulation. If we now define $\phi(x, y)$ to be the wavefront phase difference $\phi_t(x, y) - \phi_r(x, y)$, we obtain

$$I(x, y, t) = I'(x, y) + I''(x, y) \cos[\phi(x, y) + \delta(t)] \tag{14.6}$$

which is the fundamental equation for PSI. The intensity at each point varies as a sinusoidal function of the introduced phase shift $\delta(t)$ with a temporal offset given by the unknown wavefront phase.

This result can be visualized by looking at the intensity as a function of $\delta(t)$; for a linear phase shift, this is equivalent to the variation of the intensity with time. The intensity of the interferogram at an individual measurement point varies sinusoidally with phase (or time) as shown in Figure 14.1, but more importantly, the three unknowns in Eq. (14.6) are easy to identify in this signal. The constant term $I'(x, y)$ is the intensity bias, $I''(x, y)$ is half the peak-to-valley intensity modulation, and the unknown phase $\phi(x, y)$ is related to the temporal phase shift of this sinusoidal variation. The wavefront phase at this location can be easily computed from this temporal delay. The entire map of the unknown wavefront phase $\phi(x, y)$ can be measured by monitoring and comparing this temporal delay at all the required measurement points in the interferogram. In this manner, the intensity modulation and bias terms are separated from this delay, and the measured wavefront phase is therefore also decoupled from these terms.

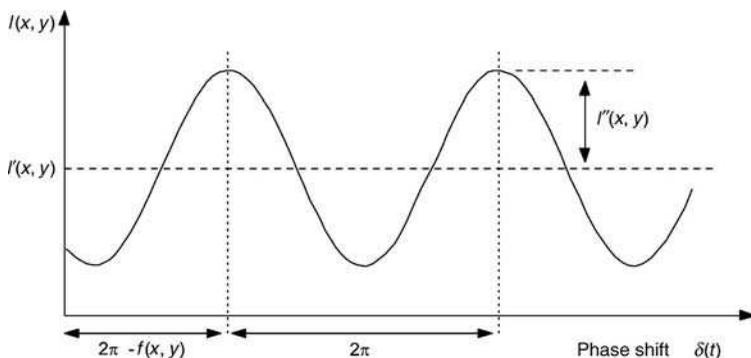


FIGURE 14.1. The variation of intensity with the reference phase at a point in an interferogram.

If the interferogram is viewed during the phase shifting operation, the fringe pattern will appear to move across the field. This is exactly analogous to the fringe motion generated by pushing on the test flat to determine the high and low points when observing Newton's rings. It is important to note that this fringe motion is an artifact of the sinusoidal intensity variations occurring at the same temporal frequency but with different phases at each point in the interferogram. PSI does not use this fringe motion to compute the wavefront; it does not find and track fringe centers. PSI uses the intensity variations at each point for this computation.

14.3. ADVANTAGES OF PSI

Traditionally, interferograms have been analyzed by noting the straightness of the fringes or by identifying the fringe centers and assigning a constant surface height along each fringe. Adjacent fringes represent a height change of a half wave. Finding the fringe centers has been the inherent limit to the precision of the technique and has also restricted the amount of data processing that can be done with the results. Visually, the location of the fringe centers can be identified to within a tenth or a twentieth of a fringe spacing. It is difficult to precisely locate the maxima or minima of the fringe intensity pattern. With a digitized interferogram produced by electronic camera or an input scanner, the situation can be improved to about a fiftieth of a wave under ideal situations. In a conventional interferometer, the observed fringes represent an ensemble of errors in the test part, the reference surface and optical errors in the interferometer. Any one of these errors may be well below the inherent resolution of the interferometer system. In addition, the apparent location of the fringe centers can be influenced by the intensity variations across the interferogram, and spatial sensitivity variations and fixed pattern noise in the detector. The best measurement precision occurs for an interferogram with a few widely spaced fringes, but since the data is collected only along the fringe centers, widely spaced data points are obtained along with very limited information. The number of data points can be increased by adding more tilt fringes to the interferogram pattern, but this is at the cost of reduced repeatability in finding the fringe centers. This trade-off of precision and amount of data is further complicated by the fact that data analysis programs require that the input data be sampled on a regular grid, not along a few curved lines across the part. An interpolation routine is needed to convert the fringe center data to a map of the wavefront on a regular grid.

A further detail needed for analyzing static interferograms is defining the sense of the part. Is it concave or convex? This determination requires another piece of information; for example, the direction in which the fringes move when the reference surface is pushed. With PSI, the phase between the reference and test wavefronts is varied in a known manner and direction during the data collection, and the sense of the part is automatically and unambiguously determined. One way to think of this is that PSI requires that someone "push" on the reference mirror as the data is collected. The use of programs that automatically locate fringe centers sometimes place the requirement on the interferogram that it should contain no closed loop

fringes. This is usually accomplished by adding tilt to the test setup and is needed so that the program can order the fringes from the high point to the low point on the surface.

Since the analysis required for PSI is not dependent on finding fringe centers or following fringes, any type of fringe pattern can be used which is a more practical situation. Even a fringe pattern with no fringes (one very broad fringe covering the entire field of view) or with a complicated series of closed fringes is analyzed correctly. As we will see later, more accuracy is obtained when the number of fringes in the field is minimized. For the same reason, data can be collected and analyzed on any sampling grid that is desired. The most common grid is the regular array of pixels on a solid state sensor, which is compatible with most data processing programs and packages. An irregular grid can also be used if the situation demands it.

Another advantage of PSI is that it is insensitive to spatial variations of intensity, detector sensitivity and fixed pattern noise. This can be seen by modifying Equation 14.6 to include these effects. Fixed pattern noise is a spatial variation in the bias signal that is unrelated to the incident intensity. The detector sensitivity and fixed pattern noise are denoted by $s(x, y)$ and $n(x, y)$, respectively. The spatial intensity variations are already included in the (x, y) dependence of the terms $I'(x, y)$ and $I''(x, y)$. The resulting output from the detector $v(x, y, t)$ is

$$v(x, y, t) = s(x, y)\{I'(x, y) + I''(x, y) \cos[\phi(x, y) + \delta(t)]\} + n(x, y) \quad (14.7)$$

which can easily be rewritten as

$$v(x, y, t) = [s(x, y)I'(x, y) + n(x, y)] + [s(x, y)I''(x, y)] \cos[\phi(x, y) + \delta(t)] \quad (14.8)$$

These additional terms have changed the intensity bias and the intensity modulation of the signal recorded at each measurement point, but have little effect on the measured temporal delay. Since the unknown wavefront phase is encoded entirely in this delay, the precision of the interferometric measurement is generally not degraded by these types of spatial variations. The only reason for these last two statements not being expressed as “no effect” is that there are signal-to-noise considerations that may change the repeatability of the measurement. These effects are especially important if the spatial variations are large or if the output signal modulation is small relative to the bias. Since these detector spatial variations show up as terms indistinguishable from the intensity bias and modulation terms already present in Eq. (14.6), the discussion of algorithms and noise considerations in later sections will include these terms implicitly in $I'(x, y)$ and $I''(x, y)$. Time-varying noise sources, such as detector dark current and photon noise, must also be included. These signal-to-noise considerations are discussed at greater length in Section 14.9.

It should not be surprising to learn that the repeatability or precision of PSI can be much greater than that of static interferogram analysis. Since a series of data is collected over time, some advantage is gained through averaging the sequentially recorded interferograms, but this simple explanation does not give the full reason.

The data collection and analysis procedure for PSI is fundamentally different from that used for the analysis of a static interferogram and results in a large improvement in the repeatability for PSI. The method used for PSI is much more analogous to that used with lock-in amplifiers or any of the local-oscillator heterodyne detection schemes used for electronic signals. Repeatabilities of a hundredth of a wavelength are routine with PSI, and the values of a thousandth of a wavelength have been reported and are fairly easy to obtain with today's commercial equipment and a little care. Once again, the particulars of this discussion can be found in Section 14.9.

In Eq. (14.6) and (14.8), we have assumed that the only term which has any time dependence is $\delta(t)$, and any deviation from this assumption will introduce errors into the PSI analysis. The unknown wavefront phase $\phi(x, y)$ is encoded in the measured time delay at each point, and temporal variations occurring during the measurement interval will degrade the performance of a PSI system. In addition to random noise, a common source of error is vibration and turbulence. This is a significant difference from static interferometry where a snapshot of the interferogram can be obtained. Another possible source of error is temporal laser intensity variations, but this is usually not a large problem as a warmed-up laser is fairly stable over typical measurement times, and stabilized lasers are available for exacting situations. A third source of error unique to PSI is errors or variations in the phase shift $\delta(t)$. The functional form must be correct, and the phase shift should be uniform over the interferogram.

14.4. METHODS OF PHASE SHIFTING

By far the most common method used to introduce the time-varying phase shift in a PSI system is to translate one of the mirrors or optical surfaces in the interferometer with a piezoelectric transducer (Soobitsky, 1986; Hayes, 1989). Composed of Lead-Zinc-Titanate (PZT) or other ceramic materials, these devices expand or contract with an externally applied voltage. Depending on the configuration, up to a few hundred volts may be needed to obtain the the required motion of a wavelength or less. By discretely changing the applied voltage, the induced phase shift varies through a series of steps. If the voltage is programmed to vary smoothly, a phase shift of a desired functional form can be produced. Information from the interferometric test can be used to calibrate the motion of the PZT, and this is further discussed in Section 14.9.

The phase shift configurations for three common interferometer types (the Twyman–Green, the Mach–Zehnder and the laser Fizeau) are shown in Figure 14.2. The first two achieve the phase shift by translating one of the mirrors in the reference arm of the interferometer. At normal incidence in the Twyman–Green, a phase shift of a full wavelength occurs for each half wavelength of translation. Because of the non-normal angles in the Mach–Zehnder, the induced phase shift is reduced by a factor of the cosine of the angle of incidence (Fig. 14.3). Since the reference wavefront is plane, the small lateral displacement is ignored. In the laser Fizeau interferometer, the transmissive reference objective or the test piece is

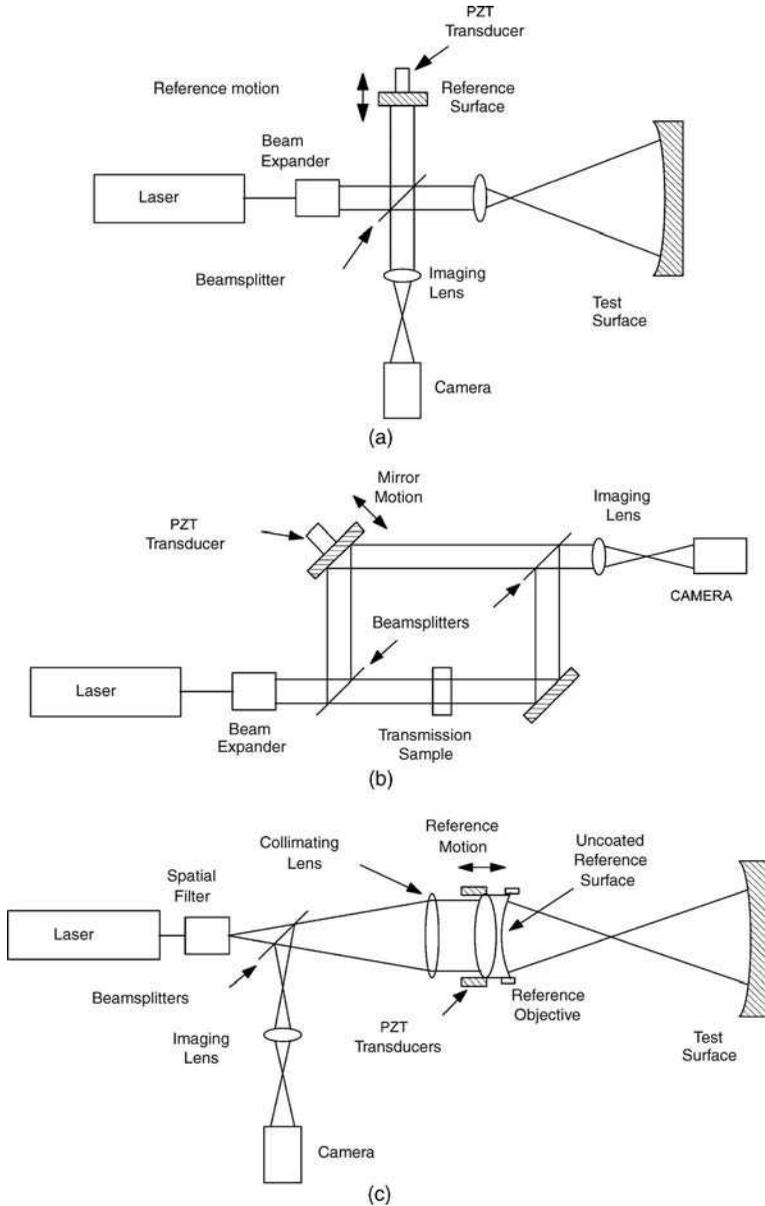


FIGURE 14.2. Three common interferometer configurations: (a) Twyman–Green, (b) Mach–Zehnder, and (c) laser Fizeau.

translated for the phase shift. The reference surface is the final uncoated surface of the objective, and this surface is concentric with the focus. For fast F-numbers, errors or nonuniformities in the phase shift result because the ray directions at the edge of the pupil are not parallel to the translation direction (Moore and Slaymaker, 1980).

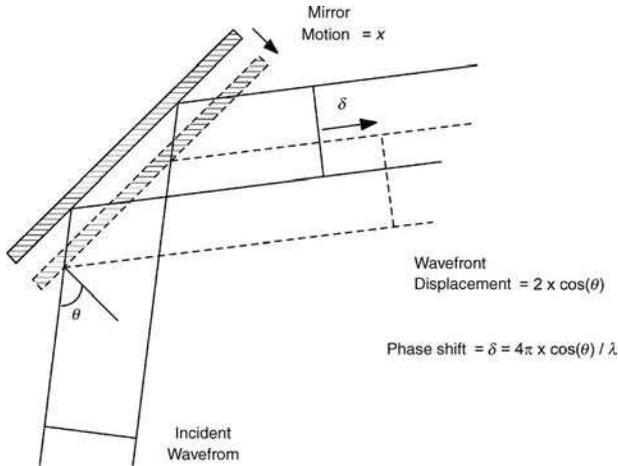


FIGURE 14.3. The induced phase due to a mirror translation at nonnormal incidence.

An alternative method for producing either a stepped or continuous phase shift is to use a tilted plane-parallel plate in the reference beam of the interferometer (Wyant and Shagam, 1978). The optical path within the plate increases as the tilt angle is increased. To avoid introducing aberrations into the reference beam, this method is useful only for a collimated reference beam.

A different type of phase shift can be produced between the reference and test beams by introducing an optical frequency difference between these two beams. If the two optical frequencies are ν and $\nu + \Delta\nu$, Eqs. (14.2) and (14.3) can be rewritten as follows to include this time dependence:

$$w_r(x, y, t) = a_r(x, y)e^{i[\phi_r(x, y) - 2\pi(\nu + \Delta\nu)t]} \quad (14.9)$$

$$w_t(x, y, t) = a_t(x, y)e^{i[\phi_t(x, y) - 2\pi\nu t]} \quad (14.10)$$

No other temporal phase shift is included. The resulting interferogram intensity pattern, analogous to Eq. (14.6), is

$$I(x, y, t) = I'(x, y) + I''(x, y) \cos[\phi(x, y) + 2\pi\Delta\nu t] \quad (14.11)$$

The frequency difference gives rise to a linear phase shift between the test and reference beams:

$$\delta(t) = 2\pi\Delta\nu t \quad (14.12)$$

and the intensity at a given location varies sinusoidally at the difference in frequency $\Delta\nu$. As before, the wavefront phase $\phi(x, y)$ is a spatially varying delay between signals at the various measurement points.

Since even a small optical frequency shift (as a fraction of the optical frequency) can result in a large difference frequency, this phase shifting method is very useful for situations where dynamic measurements are required. In these situations, the phase shift must occur faster than the changes that are being monitored. Two important applications are the measurement of turbulent flow and distance measuring interferometry. The two primary methods to produce a frequency shift are the Zeeman-split laser and Doppler shifts introduced by moving gratings. The application and use of this type of phase shift is discussed in Sections 14.14.3 and 14.16.1.

The output frequency of a laser can be split into two orthogonally polarized output frequencies by the application of a DC magnetic field (Burgwald and Kruger, 1970). The frequency separation of a Zeeman-split two-frequency laser is controlled by the magnetic field, and the values of about 1.8–5 MHz are found for this effect in helium–neon lasers.

When a diffraction grating is translated through a beam of light, a Doppler shift is introduced in the diffracted beams (Suzuki and Hioki, 1967; Stevenson, 1970; Bryngdahl, 1976). The translation direction is perpendicular to the propagation direction (see Fig. 14.4). The frequency shift is proportional to the diffraction order m and the velocity v , and inversely proportional to the grating period d , or

$$\Delta\nu = mv/d \quad (14.13)$$

The undiffracted beam has no frequency shift, beams diffracted in the same direction as the translation see a positive frequency shift, and beams diffracted in the opposite direction have their frequency decreased. One of the diffracted orders is selected, and interfered with the original frequency to produce a phase shifting interferometer.

Mechanical translations, including rotating a radial grating about its center, can be used, but relatively low frequency shifts are obtained. Similarly, a small frequency shift equal to $2v/\lambda$ is produced by the reflection from a moving mirror. Frequency shifts even larger than those of the Zeeman-split laser can be produced by an acousto-optic (AO) Bragg cell (Massie and Nelson, 1978; Wyant and Shagam, 1978; Shagam, 1983). Acoustic waves traveling through the cell create periodic index of refraction variations due to the sound pressure, and a moving diffraction grating is set up in the cell. Since the velocity divided by the period of the waves is equal to the acoustic frequency f , the frequency shift for this AO modulator is

$$\Delta\nu = mf \quad (14.14)$$

Extremely high phase shift rates are achievable with this type of device; some commercially available instruments have operating frequencies of 20–30 MHz.

Phase shifts can also be obtained through the use of rotating polarizers (See Fig. 14.5) or phase retarders, but mechanical limitations generally limit operation to a few kilohertz, and this method is not widely used (Crane, 1969; Okoomian, 1969; Bryngdahl, 1972a; Sommargren, 1975; Shagam and Wyant, 1978; Hu, 1983, Kothiyal and Delisle, 1984, 1985a).

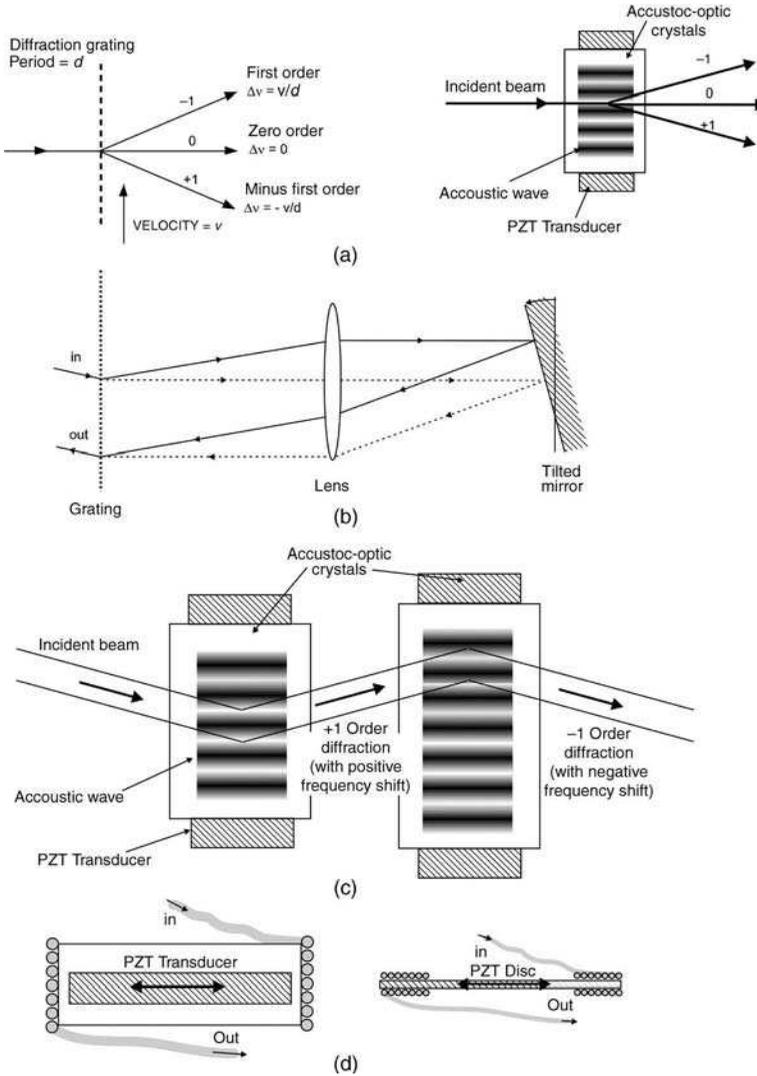


FIGURE 14.4. Various modern phase-shifter implementations: (a) doppler frequency shifts introduced by a moving grating or a acousto optical modulator; (b) frequency domain optical delay line; (c) phase-shifter based on two acousto optical modulators; (d) optical delay using fiber-stretchers.

An achromatic optical phase shifter-modulator was demonstrated by Zvyagin (2001) based on a frequency-domain optical delay. The device able to maintain zero group delay consisting of a grating, lens and a mirror. Tilting the mirror allows for adjusting phase delay, without change to the group velocity of the beam as shown in Figure 14.4(b). This especially is an advantage for multiple wavelength PSI as described later.

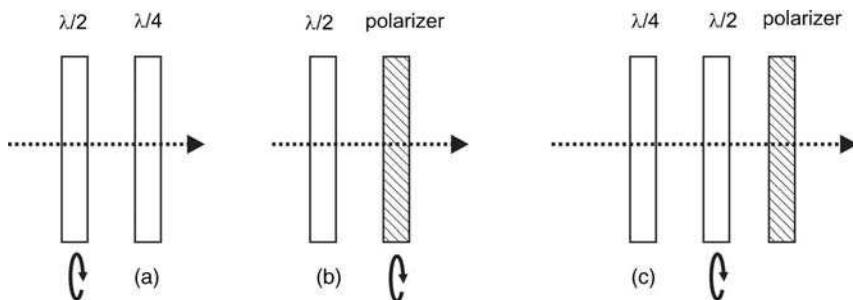


FIGURE 14.5. Polarization phase shifter arrangements (a) for the interferometer input (b,c) for the output of the interferometer (compare Kothiyal, 1985).

The advances in fiber-optics around the turn of the century brought additional approaches for introducing phase shifts. Improved accusto optical modulators allow for a compact on axis device, that enables full control of the phase in a collimated beam. An accusto optical modulator consists of a crystal that by applying a frequency using a piezo-transducer exhibits density modulation in the crystal along the wave propagation. The incoming beam orthogonal to the propagating wave now will diffract on these density modulations. A single accustom optical modulator can be used much like a grating as described earlier, but the frequency of the grating can be changed, for example, switched on and of and allows for modification of the otherwise static grating behavior. By combining two accustom-optical modulators in series, the addition of a positive and negative first diffraction order is on axis again and allows for a compact achromatic phase shifter design as showed in Figure 14.4(c) (Li, 2005).

The use of optical fibers also allow for another method by adjusting the path inside a fiber using stretching of the fiber itself for a small amount. Usually a fiber coil is expanded using a piezoelectric transducer or a fiber coil is fixed to a piezoelectric disc for allowing expansion or compression as shown in Figure 14.4(d). These methods allows for fast scanning rate of 4 kHz and achieving up to 4.5 millimeters of fiber stretch for a coil of 20 meter fiber (Henderson et al., 2004). Figure 14.5 shows common phase shifter arrangements using polarization optical components like quarter/half wave plates and polarizers, allowing a variable phase shift by changing the selected directions of polarization.

Another recent advance is the direct modulation of the output wavelength of a laser diode using temperature or current to produce phase shifts (Ishii et al., 1987; Suzuki et al., 2001; Onodera and Ishii, 2003). The operation and implications for this method are discussed in Section 14.6.1.3.

14.5. DETECTING THE WAVEFRONT PHASE

The major difference between the various PSI detection schemes is the manner in which the reference phase is varied, and the number of times and the rate at which the interference pattern is measured. All of the variations we will discuss aim to

find the “best” way to collect and analyze the interferometric data so that Eq. (14.6) may be solved for the unknown wavefront phase. Of course, the best solution for a particular application will depend on a number of factors including computational complexity and speed, sensitivity to phase shift errors and noise, data rates, and compatibility with the detection scheme. To generate a fundamental understanding of the analysis process, we will start with a conceptionally and analytically simple algorithm.

The *four step algorithm* requires that four separate interferograms of the part under test are recorded and digitized. A 90° optical phase shift is introduced into the reference beam between each of the sequentially recorded interferograms. Since these are now discrete measurements, the time dependence has been changed to the phase step index i . The function $\delta(t)$ now takes on four discrete values:

$$\delta_i = 0, \pi/2, \pi, 3\pi/2; \quad i = 1, 2, 3, 4 \quad (14.15)$$

Substituting each of these four values into Eq. (14.6) results in four equations describing the four measured interferogram intensity patterns:

$$I_1(x, y) = I'(x, y) + I''(x, y) \cos[\phi(x, y)] \quad (14.16)$$

$$I_2(x, y) = I'(x, y) + I''(x, y) \cos[\phi(x, y) + \pi/2] \quad (14.17)$$

$$I_3(x, y) = I'(x, y) + I''(x, y) \cos[\phi(x, y) + \pi] \quad (14.18)$$

and

$$I_4(x, y) = I'(x, y) + I''(x, y) \cos[\phi(x, y) + 3\pi/2] \quad (14.19)$$

A simple trigonometric identity yields

$$I_1(x, y) = I'(x, y) + I''(x, y) \cos[\phi(x, y)] \quad (14.20)$$

$$I_2(x, y) = I'(x, y) - I''(x, y) \sin[\phi(x, y)] \quad (14.21)$$

$$I_3(x, y) = I'(x, y) - I''(x, y) \cos[\phi(x, y)] \quad (14.22)$$

and

$$I_4(x, y) = I'(x, y) + I''(x, y) \sin[\phi(x, y)] \quad (14.23)$$

These four equations in three unknowns [$I'(x, y)$, $I''(x, y)$ and $\phi(x, y)$] can now be solved for the value of $\phi(x, y)$ at each point in the interferogram. As will be shown later, three interferograms are all that is required to solve for the wavefront phase, but the fourth is included for computational ease. The intensity bias term $I'(x, y)$ is eliminated by subtracting the equations in pairs:

$$I_4 - I_2 = 2I''(x, y) \sin[\phi(x, y)] \quad (14.24)$$

and

$$I_1 - I_3 = 2I''(x, y) \cos[\phi(x, y)] \quad (14.25)$$

Taking the ratio of these two equations eliminates the intensity modulation term $I''(x, y)$ to produce a result that contains only the unknown phase $\phi(x, y)$ and the four measured intensities:

$$\frac{I_4 - I_2}{I_1 - I_3} = \frac{\sin[\phi(x, y)]}{\cos[\phi(x, y)]} = \tan[\phi(x, y)] \quad (14.26)$$

This equation can now be rearranged to produce the result for the four step PSI algorithm:

$$\phi(x, y) = \tan^{-1} \left[\frac{I_4 - I_2}{I_1 - I_3} \right] \quad (14.27)$$

This simple equation is evaluated at each measurement point to obtain a map of the measured wavefront. To show the simplicity of the algorithm and to allow easier comparison with other algorithms, the (x, y) dependence of the four measurements is implied. This wavefront can be easily related to the surface height of the part (Eq. 14.1) or the optical path difference (OPD):

$$\text{OPD}(x, y) = \lambda\phi(x, y)/2\pi \quad (14.28)$$

The measurement array that is often used for PSI is the rectangular array of pixels on a solid state imaging detector.

Figure 14.6 shows the photographs of four interferograms recorded with a 90° phase shift. The fringes appear to walk across the frame, but more importantly, the intensity at any point varies with phase. Since there is a 360° phase shift between the first and fifth frames, they appear nominally identical.

At this point, the differences between PSI and conventional interferometric analysis should be very apparent. PSI calculates the wavefront phase at every

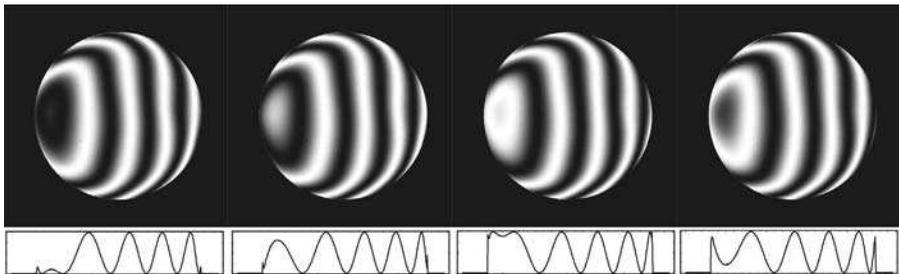


FIGURE 14.6. Four interferograms recorded with phase shifts of 90° and a horizontal Line-Scan through their center.

measurement location from the time varying intensity measured at that point. The result at a point is the arctangent of the ratio of the intensities measured at that point, and there is no need to find the fringe centers or to order the fringes. In fact, the fringe pattern is somewhat irrelevant to PSI; there is no requirement on the minimum number of fringes or their shape, and the system precision is relatively independent of the fringe frequency in the interferogram. Since the data can be collected on a regular grid, PSI is well suited for further analysis and processing.

If desired, the intensity data can also be evaluated to determine the data modulation $\gamma(x, y)$ across the interferogram:

$$\gamma(x, y) = \frac{I''(x, y)}{I'(x, y)} \quad (14.29)$$

Starting with Eqs. (14.24) and (14.25) and also using Eqs. (14.20)–(14.23), we obtain:

$$\gamma(x, y) = \frac{2[(I_4 - I_2)^2 + (I_1 - I_3)^2]^{1/2}}{I_1 + I_2 + I_3 + I_4} \quad (14.30)$$

where the (x, y) dependence of the measurements is implied. The numerator is the intensity modulation, and the denominator is the average intensity or the intensity bias. Once again, all of the detector characteristics such as sensitivity and bias are implicitly included in this analysis as part of $I'(x, y)$ and $I''(x, y)$. This information is useful for evaluating the quality of the data that has been collected. A data modulation near one is good, and a low modulation is bad. Data points with modulations below some threshold will have insufficient signal to noise, and the wavefront phase cannot be reliably calculated at that point. These points are excluded from the analysis, and typical data modulation thresholds are in the range of 5–10%. This value is, of course, application dependent.

14.6. DATA COLLECTION

As described earlier, there are actually many ways to acquire the intensity signal. This section summarizes several methods in common use. The fundamental idea behind phase shifting is that a minimum of 3 different intensity patterns are required to evaluate the desired phase. A phase shift can be created temporally by changing the path difference between two interferometer arms or by changing the wavelength of the light source. A phase shift can also be created instantaneously by spatial displacement or by a change in optical path length over the aperture of the interferometer resulting in a phase shifting spatially.

14.6.1. Temporal Methods

There are several ways to temporally change the reference phase. These are separated into discrete stepping and continuous scanning by a phase shifter or stepping and scanning of the wavelength.

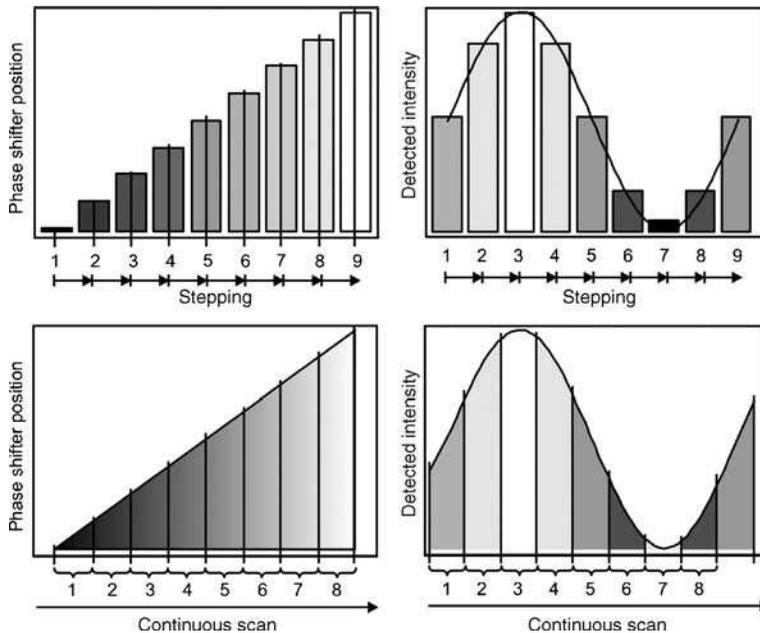


FIGURE 14.7. Stepping the phase (top) and integrating over a phase shift (bottom) are alternative methods for temporal phase shifting.

Stepping the Phase. To illustrate the phase stepping process, the four step algorithm is used in Section 14.5. The reference phase is stepped through a series of discrete values and an interferogram is recorded at each step, as shown in Figure 14.7. For a clarity, Fig. 14.7 shows 8 steps within one period. The upper left picture shows phase shifter steps. The upper right illustrates the intensity recorded by a detector pixel corresponding to changes in the reference phase.

Scanning the Phase. For practical reasons, it is often desirable to vary the phase smoothly between steps as the interferograms are recorded. One reason to do this is that after a commanded phase step, the PZT transducer may tend to “ring,” where the reference mirror position may oscillate until transients die out. (Seligson et al., 1984). This is a likely situation when the reference surface has a large mass or the control electronics are not well damped. It may therefore be necessary to wait after each phase step before recording the data, which may greatly increase collection time.

The *integrating bucket* data collection scheme, which was first proposed by Wyant (1975), allows the reference phase to vary linearly with time while a series of interferograms are collected. The reference phase changes during the integration time for a single measurement. This integration time may be chosen to equal the frame rate of the sensor, or it could be the time required to collect adequate signal.

For this analysis, we first rewrite the general expression for an interferogram in Eq. (14.6) as a function of the reference phase δ instead of time:

$$I(x, y, \delta) = I'(x, y) + I''(x, y) \cos[\phi(x, y) + \delta] \quad (14.31)$$

If the reference phase changes by an amount Δ during the integration or exposure time, the recorded interferogram is found by integrating the intensity over this interval (Greivenkamp, 1984):

$$I_i(x, y) = \frac{1}{\Delta} \int_{\delta_i - \Delta/2}^{\delta_i + \Delta/2} \{I'(x, y) + I''(x, y) \cos[\phi(x, y) + \delta]\} d\delta \quad (14.32)$$

where δ_i is the phase shift at the center of the integration period, and $I_i(x, y)$ is the corresponding recorded interferogram. The term $1/\Delta$ in this equation normalizes the results so that the average recorded signal is independent of Δ . Integration yields

$$I_i(x, y) = I'(x, y) + \left(\frac{1}{\Delta}\right) I''(x, y) \left\{ \sin \left[\phi(x, y) + \delta_i + \frac{\Delta}{2} \right] - \sin \left[\phi(x, y) + \delta_i - \frac{\Delta}{2} \right] \right\} \quad (14.33)$$

which, after the use of a trigonometric identity, can be rewritten as

$$I_i(x, y) = I'(x, y) + I''(x, y) \text{sinc}[\phi(x, y) + \delta_i] \quad (14.34)$$

where

$$\text{sinc}(\beta) = \frac{\sin(\pi\beta)}{\pi\beta} \quad (14.35)$$

Under this definition, the sinc function has a value of one at zero and goes to zero whenever the argument is an integer. Comparing this result to the original expressions in Equations 14.31 or 14.6 indicates that the only change due to integrating the intensity while the phase is changing is to reduce the modulation of the recorded intensity at each point in the interferogram by the value of the sinc function. This expression is the most general of those we have discussed. Note that the value of Δ must be the same for all frames recorded as part of a data set. The lower row of Figure 14.7 illustrates the continuous movement of the phase and the recorded intensity per detector pixel.

The same algorithms can be used for data collected by the phase step method or the integrating bucket method. In fact, we can think of phase stepping as an integrating bucket of zero width ($\Delta = 0$). In this case, the integrating bucket result reduces to phase stepping result.

There are some signal-to-noise concerns with using the integrating bucket approach since the PSI analysis depends on having data with good modulation.

Some data points may fall below the threshold required for proper analysis as a result of this reduction in modulation. The sinc function serves as a modulation transfer function (MTF) for the measurement. As Δ increases, the MTF decreases.

For phase stepping ($\Delta = 0$), it is unity. Small integration periods have minimal effects; for example, the modulation for $\Delta = \pi/4$ (45°) and $\pi/2$ (90°) are 97% and 90%, respectively. The $\Delta = \pi/2$ situation corresponds to collecting four contiguous data frames while the reference phase changes by a full wave and approximates the condition found in many available interferometers. Integrating the intensity over a full period ($\Delta = 2\pi$) of reference phase shift reduces the signal modulation to zero, and there is no useful information about the wavefront phase in the resulting data. To obtain good data modulation, we are usually restricted to integration periods between 0 and π for each frame. There is no advantage to the integrating bucket approach unless there are dynamic or control issues associated with the data collection. A 10% reduction in effective modulation is a price that must be weighed against other factors.

Stepping or Scanning the Wavelength. For unequal path interferometers like a laser-Fizeau or unequal path Twyman–Green configurations changing the wavelengths allows for a change of the phase in the interferogram pattern (de Groot, 2000, Dunn, 2005). As shown by Deck (2001) for a given difference L between test and reference arm in an interferometer, the resulting change in phase is a linear function of the change in laser frequency:

$$\Delta\phi = \phi_{test} - \phi_{reference} = 2\frac{2\pi}{\lambda}n \cdot L = \frac{4\pi n}{c}L \cdot v \quad \Rightarrow \quad \frac{\partial\Delta\phi}{\partial t} = \frac{4\pi n}{c}L \cdot \frac{\partial v}{\partial t} \quad (14.36)$$

Therefore the wavelength change has to be large enough to allow enough phase change for the given interferometer cavity length.

Figure 14.8 shows two of the most common external cavity feedback options for tuning laser diodes (Duarte, 2003). In both cases, the laser beam emitted from

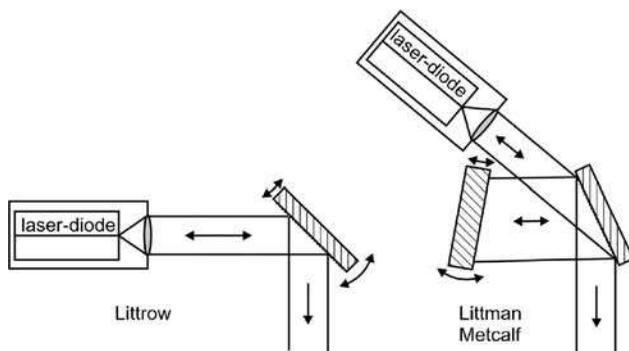


FIGURE 14.8. Two of the most common external feedback configurations for wavelength tuning using laser-diodes.

the laser diode is collimated before it hits the grating. In the Littrow configuration, the zero order exits the cavity and the first order is reflected back into the laser. This external feedback of the only wavelength at which dispersion angle is reflected back into the cavity causes the diode laser to resonate at exactly that wavelength. By turning the grating using a piezoelectric transducer the wavelength reflected back into the laser can be changed and so allows to adjust the emitting wavelength.

In the Littman/Metcalf configuration the first-order beam is coupled out of the laser, while the zero-order beam is coupled to a rotatable mirror which reflects the beam back onto the grating. There, the first-order reflection reflects back into the laser cavity. The changing of the mirror angle using a piezoelectric transducer allows to tune the lasers resonant wavelength within the gain curve.

Both configurations have advantages and disadvantages. In the Littrow configuration, which delivers the higher output power, the exiting beam of the laser grating combination changes angle when rotating the grating. In the Littman/Metcalf configuration the exit angle stays constant; however, the zero order from the second gratings reflection is lost and so limits the overall output power of the system.

The wavelength tuning method has significant advantages in Fizeau configurations, because there are no moving parts in the system except for the ones in the light-source. It especially allows to avoid the often used shifting of one of the collimation optics in the interferometer and therefore it introduces decollimation in doing so.

The wavelength tuning approach allows to separate out different reflections from different surfaces like front and back surface of optics under test as described by Okada et al. (1990) and de Groot (1994). In principle, different surface reflections will have a different phase change, because the wavelength change introduces a phase change relative to the cavity length. And the cavity length for different surface reflections is different.

14.6.2. Spatial Methods

Besides the previous described methods where temporally a phase change is introduced, there are several methods which use a spatial approach to introduce the required phase-difference. The spatial methods became increasingly popular when higher resolution detectors became available. One advantage of the spatial methods is that they do not require the signal phase to be stable over a specific time, because usually they are instantaneous methods whose actual data taking time is only determined by the integration time of the detector.

Spatial Dense Detection. Having a detector with sufficient resolution allows for intruding a spatial carrier frequency into the detected intensity signal. This should not be done physically using a large amount of tilt between the test and reference arm, because the beam would propagate using different parts in the viewing system introducing additional aberrations in the detected interferogram. Instead it can be done directly in front of the detector by using left handed and right handed circular polarized light in the test and reference arms and the placing an optical element

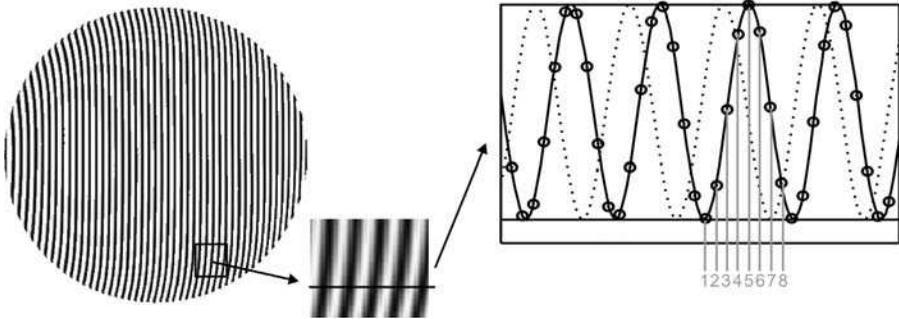


FIGURE 14.9. Carrier frequency allows the use of spatial phase-shifting algorithms.

introducing tilt between the orthogonal directions like a Wollaston prism followed by an analyzer in front of the detector.

If the wavefront slope in relation to the introduced tilt fringes is small, then the wavefront can be locally considered flat, which results in a linear phase variation and the phase difference between adjunct pixels is constant as described in Shough, 1990. This allows to use phase-shift algorithms. To reduce the errors introduced by larger wavefront slopes, which result in the phase difference between adjunct pixels not being equal the error compensating algorithms, that allow for an error in the phase-shift should be used. A comprehensive discussion is given in Kuechel's publications (1990, 1992a, 1992b) and Malacara-Doblado, Kuechel's publication in 2001.

Figure 14.9 illustrates the spatial dense methods. The same wavefront as already displayed in Figure 14.6 is now recorded with a significant amount of additional tilt. If looking locally, the phase difference between adjunct pixels is nearly constant. This allows to evaluate the phase for these pixels. Usually because error compensating algorithms should be used, this method uses just the 4 phase values more to include and allow for the wide range-detuning insensitive algorithms. As described in Kuechel's publications (1990, 1992a, 1992b) this method can be implemented electronically doing fast evaluations, resulting in video-realtime display of the evaluated phase values, enabling fast averaging and vibration insensitivity. Because the method calculates one phase point out of several neighboring pixels, the spatial resolution in the axis of the carrier frequency is reduced.

Laterally Displaced Images. The first of these instantaneous systems uses polarizing elements to produce the desired PSI signals (Smythe and Moore, 1984).

A polarization interferometer is used where the two polarizations are each sent to one arm of the interferometer. The returning beams are sent to the decoding module shown in Figure 14.10. The halfwave plate rotates both polarizations by 45° so that they are at 45° with the axes of the polarizing components. An intensity beamsplitter (BS) creates two beams, and the reflected beam is further split by a polarizing beamsplitter (PBS). Two sets of interference fringes are formed with a phase shift of 180° due to conservation of energy. The other beam from the intensity

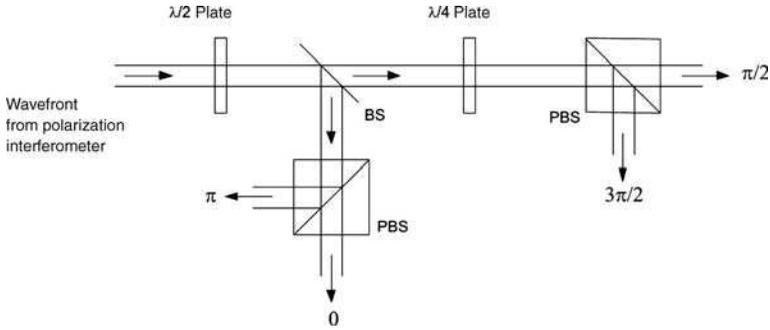


FIGURE 14.10. Polarization decoding scheme for instantaneous PSI measurements.

beam splitter goes through a quarterwave plate whose fast axis is aligned with one of the orthogonal polarizations; this introduces a 90° phase shift between the two polarizations. When this beam is analyzed by a second polarization beam splitter, the two interferograms will be 180° out of phase and also shifted by 90° from the first set of fringe patterns. These four interferograms with 90° phase shifts can be recorded and wavefront can be calculated with the four step algorithm (Eq. 14.27). Similar polarization encoding schemes have been applied to distance measuring interferometers (Dorsey et al., 1983; Crosdale and Palum, 1990).

Hettwer et al. showed in 2000, that the combination of a grating followed by quarter-wave plates and an analyzer can be used to do instantaneous phase shifting on a single detector (See also Sec. 2.8). The grating acts in this set-up as a beam splitter. A lens in the right placement and focal length collimates the zero and plus/minus first order diffracted beams back onto a detector. Before the detector, an analyzer is used to select out a single interfering polarization. The plus and minus first orders have quarter wave plates inserted that rotate the plus and minus first order in opposite directions 90° . This results in 3 spatially separated beam on the detector with -90° , 0° , and $+90^\circ$ phase shift between them. In this arrangement if a phase grating is used, the diffraction intensities can be engineered in a way that the first order beam with the analyzer in place have the same intensities than the zero order beam without the analyzer.

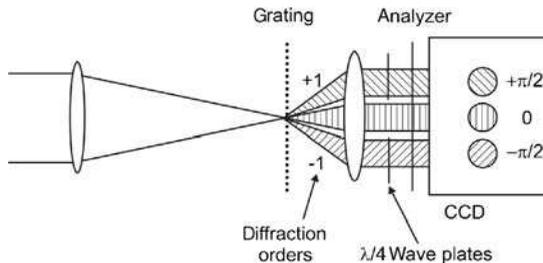


FIGURE 14.11. Schematic set-up of inline lateral displaced phase shifted images using one detector (Hettwer et al., 2000).

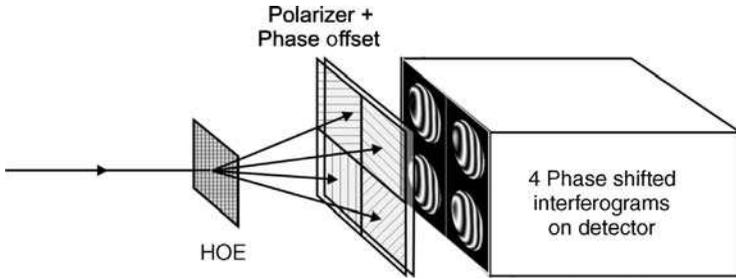


FIGURE 14.12. Simultaneous phase shifting using a holographic optical element (HOE) and 4 spatially separated polarizers before the detector.

This three frame approach does not allow to use error compensating algorithms nor are the usually square detectors used efficiently.

Millerd et al. (2004, 2005) published a very compact approach allowing for 4 phase shifted interferograms, enabling the use of error compensating algorithms and also space wise using the detector geometry more effectively. They used a specially designed holographic element to split the beam-up into 4 separate beams laterally displaced in x and y . An additional phase offset and polarizer mask placed in front of the detector allows introducing the desired phase shift in the interferogram pattern as shown in Figure 14.12.

Pixel Displaced. Millerd et al. (2005) also showed that wire grid polarizers can be used rotated for each pixel to introduce different phase shift. For 4 phase steps this would be 4 different analyzer rotations. In this way, using a high resolution detector 4, instantaneous interferogram can be obtained without an additional holographic optical element to split the beam. Figure 14.13 illustrates the basic principle of this

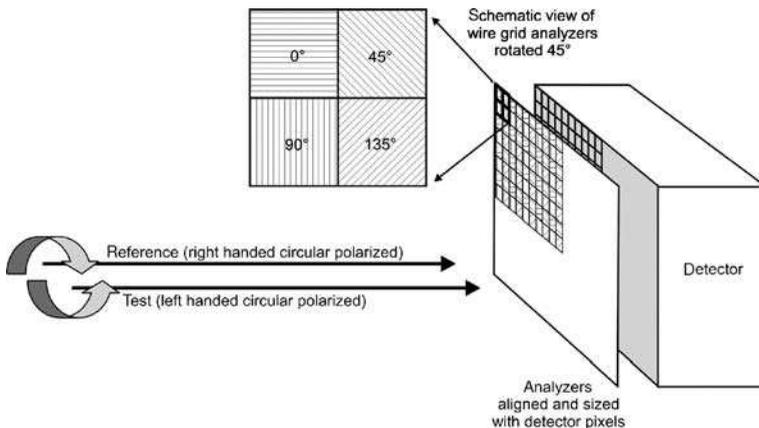


FIGURE 14.13. Wire grid polarizer array directly placed in front of camera.

set-up. The polarizing interferometer has different left handed or right handed polarization coming from test or reference arm respectively. Both beams then path through an array with matching wire grid analyzers aligned to the detector pixels. Every analyzer can have a different rotation per pixel, selecting a different phase-difference between the interferometer arms and also resulting in different phase shifts per pixel. In this way, four high-resolution instantaneous interferograms can be recorded by sorting out the pixels belonging to 0° , 45° , 90° , and 135° analyzer rotation.

14.7. PSI ALGORITHMS

In Section 14.6, we described several methods for collected phase-shifted intensity patterns. A number of different data evaluation strategies or algorithms have been developed, of which the four step algorithm (14.27) is just one example. All of these algorithms share common characteristics: They require that a series of interferograms are recorded as the reference phase is varied. The wavefront phase, modulo 2π , is then calculated at each measurement point as the arctangent of a function of the interferogram intensities measured at that same measurement point. The final wavefront map is then obtained by “unwrapping” the phases to remove the 2π phase discontinuities. While there are many different approaches to calculating the phase, the differences between the various algorithms relate to the number of interferograms recorded, the phase shift between these interferograms, and the susceptibility of the algorithm to the errors in the phase shift, environmental noise such as vibration and turbulence as well as nonlinearities of the detector when recording the intensities.

Since we have shown that the difference between the phase step and integrating bucket methods relates to the effective modulation of the recorded intensities, all of the algorithms that we will discuss are equally valid for both approaches. To simplify the notation, we will rewrite the general expression in Eq. (14.34) as

$$I_i(x, y) = I'(x, y) + I''(x, y, \Delta) \cos[\phi(x, y) + \delta_i] \quad (14.37)$$

where

$$I''(x, y, \Delta) = I''(x, y) \text{sinc}(\Delta/2\pi) \quad (14.38)$$

The dependence of the modulation on the integration period Δ is now implicit in $I''(x, y, \Delta)$. In this section, we are mainly concerned with algorithm development and will highlight the sensitivity of each of the algorithms to the most common error sources. Error sources themselves are more closely examined in Section 14.10.

With these PSI algorithms, the starting value of the reference phase, while arbitrary, is often chosen to produce a simpler mathematical expression for the measured wavefront phase. In practice, we do not know (or need to know) the absolute

reference phase; what is important is the phase shift between measurements. We merely define the starting position of the reference mirror to be at the first required phase value, and proceed from there.

14.7.1. Three Step Algorithms

Since there are three unknowns in Eq. (14.39), the minimum number of measurements of the interferogram intensity required to reconstruct the unknown wavefront phase is three. The general case can be solved using equal phase steps of size α (Creath, 1988). In this case,

$$\delta_i = -\alpha, 0, \alpha; \quad i = 1, 2, 3 \quad (14.39)$$

and

$$I_1 = I'(x, y) + I''(x, y, \Delta) \cos[\phi(x, y) - \alpha] \quad (14.40)$$

$$I_2(x, y) = I'(x, y) + I''(x, y, \Delta) \cos[\phi(x, y)] \quad (14.41)$$

and

$$I_3(x, y) = I'(x, y) + I''(x, y, \Delta) \cos[\phi(x, y) + \alpha] \quad (14.42)$$

Using the trigonometric addition identities,

$$I_1(x, y) = I'(x, y) + I''(x, y, \Delta) \{ \cos[\phi(x, y)] \cos(\alpha) + \sin[\phi(x, y)] \sin \alpha \} \quad (14.43)$$

$$I_2(x, y) = I'(x, y) + I''(x, y, \Delta) \cos[\phi(x, y)] \quad (14.44)$$

and

$$I_3(x, y) = I'(x, y) + I''(x, y, \Delta) \{ \cos[\phi(x, y)] \cos(\alpha) - \sin[\phi(x, y)] \sin(\alpha) \} \quad (14.45)$$

These three equations can easily be solved for the unknown wavefront phase at each location:

$$\phi(x, y) = \tan^{-1} \left\{ \left[\frac{1 - \cos(\alpha)}{\sin(\alpha)} \right] \frac{I_1 - I_3}{2I_2 - I_1 - I_3} \right\} \quad (14.46)$$

The signal modulation can also be derived by first solving Eqs. (14.42)–(14.47) for the unknowns $I'(x, y)$ and $I''(x, y, \Delta)$:

$$I'(x, y) = \frac{I_1 + I_3 - 2I_2 \cos(\alpha)}{2[1 - \cos(\alpha)]} \quad (14.47)$$

and

$$I''(x, y, \Delta) = \frac{\{[1 - \cos(\alpha)(I_1 - I_3)]^2 + [\sin(\alpha)(2I_2 - I_1 - I_3)]^2\}^{1/2}}{2 \sin(\alpha)[1 - \cos(\alpha)]} \quad (14.48)$$

The ratio of these two intensities gives the recorded signal modulation at each point in the data set:

$$\begin{aligned} \gamma(x, y) &= \frac{I''(x, y, \Delta)}{I'(x, y)} \\ &= \frac{\{[1 - \cos(\alpha)(I_1 - I_3)]^2 + [\sin(\alpha)(2I_2 - I_1 - I_3)]^2\}^{1/2}}{[I_1 + I_3 - 2I_2 \cos(\alpha)] \sin(\alpha)} \end{aligned} \quad (14.49)$$

It is important to remember that the modulation being measured here is that of the recorded signal, not the intensity incident on the detector, and it includes the MTF reduction due to using the integrating bucket method and other terms such as fixed pattern noise from the detector. This statement also applies to the modulation previously calculated for the four step algorithm (Eq. 14.30). Calculation of the modulation at each point allows data points with insufficient modulation to be identified and excluded from the analysis.

Two commonly used phase step sizes with the three step algorithm are 90° and 120° . For these values of α , the solutions for the wavefront phase are, when $\alpha = \pi/2$:

$$\phi(x, y) = \tan^{-1} \left(\frac{I_1 - I_3}{2I_2 - I_1 - I_3} \right) \quad (14.50)$$

where the signal modulation is

$$\gamma(x, y) = \frac{[(I_1 - I_3)^2 + (2I_2 - I_1 - I_3)^2]^{1/2}}{I_1 + I_3} \quad (14.51)$$

and when $\alpha = 2\pi/3$

$$\phi(x, y) = \tan^{-1} \left(\sqrt{3} \frac{I_1 - I_3}{2I_2 - I_1 - I_3} \right) \quad (14.52)$$

and the corresponding signal modulation for this value is easily determined by using Eq. (14.49).

An interesting variation on the three step algorithm results from using 90° phase steps and a phase offset of 45° (Wyant et al., 1984; Bhushan et al., 1985). The phase offset is included to take advantage of a trigonometric identity and for computational convenience. It is irrelevant for the operation of the interferometer since we are only interested in the relative phases of the steps. For these

conditions, the phase steps are

$$\delta_i = \pi/4, 3\pi/4, 5\pi/4; \quad i = 1, 2, 3 \quad (14.53)$$

and the resulting wavefront phase is

$$\phi(x, y) = \tan^{-1} \left(\frac{I_3 - I_2}{I_1 - I_2} \right) \quad (14.54)$$

The result for the signal modulation is identical to the result we obtained in Eq. (14.51). This should not be surprising since the modulation of the signal is independent of the starting value of the reference phase ($-\pi/2$ vs. $\pi/4$). Similarly, the difference between analyzing the three intensity frames with Eq. (14.50) instead of Eq. (14.54) is a constant wavefront phase offset of $3\pi/4$. These later expressions are useful as they are analytically less complex, and can be used for any set of three measurements with $\pi/2$ phase shifts regardless of the starting phase. An unimportant phase offset or piston term will appear in the result of the calculation.

The three step algorithm requires the minimum amount of data and is the simplest to use. However, as we will see later, this algorithm is also the most sensitive to errors in the phase shift between frames.

14.7.2. Least-Squares Algorithms

As we have noted several times, the measured intensity at a given location in the interferogram varies as a sinusoidal function of the reference phase with a known period and three unknowns: the average signal, the data modulation and the unknown reference phase. It should not be surprising that the wavefront phase can be determined through a least-squares fit of the measured intensities to a sinusoidal function.

The general least-squares solution for a series of N interferograms recorded at phase shifts δ_i was first discussed by Bruning, Herriott, Gallagher, Rosenfeld, White and Brangaccio (1974) and more rigorously by Greivenkamp (1984). Eq. (14.37) is first rewritten as

$$I_i(x, y) = I'(x, y) + I''(x, y, \Delta) \cos[\phi(x, y)] \cos(\delta_i) - I''(x, y, \Delta) \sin[\phi(x, y)] \sin(\delta_i) \quad (14.55)$$

or

$$I_i(x, y) = \alpha_0(x, y) + \alpha_1(x, y) \cos(\delta_i) + \alpha_2(x, y) \sin(\delta_i) \quad (14.56)$$

where

$$\alpha_0(x, y) = I'(x, y) \quad (14.57)$$

$$\alpha_1(x, y) = I''(x, y, \Delta) \cos[\phi(x, y)] \quad (14.58)$$

and

$$\alpha_2(x, y) = -I''(x, y, \Delta) \sin[\phi(x, y)] \quad (14.59)$$

Eq. (14.56) is in the proper form for a generalized least-squares fit to the measured intensities $I_i(x, y)$ at each location. The squared difference between the measured intensities and the intensities predicted by Eq. (14.56) for a choice of the three unknowns or variables $\alpha_0(x, y)$, $\alpha_1(x, y)$ and $\alpha_2(x, y)$ is

$$E^2 = \sum_{i=1}^N [I_i(x, y) - \alpha_0(x, y) - \alpha_1(x, y) \cos(\delta_i) - \alpha_2(x, y) \sin(\delta_i)] \quad (14.60)$$

This error is minimized by differentiating with respect to each of the three unknowns and equating these three results to zero. The simultaneous solution of these three equations produces the least-squares result, and is given by the following matrix equation:

$$\begin{bmatrix} \alpha_0(x, y) \\ \alpha_1(x, y) \\ \alpha_2(x, y) \end{bmatrix} = \mathbf{A}^{-1}(\delta_i) \mathbf{B}(x, y, \delta_i) \quad (14.61)$$

where the component matrices are

$$\mathbf{A}(\delta_i) = \begin{bmatrix} N & \sum \cos(\delta_i) & \sum \sin(\delta_i) \\ \sum \cos(\delta_i) & \sum \cos^2(\delta_i) & \sum \cos(\delta_i) \sin(\delta_i) \\ \sum \sin(\delta_i) & \sum \cos(\delta_i) \sin(\delta_i) & \sum \sin^2(\delta_i) \end{bmatrix} \quad (14.62)$$

and

$$\mathbf{B}(x, y, \delta_i) = \begin{bmatrix} \sum I_i \\ \sum I_i \cos(\delta_i) \\ \sum I_i \sin(\delta_i) \end{bmatrix} \quad (14.63)$$

All of the summations are from 1 to N. The matrix $\mathbf{A}(\delta_i)$ is a function of the reference phase shifts only and incorporates the information about the data collection scheme, and therefore, only needs to be calculated and inverted once. The matrix $\mathbf{B}(x, y, \delta_i)$ is composed of weighted sums of the measured interferogram intensities.

Once the values of the three unknowns $\alpha_0(x, y)$, $\alpha_1(x, y)$ and $\alpha_2(x, y)$ are determined at each measurement location, the wavefront phase and the signal or data

modulation can be easily found from Eqs. (14.61)–(14.63):

$$\phi(x, y) = \tan^{-1} \left(\frac{-\alpha_2(x, y)}{\alpha_1(x, y)} \right) \quad (14.64)$$

and

$$\gamma(x, y) = \frac{I''(x, y, \Delta)}{I'(x, y)} = \frac{[\alpha_1(x, y)^2 + \alpha_2(x, y)^2]^{1/2}}{\alpha_0(x, y)} \quad (14.65)$$

One feature of PSI that this general algorithm illustrates is that any combination of three or more reference phase values can be used to measure the wavefront. These values need not to be evenly spaced and can be spread over a range greater than 2π . However, the choice of the phase shift positions does influence the system signal-to-noise performance and are specific to the $\mathbf{A}(\delta_i)$ matrix.

A specific algorithm that follows from the least-squares analysis is for N evenly spaced phase steps over one period of phase shift (Bruning et al., 1974; Morgan, 1982; Greivenkamp, 1984), or

$$\delta_i = i2\pi/N; \quad i = 1, \dots, N \quad (14.66)$$

In this special case, all of the off-diagonal terms in the matrix $\mathbf{A}(\delta_i)$ are zero, and the simple solution for the wavefront phase and the recorded data modulation are

$$\phi(x, y) = \tan^{-1} \left(\frac{-\sum I_i \sin(\delta_i)}{\sum I_i \cos(\delta_i)} \right) \quad (14.67)$$

and

$$\gamma(x, y) = \frac{2\{[\sum I_i \cos(\delta_i)]^2 + [\sum I_i \sin(\delta_i)]^2\}^{1/2}}{\sum I_i} \quad (14.68)$$

The similarity between these results and the four step algorithm in Eqs. (14.27) and (14.30) should be apparent.

All of the three- and four-step algorithms presented in the previous sections of this chapter are derived without the use of the least-squares criterion; they are analytic solutions. However, these algorithms are precisely equivalent to the least squares solutions and can be derived from Eqs. (14.61)–(14.66).

We will discuss additional families of algorithms later in this chapter, but wish to mention here that Eq. (14.68) is already the basis for a series of PSI-algorithms referred to as N -frame or N -bucket, where the 4-bucket is the most common one.

TABLE 14.1. Members of the N -bucket family of algorithms derived from Eq. (14.61).

No. of frames	Algorithm	Phase shift
3	$\tan \phi = \frac{\sqrt{3}(I_2 - I_3)}{-2I_1 + I_2 + I_3}$	$\delta_i = (i - 1) \frac{2\pi}{3}$
4	$\tan \phi = \frac{I_2 - I_4}{I_3 - I_1}$	$\delta_i = (i - 1) \frac{\pi}{2}$
6	$\tan \phi = \frac{-\sqrt{3}(I_2 + I_3 - I_5 - I_6)}{2I_1 + I_2 - I_3 - 2I_4 - I_5 + I_6}$	$\delta_i = (i - 1) \frac{\pi}{3}$

14.7.3. Carré Algorithm

In all of the PSI algorithms discussed so far, it is assumed that the phase shift between measurements is known. As we will see later, if the actual phase shift differs from the assumed value, errors are introduced into the reconstruction. An error of this sort could result from a change in the slope of the PZT displacement versus voltage curve, where all phase steps would be the same, but of incorrect size (i.e., 88° instead of 90°). Nonuniformities of the phase shift across the pupil can result when a linear phase shift in a converging or diverging beam (as discussed in Section 14.4) are another source of this type of error. Several PSI algorithms have been developed to minimize the effects of these phase shift variations.

The first of these algorithms is the Carré algorithm (Carré, 1966). This algorithm is a variation of the four-step algorithm, but instead of requiring that the data be collected at 90° increments, the reference phase shift between measurements is treated as an unknown and solved for in the analysis. A linear phase shift of 2α is assumed between each step, so that the values of the reference phase are

$$\delta_i = -3\alpha, -\alpha, \alpha, 3\alpha; \quad i = 1, 2, 3, 4 \quad (14.69)$$

As noted earlier, these four values are chosen to give simpler expression for the result. The four measured intensity frames here are represented by

$$I_1(x, y) = I'(x, y) + I''(x, y, \Delta) \cos[\phi(x, y) - 3\alpha] \quad (14.70)$$

$$I_2(x, y) = I'(x, y) + I''(x, y, \Delta) \cos[\phi(x, y) - \alpha] \quad (14.71)$$

$$I_3(x, y) = I'(x, y) + I''(x, y, \Delta) \cos[\phi(x, y) + \alpha] \quad (14.72)$$

and

$$I_4(x, y) = I'(x, y) + I''(x, y, \Delta) \cos[\phi(x, y) + 3\alpha] \quad (14.73)$$

This series of four equations contains four unknowns, the three we have been considering up to this point with the fourth, the unknown reference phase.

The solution for the reference phase shift can be found by expanding these four equations and applying the trigonometric identity for sine or cosine of 3α :

$$\alpha(x, y) = \tan^{-1} \left[\frac{3(I_2 - I_3) - (I_1 - I_4)}{(I_1 - I_4) + (I_2 - I_3)} \right]^{1/2} \quad (14.74)$$

Note that since this equation can be solved at each measurement point, the reference phase shift $2\alpha(x, y)$ can also be determined at each point. This allows spatial variations of the phase shift to be identified. The solution for the wavefront phase at each measurement location is

$$\phi(x, y) = \tan^{-1} \left\{ \tan[\alpha(x, y)] \frac{(I_1 - I_4) + (I_2 - I_3)}{(I_2 + I_3) - (I_1 + I_4)} \right\} \quad (14.75)$$

or combining these two results

$$\phi(x, y) = \tan^{-1} \left\{ \frac{\{[3(I_2 - I_3) - (I_1 - I_4)][(I_1 - I_4) + (I_2 - I_3)]\}^{1/2}}{(I_2 + I_3) - (I_1 + I_4)} \right\} \quad (14.76)$$

where the (x, y) dependence of the measured intensities is implied. The value of this algorithm is that it compensates for errors in the amount of phase shift as well as for spatial variations of the phase shift. However, this algorithm requires that the phase shift increments at a given location be equal.

The conversion of the result of the arctangent in the Carré algorithm to the wavefront phase modulo 2π (see Section 14.5) is not as straight forward as for the other algorithms discussed in this chapter. With these other algorithms, terms proportional to the sine and cosine of the wavefront phase $\phi(x, y)$ are contained in the numerator and denominator of the arctangent, and these values are used with Table 14.11 to correct the calculated phase. In the Carré algorithm (Eq. 14.76), the square root in the numerator produces the absolute value of the $\sin[\phi(x, y)]$, not the sine. In fact, since the denominator can be either positive or negative, the sign of the phase produced by this equation can be wrong. The conversion to the phase modulo 2π for the Carré algorithm should therefore be based upon the absolute value of this calculation, and the appropriate entries in the ‘‘Corrected Phase’’ column of Table 14.11 are $0, |\phi(x, y)|, \pi/2, \pi - |\phi(x, y)|, \pi, \pi + |\phi(x, y)|, 3\pi/2,$ and $2\pi - |\phi(x, y)|$. The entries that have changed their form are those where the sine and cosine are of opposite sign. In addition, to use this revised table, terms proportional to the sine and cosine of the wavefront phase must also be constructed in order to determine the signs and select the appropriate table entries (Creath, 1985). One such set of terms is

$$\begin{aligned} \sin[\phi(x, y)] &\propto (I_2 - I_3) \\ &= 2I''(x, y, \Delta) \sin[\alpha(x, y)] \sin[\phi(x, y)] \end{aligned} \quad (14.77)$$

and

$$\begin{aligned}\cos[\phi(x, y)] &\propto (I_2 + I_3) - (I_1 + I_4) \\ &= 8I''(x, y, \Delta) \sin^2[\alpha(x, y)] \cos[\alpha(x, y)] \cos[\phi(x, y)]\end{aligned}\quad (14.78)$$

With these values, the revised Table 14.11 can be used to calculate the wavefront phase modulo 2π for the Carré algorithm. It should also be noted that this revised correction table can be used with any of the other algorithms.

The most commonly used reference phase shift is 90° between measurements ($\alpha = 45^\circ$), and the recorded data modulation can then be calculated with Eq. (14.30), the result for the four step algorithm. (Once again, the modulation must be independent of the initial value of the phase shift.) This expression can be used even in the presence of fairly large variations in the size of the phase step as only small errors in the calculated modulation will result. For a $\pm 10\%$ change in the phase shift, the maximum error in the modulation is approximately $\pm 5\%$. This is quite tolerable as the modulation is generally only used to sort data points with insufficient modulation from the analysis.

14.7.4. Family of Averaging Algorithms

As we will see in Section 14.10, a linear phase shift error results in an sinusoidal error in the reconstruction of the wavefront phase that has a frequency twice that of the interferogram fringe frequency. Averaging two calculations of the wavefront phase with a 90° reference phase offset between them can reduce this error (Schwider et al., 1983; Wyant and Prettyjohns, 1987). This 90° phase shift will move the fringe pattern by a quarter period, and the reconstruction error will also be shifted by this same physical distance between the two calculations (Eq. (14.79)). Since this error occurs at twice the fringe frequency, the periodic error between the two measurements will be offset by half its period. The errors in the two calculations will therefore tend to cancel when they are averaged together.

$$\tan \phi = \frac{N}{D} = \frac{N_0 + N_{\pi/2}}{D_0 + D_{\pi/2}}\quad (14.79)$$

One implementation of this technique is to record one additional intensity measurements with a $\pi/2$ phase shift to the 3 or 4 required intensities. The first three or four of these frames are analyzed using the three-step algorithm (Equation 14.44) or 4-frame algorithms (Eq. (14.27)) to get the first measurement of the wavefront phase. The second measurements are now analyzed with the same equation, substituting I_2 for I_1 , I_3 for I_2 and so on to get the second calculation of the wavefront phase. These two results are then averaged at each measurement location to obtain a result that is significantly less susceptible to phase shift errors that may exist in either of the two individual calculations.

For example Eq. (14.79) can also be used to derive new implicit error compensating algorithms from a 3 or 4 frame base algorithm (Schwider et al., 1993). If the 3-frame base algorithm is used, the first derived algorithm was shown by

$$\tan\left(\phi + \frac{\pi}{4}\right) = \frac{2(I_3 - I_2)}{I_1 - I_2 - I_3 + I_4} \quad (14.80)$$

and if the 4-frame base algorithm is used, the first derived algorithm is shown (Schwider et al., 1983) to be

$$\tan(\phi) = \frac{2(I_4 - I_2)}{I_1 - 2I_3 + I_5} \quad (14.81)$$

Schmit and Creath showed 1995 that the averaging technique can be extended to obtain further compensations. Using this “extended averaging” technique, a $N + 2$ bucket version can be obtained, meaning 2 additional buckets are needed relative to the base algorithm:

$$\tan \phi = \frac{N}{D} = \frac{(N_0 + N_{\pi/2}) + (N_0 + N_{\pi/2})_{\pi/2}}{(D_0 + D_{\pi/2}) + (D_0 + D_{\pi/2})_{\pi/2}} = \frac{N_0 + 2N_{\pi/2} + N_{\pi}}{D_0 + 2D_{\pi/2} + D_{\pi}} \quad (14.82)$$

In 1996, Schmit and Creath showed there is also a possibility to mix-and-match different algorithms and to average them. This multiaveraging technique can be used to derive several algorithms with optimized sampling window functions of the intensities:

$$\tan \phi = \frac{N}{D} = \frac{\sum_i N_i}{\sum_i D_i} \quad (14.83)$$

We summarize the various algorithms discussed above and others in Table 14.2 by separating them into two classes. Class A is derived from the 4-frame algorithm using a 90° phase step (Eq. (14.27)). Group B is derived using the 3-step base algorithm (Eq. (14.52)).

14.7.5. Hariharan Algorithm

Another approach to producing the identical PSI algorithm to Eq. (14.81) that is insensitive to reference phase shift calibration errors is provided by Hariharan et al. (1987). They use five measurements of the interferogram intensity, and initially assume a linear phase shift of α between frames:

$$\delta_i = -2\alpha, -\alpha, 0, \alpha, 2\alpha; \quad i = 1, 2, 3, 4, 5 \quad (14.84)$$

TABLE 14.2. Algorithms derived using a base algorithm and the averaging technique.

Frames	Class A: $\tan \phi =$	Class B: $\tan\left(\phi + \frac{\pi}{4}\right) =$	Phase shift
base	$\frac{I_2 - I_4}{I_3 - I_1}$	$\frac{I_3 - I_2}{I_1 - I_2}$	$\delta_i = (i - 1) \frac{\pi}{2}$
4		$\frac{2(I_3 - I_2)}{I_1 - I_2 - I_3 + I_4}$	$\delta_i = (i - 1) \frac{\pi}{2}$
5	$\frac{2(I_4 - I_2)}{I_1 - 2I_3 + I_5}$	$\frac{3(I_3 - I_2) + I_4 - I_5}{I_1 - I_2 - 3I_3 + 3I_4}$	$\delta_i = (i - 1) \frac{\pi}{2}$
6	$\frac{4I_4 - 3I_2 - I_6}{I_1 - 4I_3 + 3I_5}$	$\frac{4(I_3 - I_2 + I_4 - I_5)}{I_1 - I_2 - 6I_3 + 6I_4 + I_5 - I_6}$	$\delta_i = (i - 1) \frac{\pi}{2}$
7	$\frac{4(2I_4 - I_2 - I_6)}{I_1 - 7I_3 + 7I_5 - I_7}$	$\frac{5I_3 - 5I_2 + 10I_4 - 10I_5 - I_6 + I_7}{I_1 - I_2 - 10I_3 + 10I_4 + 5I_5 - 5I_6}$	$\delta_i = (i - 1) \frac{\pi}{2}$
8	$\frac{15I_4 - 5I_2 - 11I_6 - I_8}{I_1 - 11I_3 + 15I_5 - 5I_7}$	$\frac{-2(3I_2 - 3I_3 - 10I_4 + 10I_5 + 3I_6 - 3I_7)}{I_1 - I_2 - 15(I_3 - 15I_4 - 15I_5 + 15I_6) - I_7 + I_8}$	$\delta_i = (i - 1) \frac{\pi}{2}$

Then,

$$I_1(x, y) = I'(x, y) + I''(x, y, \Delta) \cos[\phi(x, y) - 2\alpha] \tag{14.85}$$

$$I_2(x, y) = I'(x, y) + I''(x, y, \Delta) \cos[\phi(x, y) - \alpha] \tag{14.86}$$

$$I_3(x, y) = I'(x, y) + I''(x, y, \Delta) \cos[\phi(x, y)] \tag{14.87}$$

$$I_4(x, y) = I'(x, y) + I''(x, y, \Delta) \cos[\phi(x, y) + \alpha] \tag{14.88}$$

and

$$I_5(x, y) = I'(x, y) + I''(x, y, \Delta) \cos[\phi(x, y) + 2\alpha] \tag{14.89}$$

These five equations are expanded and combined to produce the intermediate result:

$$\frac{\tan[\phi(x, y)]}{2 \sin(\alpha)} = \frac{I_2 - I_4}{2I_3 - I_5 - I_1} \tag{14.90}$$

Since the choice of the phase shift α is open, we can choose it to minimize the variation of this expression to errors in the phase shift. This function is plotted in Figure 14.14 for $\phi(x, y) = \pi/4$. Differentiating this equation with respect to α , we find

$$\frac{d}{d\alpha} \left\{ \frac{\tan[\phi(x, y)]}{2 \sin(\alpha)} \right\} = \frac{-\cos(\alpha) \tan[\phi(x, y)]}{2 \sin^2(\alpha)} \tag{14.91}$$

which goes to zero when $\alpha = \pi/2$. When using this value for the phase shift, Eq. (14.90) becomes insensitive to phase shift calibration errors, and Eq. (14.90) reduces

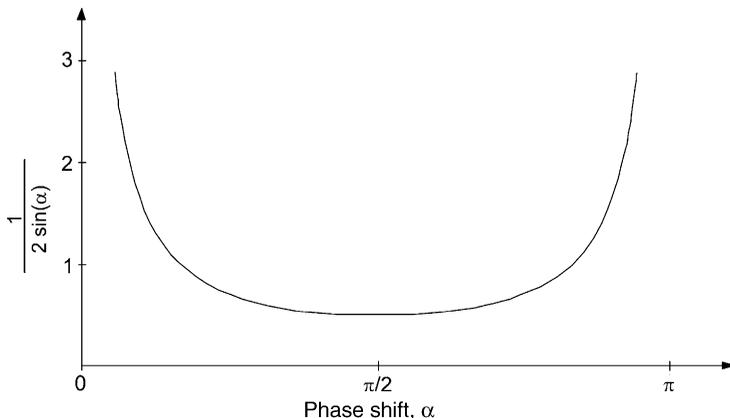


FIGURE 14.14. A plot of Eq. (14.90) as a function of the value of the phase shift.

to the final algorithm,

$$\phi(x, y) = \tan^{-1} \left[\frac{2(I_2 - I_4)}{2I_3 - I_5 - I_1} \right] \tag{14.92}$$

The data modulation at each measurement point can also be determined:

$$\gamma(x, y) = \frac{3[4(I_4 - I_2)^2 + (I_1 + I_5 - 2I_3)^2]^{1/2}}{2(I_1 + I_2 + 2I_3 + I_4 + I_5)} \tag{14.93}$$

A certain similarity should be noted between the form of these two results and the results for the four step algorithm (Eq. (14.27) and (14.30)), which also uses 90° phase shifts. Since I_1 and I_5 are nominally identical, the appearance of both of these terms is offset by an extra I_3 , which is the term 180° out of phase.

The minimum in the plot of Eq. (14.90) (Figure 14.14) is centered at 90° and is quite broad. As a result, the Hariharan algorithm can tolerate fairly large errors in the phase shift before significant errors appear in the calculated wavefront phase (Hariharan et al. 1987). If the actual phase shift between measurements is $\pi/2 + \epsilon$, then there is a corresponding measured phase $\phi'(x, y) = \phi(x, y) + \Delta\phi(x, y)$ which can be approximated using Eq. (14.90) and assuming that ϵ is small:

$$\tan[\phi'(x, y)] \cong [1 + (\epsilon^2/2)] \tan[\phi(x, y)] \tag{14.94}$$

The error in the measurement of the phase can now also be easily determined:

$$\Delta\phi(x, y) = \phi'(x, y) - \phi(x, y) \cong (\epsilon^2/4) \sin[2\phi(x, y)] \tag{14.95}$$

The error is a function of the wavefront phase, and the maximum error for a 2° change in the linear phase shift between measurements (88° instead of 90°) is about 0.02°.

This same phase shift error for one of the standard algorithms, such as the three-step algorithm, results in a measurement error on the order of 1° .

14.7.6. 2 + 1 Algorithm

PSI algorithms require a series of intensity measurements to calculate the wavefront phase. Since a video camera is normally used to record the intensity patterns, the optical setup must remain stable and air turbulence must be minimized during its image capture time. Failure to attain these conditions may result in measurement errors or even a data set that cannot be analyzed. In many situations, such as the testing of large optics, this requirement can be difficult or impractical.

The 2 + 1 algorithm (Angel and Wizinowich, 1988; Wizinowich, 1989, 1990) attacks the problem of PSI testing in the presence of vibration by first rapidly collecting two time-critical interferograms with a 90° phase shift to monitor the intensity modulation, and later recording a third interferogram that gives the average intensity across the field. This third interferogram is the average of two interferograms with a 180° phase shift; the peaks of one set of fringes fall in the troughs of the other set, canceling out the fringe pattern. With this system, the phase shifts are

$$\delta_i = 0, -\pi/2, 0 \text{ and } \pi; \quad i = 1, 2, 3 \quad (14.96)$$

and the three recorded interferograms are

$$I_1(x, y) = I'(x, y) + I''(x, y) \cos[\phi(x, y)] \quad (14.97)$$

$$\begin{aligned} I_2(x, y) &= I'(x, y) + I''(x, y) \cos[\phi(x, y) - \pi/2] \\ &= I'(x, y) + I''(x, y) \sin[\phi(x, y)] \end{aligned} \quad (14.98)$$

and

$$\begin{aligned} I_3(x, y) &= \frac{1}{2} \{I'(x, y) + I''(x, y) \cos[\phi(x, y)]\} \\ &\quad + \frac{1}{2} \{I'(x, y) + I''(x, y) \cos[\phi(x, y) + \pi]\} \end{aligned} \quad (14.99)$$

These three equations can be easily solved for the wavefront phase and the recorded data modulation at each measurement point:

$$\phi(x, y) = \tan^{-1} \left(\frac{I_2 - I_3}{I_1 - I_3} \right) \quad (14.100)$$

and

$$\gamma(x, y) = \frac{\left[(I_2 - I_3)^2 + (I_1 - I_3)^2 \right]^{1/2}}{I_3} \quad (14.101)$$

Perhaps more interesting than the algorithm itself is the hardware needed to implement the $2 + 1$ algorithm. This system is described in detail by Wizinowich (1990), and two of the key design features are reviewed here. The first requirement is that the two time critical interferograms (I_1 and I_2) be recorded with as small a time lag between them as possible. A common solid state sensor architecture is an interline-transfer charge coupled device. Each photosite on the sensor is accompanied by an adjacent storage pixel. The storage pixels are read out to produce the video signal while the active photosites are integrating the light for the next video field. At the end of this integration period, the charge collected in the active pixels is quickly transferred to the now empty storage sites, and the next video field is collected. This transfer takes place in about a microsecond. It is possible with a synchronized shutter and this sensor to record two exposures of about a millisecond in a duration separated by a microsecond; the first is recorded just before the transfer, and the second is recorded just after the transfer. These two recorded interferograms are read out at standard video rates and can be digitized.

The second system requirement is that the appropriate phase shift is applied to each of these two exposures. The system that has been used is diagrammed in Figure 14.15. The rotating turntable with the two right-angle prisms introduces equal and opposite Doppler frequency shifts of $\pm\Delta\nu$ into the upper and lower beams. Because of the quarterwave plates, the recombined beams are also orthogonally polarized. The Pockels cell selects one or the other or both of the polarizations to enter the interferometer, and the shutter provides the appropriate exposure. All of these devices must be synchronized with the detector array. The

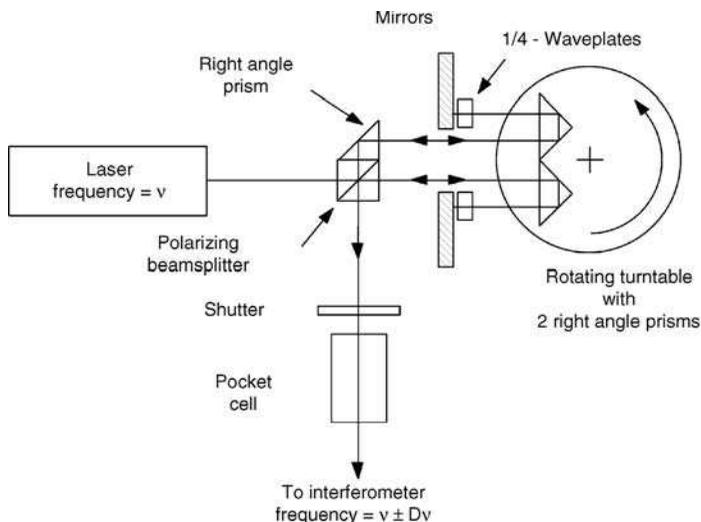


FIGURE 14.15. The system used to generate the phase shift required for the $2 + 1$ algorithm (From Wizinowich, 1990).

interferometer used with system is non-polarizing so that any input polarization will enter both the reference and test arms of the interferometer (this is not the situation described in Section 14.4, Eq. (14.12)). The individual interferograms are, therefore, produced by just one of the two different optical wavelengths or frequencies, and the relative phase shift α between interferograms recorded with the two frequencies is related to the optical path difference OPD between the two arms of the interferometer:

$$\alpha = 2\pi \text{ OPD } \Delta\nu/c \quad (14.102)$$

where c is the speed of light. This phase shift is equal to the difference in the number of optical cycles that fit into the OPD for the two optical frequencies. The frequency shift of the two beams $\Delta\nu$ is a function of the turntable rotation speed and its size. For path length differences of 10 m, a small frequency shift of about one part in 10^8 is needed. The turntable rotation rate is chosen to produce the desired phase shift.

The system operation is as follows. The phase shift is set for 90° , and the two time-dependent interferograms are recorded on each side of the interline transfer of the CCD. The Pockels cell is switched between these exposures, so that only one optical frequency is used for each. After these two interferograms are digitized and stored, the turntable speed is changed for a 180° phase shift. At this point, the Pockels cell is set to allow both frequencies into the interferometer, and the third or average interferogram is recorded (both interferograms are imaged on the detector). This system has been shown to freeze out the effects of vibration and allow for PSI analysis.

Another approach to PSI testing in the presence of vibration is to speed up one of the standard algorithms through the use of a high-speed video cameras (Greivenkamp, 1987a). The speed of the phase shifter must also be increased, but this is not a problem particularly with the integrating bucket approach where stop-start motions are not required. It should be noted, however, that speeding up the frame rate does not guarantee freedom from errors.

14.7.7. Methods to Generate Algorithms

Two generic methods to develop PSI algorithms have already been discussed earlier in this chapter: The original N -bucket technique is described by Bruning et al. (1974), and the averaging technique by Schwider et al. (1983), as well as the extended averaging techniques by Schmit and Creath (1995 and 1996), respectively.

Larkin and Oreb (1992) and later Surrel (1993, 1996,1998) and Hibino et al. (1995, 1997,1998) developed a sophisticated theory and a general approach to algorithm design. While similar, they differ in the way the algorithms are derived, the number of samples that are needed, and the algorithmic sensitivities to various error sources. The first family of algorithms to classify here is the so called $N + 1$ bucket class of symmetrical phase shift algorithms that comes in two types, A and B.

Type A (Eq. (14.103)) and B (E. (14.104)) were first introduced by Larkin and Oreb (1992) and later refined by Surrel (1993).

$$\tan \phi = \frac{\sum_{i=2}^N I_i \sin\left(i \frac{2\pi}{N}\right)}{\frac{I_1 + I_{N+1}}{2} + \sum_{i=2}^N I_i \cos\left(i \frac{2\pi}{N}\right)} \tag{14.103}$$

$$\tan \phi = \frac{\frac{I_{N+1} - I_1}{2} \cot\left(\frac{2\pi}{N}\right) - \sum_{i=2}^N I_i \sin\left(i \frac{2\pi}{N}\right)}{\frac{I_1 + I_{N+1}}{2} + \sum_{i=2}^N I_i \cos\left(i \frac{2\pi}{N}\right)} \tag{14.104}$$

The difference between the two algorithms is the different weighting of the first and last frame. Type B is in general correcting for phase shift errors.

Table 14.3 shows the 4,5 and 7 frame versions of the $N + 1$ bucket algorithm. Note that the 5 frame version is common to both the Type A and B, and is based on the fact that the cotangent function in Eq. (14.98) becomes zero for $N = 4$, resulting in the $N + 1 = 5$ frame algorithms. This algorithm is also identical to the 5 bucket algorithms by Schwider et al. (1983) (Eq. (14.81)) and Hariharan et al. (1987b) (Eq. (14.92)) and is reducing errors introduced by the phase shifting.

A more refined approach called the windowed DFT algorithm (WDFT) was presented by Surrel in 1996. Algorithms derived from Eq. (14.105) are insensitive to harmonics up to the order of $j = N - 2$ even in the presence of a phase-shift miscalibration.

$$\tan \phi = \frac{-\sum_{i=1}^{N-1} i(I_i - I_{2N-i}) \sin\left(i \frac{2\pi}{N}\right)}{NI_N - \sum_{i=1}^{N-1} i(I_i + I_{2N-i}) \cos\left(i \frac{2\pi}{N}\right)} \tag{14.105}$$

TABLE 14.3. $N + 1$ bucket.

Frames	Type A algorithm $\tan \phi =$	Type B algorithm $\tan \phi =$	Phase shift
4	$\frac{\sqrt{3}(I_2 - I_3)}{-I_1 + I_2 + I_3 - I_4}$	$\frac{I_1 - 3I_2 + 3I_3 - I_4}{\sqrt{3}(I_1 - I_2 - I_3 + I_4)}$	$\delta_i = (i - 1) \frac{2\pi}{3}$
5	$\frac{2(I_4 - I_2)}{I_1 - 2I_3 + I_5}$		$\delta_i = (i - 1) \frac{\pi}{2}$
7	$\frac{-\sqrt{3}(I_2 + I_3 - I_5 - I_6)}{I_1 + I_2 - I_3 - 2I_4 - I_5 + I_6 + I_7}$	$\frac{-(I_1 + 3I_2 + 3I_3 - 3I_5 - 3I_6 - I_7)}{\sqrt{3}(I_1 + I_2 - I_3 - 2I_4 - I_5 + I_6 + I_7)}$	$\delta_i = (i - 1) \frac{\pi}{3}$

TABLE 14.4. First three algorithms of the WDFT family of algorithms.

Frames	Algorithm	Phase shift
5	$\tan \phi = \frac{\sqrt{3}(I_1 - 2I_2 + 2I_4 - I_5)}{I_1 + 2I_2 - 6I_3 + 2I_4 + I_5}$	$\delta_i = i \frac{2\pi}{3}$
7	$\tan \phi = \frac{I_1 - 3I_3 + 3I_5 - I_7}{2(I_2 - 2I_4 + I_6)}$	$\delta_i = i \frac{\pi}{2}$
11	$\tan \phi = \frac{-\sqrt{3}(I_1 + 2I_2 - 4I_4 - 5I_5 + 5I_7 + 4I_8 - 2I_{10} - I_{11})}{I_1 - 2I_2 - 6I_3 - 4I_4 + 5I_5 + 12I_6 + 5I_7 - 4I_8 - 6I_9 - 2I_{10} + I_{11}}$	$\delta_i = i \frac{\pi}{3}$

Table 14.4 shows the first 3 algorithms derived using the WDFT technique. Note the $\pi/2$ step is now at 7 frames, illustrating that for more highly developed algorithms, more frames beyond 2π are needed.

In 1997 Hibino et al. published an even more refined algorithm for nonlinear and spatially non-uniform phase shifts, which they called the $N + 3$ algorithms (Eq. (14.106)).

$$\tan \phi = \frac{\frac{1}{4}(I_0 + I_1 - I_{N+2} - I_{N+3}) \frac{\sin\left(\frac{3\pi}{N+2}\right)}{\sin^2\left(\frac{2\pi}{N+2}\right)} + \sum_{i=1}^{N+2} I_i \sin\left(\frac{2\pi}{N+2}\right) \left(i - \frac{N+5}{2}\right)}{\frac{1}{4}(I_0 - I_1 - I_{N+2} + I_{N+3}) \frac{\cos\left(\frac{3\pi}{N+2}\right)}{\sin^2\left(\frac{2\pi}{N+2}\right)} + \sum_{i=1}^{N+2} I_i \cos\left(\frac{2\pi}{N+2}\right) \left(i - \frac{N+5}{2}\right)}$$

(14.106)

Table 14.5 summarizes the first 3 $N + 1$ bucket algorithms for 5,6, and 8 frames, respectively.

Table 14.6 gives an overview of all the families discussed in this chapter, including the required phase-steps, their name and the major reference introducing the family of algorithms.

TABLE 14.5. $N + 3$ bucket algorithms.

No. of Frames	Algorithm	Phase shift
5	$\tan \phi = \frac{3\sqrt{3}(I_4 - I_2)}{2I_1 + I_2 - 6I_3 + I_4 + 2I_5}$	$\delta_i = (i - 3) \frac{2\pi}{3}$
6	$\tan \phi = \frac{I_1 - 3I_2 - 4I_3 + 4I_4 + 3I_5 - I_6}{I_1 + 3I_2 - 4I_3 - 4I_4 + 3I_5 + I_6}$	$\delta_i = \left(i - \frac{7}{2}\right) \frac{\pi}{2}$
8	$\tan \phi = \frac{2I_1 - I_2 - 6I_3 - 3I_4 + 3I_5 + 6I_6 + I_7 - 2I_8}{3\sqrt{3}(I_2 - I_4 - I_5 + I_7)}$	$\delta_i = \left(i - \frac{9}{2}\right) \frac{\pi}{3}$

TABLE 14.6. Overview of all introduced family of algorithms (note the increase of complexity for the generic equation).

Family of algorithms: $\tan \phi =$		Phase-steps	Reference
$\frac{-\sum_{i=1}^N I_i \sin(i \frac{2\pi}{N})}{\sum_{i=1}^N I_i \cos(i \frac{2\pi}{N})}$	$\delta_i = (i-1) \frac{2\pi}{N}$	N Bucket DFT	Bruning, 1974
$\frac{N}{D} = \frac{N_0 + N_{\pi/2}}{D_0 + D_{\pi/2}}$	$\delta_i = (i-1) \frac{2\pi}{N}$	Averaging $N+1$ technique	Schwider et al., 1983
$\frac{N}{D} = \frac{(N_0 + N_{\pi/2}) + (N_0 + N_{\pi/2})_{\pi/2}}{(D_0 + D_{\pi/2}) + (D_0 + D_{\pi/2})_{\pi/2}} = \frac{N_0 + 2N_{\pi/2} + N_{\pi}}{D_0 + 2D_{\pi/2} + D_{\pi}}$	$\delta_i = (i-1) \frac{2\pi}{N}$	Extended averaging $N+2$ technique	Schmit and Creath, 1995
$\frac{N}{D} = \frac{\sum_i N_i}{\sum_i D_i}$	$\delta_i = (i-1) \frac{2\pi}{N}$	Multiple averaging $N+m$ technique	Schmit and Creath, 1996
$\frac{\sum_{i=2}^N I_i \sin(i \frac{2\pi}{N})}{\frac{I_1 + I_{N+1}}{2} + \sum_{i=2}^N I_i \cos(i \frac{2\pi}{N})}$	$\delta_i = (i-1) \frac{2\pi}{N}$	$N+1$ bucket type A	Larkin and Oreb, 1992
$\frac{\frac{I_{N+1} - I_1}{2} \cot(\frac{2\pi}{N}) - \sum_{i=2}^N I_i \sin(i \frac{2\pi}{N})}{\frac{I_1 + I_{N+1}}{2} + \sum_{i=2}^N I_i \cos(i \frac{2\pi}{N})}$	$\delta_i = (i-1) \frac{2\pi}{N}$	$N+1$ bucket type B	Larkin and Oreb, 1992, Surrel, 1993
$\frac{-\sum_{i=1}^{N-1} i(I_i - I_{2N-i}) \sin(i \frac{2\pi}{N})}{NI_N - \sum_{i=1}^{N-1} i(I_i + I_{2N-i}) \cos(i \frac{2\pi}{N})}$	$\delta_i = i \frac{2\pi}{N}$	WDFT	Surrel, 1996
$\frac{\frac{1}{4}(I_0 + I_1 - I_{N+2} - I_{N+3}) \frac{\sin(\frac{3\pi}{N+2})}{\sin^2(\frac{2\pi}{N+2})} + \sum_{i=1}^{N+2} I_i \sin(\frac{2\pi}{N+2})(i - \frac{N+5}{2})}{\frac{1}{4}(I_0 - I_1 - I_{N+2} + I_{N+3}) \frac{\cos(\frac{3\pi}{N+2})}{\sin^2(\frac{2\pi}{N+2})} + \sum_{i=1}^{N+2} I_i \cos(\frac{2\pi}{N+2})(i - \frac{N+5}{2})}$	$\delta_i = (i-1) \frac{2\pi}{N}$	$N+3$ bucket	Hibino, 1997

14.7.8. Methods to Evaluate Algorithms

Along with the algorithm development, the theory, associated with evaluating the accuracy, speed, error-sources, and appropriateness for use in various applications, has been developed.

One of the most powerful analytical techniques was proposed by Freischlad and Koliopoulos (1990) and looks at phase evaluation like a filtering process in the frequency domain.

Fourier Description. In general, phase evaluation can be treated as correlations of the real signal $I(\phi)$ with two real functions f_N, f_D that describe the phase shift algorithms. This results in the correlation functions c_N, c_D that can be defined as

$$c_i(t) = \int_{-\infty}^{\infty} I(\phi)f_i(\phi + t)d\phi \quad i = N, D \quad (14.107)$$

In the special case of digital detection with the intensity being digitized at M discrete points this defaults to

$$c_i(t) = \sum_{n=1}^M I(\phi)f_i(\phi + t) \quad (14.108)$$

Sampling is incorporated into the filter functions through the use of a sampling comb:

$$f_x(t) = \sum_{l=0}^M a_{x,l}\delta(t - t_l) \quad (14.109)$$

where a_l is the real coefficient of the l th discrete sample and $t_l = lT_f/M$ represents equally spaced sampling positions.

Now if we apply this to a specific algorithm, for example for the 5 bucket algorithm presented by Schwider et al. (1983), we get

$$\begin{aligned} f_N(\phi) &= 2\left(\delta\left(\phi + \frac{3\pi}{2}\right) - \delta\left(\phi + \frac{\pi}{2}\right)\right) \\ f_D(\phi) &= \delta(\phi) - 2\delta(\phi + \pi) + \delta(\phi + 2\pi) \end{aligned} \quad (14.110)$$

The spectra of these filter functions can be easy calculated using the discrete Fourier transform

$$F_x(v) = \frac{1}{\sqrt{M}} \sum_{r=1}^M f_x(\phi)e^{i\phi(v/v_f)} \quad (14.111)$$

which results in the two spectra \mathcal{F}_x

$$\begin{aligned} \mathcal{F}_N(v) &= -4i \sin\left(\frac{\pi v}{2 v_f}\right) e^{-i\pi(v/v_f)} \\ \mathcal{F}_D(v) &= 2\left(\cos\left(\frac{v}{v_f}\pi\right) - 1\right) e^{-i\pi(v/v_f)} \end{aligned} \tag{14.112}$$

These spectra are plotted in Figure 14.16(b). Note that both spectra have identical amplitude at the fundamental frequency. In a similar way the three-step algorithm yields the following filter functions as shown in Figure 14.16(a),

$$\begin{aligned} \mathcal{F}_N(v) &= 2\sqrt{3}i \sin\left(v\frac{\pi}{3}\right) e^{-iv\pi} \\ \mathcal{F}_D(v) &= -2 + 2e^{-iv\pi} \cos\left(v\frac{\pi}{3}\right) \end{aligned} \tag{14.113}$$

Spectra for the Larkin and Oreb (1992) algorithms are shown in Figure 14.16c

$$\begin{aligned} \mathcal{F}_N(v) &= -2i\sqrt{3}\left(\sin\left(v\frac{\pi}{3}\right) + \sin\left(v\frac{2\pi}{3}\right)\right) e^{-iv\pi} \\ \mathcal{F}_D(v) &= -4\left(1 + 2\cos\left(v\frac{\pi}{3}\right)\right) \sin^2\left(v\frac{\pi}{3}\right) e^{-iv\pi} \end{aligned} \tag{14.114}$$

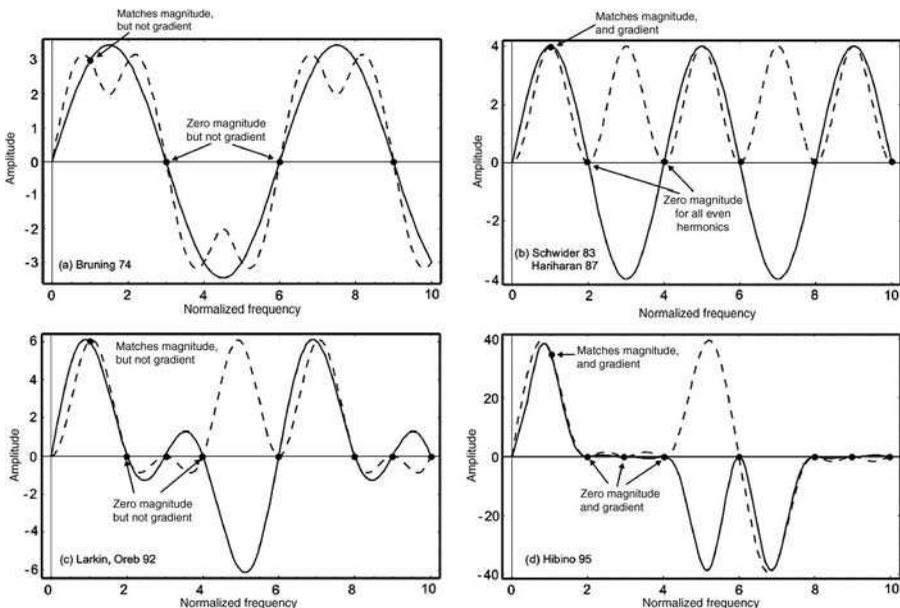


FIGURE 14.16. Spectra of the sampling functions for various algorithms.

For the more developed algorithm by Hibino, (1995) the spectral components are given below and in Figure 14.16(d),

$$\begin{aligned} \mathcal{F}_N(v) &= 32\sqrt{3} \cos^3\left(v\frac{\pi}{6}\right) \sin^2\left(v\frac{\pi}{2}\right) e^{-(3/2)v\pi} \\ \mathcal{F}_D(v) &= 16i \sin\left(v\frac{\pi}{6}\right) \left(\cos\left(v\frac{\pi}{3}\right) - 2\right) \left(2\cos\left(v\frac{\pi}{6}\right) + \cos\left(v\frac{\pi}{2}\right)\right)^2 \end{aligned} \quad (14.115)$$

To determine the phase of the input signal with respect to the filter functions, the ratio of the two filter functions is defined as

$$r = \frac{c_N}{c_D} = \frac{a_0 F_N(0) + 2\operatorname{Re}\left[\sum_{n=1}^{\infty} a_n F_N^*(nv_s)\right]}{a_0 F_D(0) + 2\operatorname{Re}\left[\sum_{n=1}^{\infty} a_n F_D^*(nv_s)\right]} \quad (14.116)$$

In order to determine the phase ϕ_m of the m th harmonic of the signal, one requires the following conditions be fulfilled:

$$\begin{aligned} a_n F_N^*(n \cdot v_g) &= -iA\delta(n, m) \\ a_n F_D^*(n \cdot v_g) &= A\delta(n, m) \end{aligned} \quad (14.117)$$

This lead to the conclusion that components a_n at $v = nv_g$ should be 90° out of phase and of equal magnitude:

$$F_N(nv_g) = iF_D(nv_g) \quad (14.118)$$

As Freischlad et al. (1990) have shown, the phase is then equal to the arctangent of the ratio c_N/c_D , independent of the amplitude $|s_m|$ and also independent of all other frequency components of the signal with a constant offset.

$$\phi = \tan^{-1}(r) + \text{const} \quad (14.119)$$

For a given fundamental frequency, v_g , Eq. (14.118) can only be satisfied at certain signal frequencies. Further, as shown by Freischlad (1990) and Larkin and Oreb (1992), the following properties can be derived from the Fourier description:

1. An algorithm is insensitive to the harmonic m , when the Fourier transforms for the sampling functions for both nominator and denominator match, that is,

$$F_N\left(\frac{v}{v_f}\right) = F_D\left(\frac{v}{v_f}\right) = 0.$$

2. If the slope of the Fourier transforms for harmonic m are equal, that is,

$$\left(\frac{d}{dv} F_N\right](v_f) \Big|_{v=v_f} = i \left(\frac{d}{dv} F_D\right](v_f) \Big|_{v=v_f}, \text{ the algorithms is insensitive to misca-}$$

libration at that harmonic.

3. Insensitivity to phase-shift errors and harmonic k are achieved when the k th derivative of the Fourier transform at the fundamental frequency are equal (condition 1) and the k -derivative at the frequency of the ($k \geq 2$) harmonic are zero as well (condition 2).

Characteristic Polynomial. A very elegant analytical method has been proposed by Surré (1996). The method uses a characteristic polynomial to describe an algorithm and is loosely related to the Fourier description presented before. The method assures that the intensity of the recorded interference signal $I(\phi)$ is a periodic function and therefore can be expanded into a Fourier series

$$I(\phi + \delta) = \sum_{m=-\infty}^{\infty} \beta_m(\phi) \exp(im\delta) \tag{14.120}$$

The various phase shift algorithms are then used to evaluate the argument of the Fourier coefficient β_1 which corresponds to the fundamental harmonic of the intensity signal. Writing the algorithms as the arctangent of the ratio of two linear combinations of values $I(\phi + \delta)$ yields

$$\phi = \tan^{-1} \left[\frac{\sum_{n=0}^N b_n I(\phi + n\delta)}{\sum_{n=0}^N a_n I(\phi + n\delta)} \right] \tag{14.121}$$

Substituting the exponential term y , the complex variable z of the characteristic polynomial, and c_n as its corresponding coefficient to the n th order

$$\phi = \arg \left[\sum_{n=1}^N (a_n + ib_n) I(\phi + n\delta) \right] = \arg \left[\sum_{m=-\infty}^{\infty} \beta_m(\phi) \underbrace{\sum_{n=1}^N \overbrace{(a_n + ib_n)}^{c_n} [\exp(im\delta)]^n}_{P(z)} \right] \tag{14.122}$$

results in the very compact form of the characteristic polynomial:

$$P(z) = \sum_{n=1}^N c_n z^{n-1} \tag{14.123}$$

The polynomial coefficients, c_n , thereby correspond to the specific coefficients of the algorithm, with a_n as the coefficients of the numerator or cosine part and b_n as the

coefficients of the denominator or sine part. For example, for the Bruning (1974) three-step algorithm using 120° steps, the characteristic polynomial yield:

$$P_{\text{Bruning } 74}(z) = -2 + z + i\sqrt{3}z + z^2 - i\sqrt{3}z^2 \quad (14.124)$$

One can draw conclusions about the behavior of an algorithm from the characteristic polynomial simply by looking at the roots. For the algorithm from Bruning (1974), there are only 2 single roots at

$$z = \left\{ 1, -\frac{1}{2} - i\frac{\sqrt{3}}{2} \right\} \quad (14.125)$$

The single root at $z = 1$ allows the algorithm to detect the fundamental frequency, but the missing existence of a double root demonstrates that the algorithm does not have the ability to compensate for miscalibration. In comparison, Schwider's (1993) algorithm has one double root, allowing it to be insensitive to phase-shifter miscalibration

$$P_{\text{Schwider } 93}(z) = (-1 - i) - (1 - 3i)z + (3 - i)z^2 - (1 + i)z^3 \quad (14.126)$$

$$z_{\text{single roots}} = \{1\} \quad z_{\text{double roots}} = \{-i\}$$

Schwider's algorithm, presented earlier in 1983, has a single root at both plus and minus m th-harmonic, which makes it insensitive to the m th-harmonic.

$$P_{\text{Schwider } 83}(z) = 1 - 2iz - 2z^2 + 2iz^3 + z^4 \quad (14.127)$$

$$z_{\text{single roots}} = \{-1, 1\} \quad z_{\text{double roots}} = \{-1 - i\}$$

Hibino's algorithm has only a single root for the first harmonic, but double roots for many of the following higher harmonics, making it insensitive to these higher harmonics:

$$P_{\text{Hibino } 95}(z) = 1 - i\sqrt{3} - z - 3i\sqrt{3}z - 7z^2 - 3i\sqrt{3}z^2 - 11z^3 + i\sqrt{3}z^3 - 6z^4 + 6i\sqrt{4}z^4 + 6z^5$$

$$+ 6i\sqrt{3}z^5 + 11z^6 + i\sqrt{3}z^6 + 7z^7 - 3i\sqrt{3}z^7 + z^8 - 3i\sqrt{3}z^8 - z^9 - i\sqrt{3}z^9$$

$$z_{\text{single roots}} = \{1\} \quad z_{\text{double roots}} = \left\{ -1, -\frac{1}{2} + i\frac{\sqrt{3}}{2}, -\frac{1}{2} - i\frac{\sqrt{3}}{2}, \frac{1}{2} + i\frac{\sqrt{3}}{2}, \frac{1}{2} - i\frac{\sqrt{3}}{2} \right\} \quad (14.128)$$

The properties using the characteristic polynomial are summarized as follows, as originally proposed by Surrel (1996):

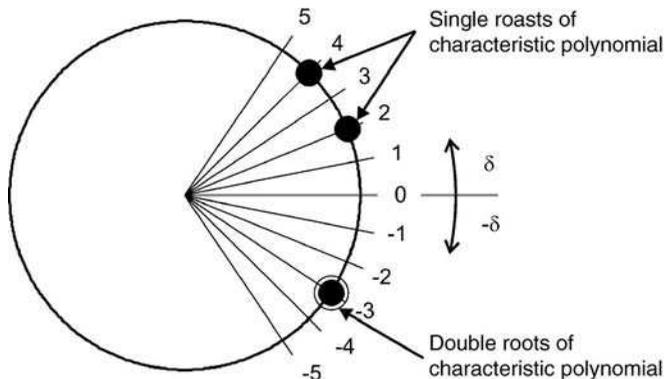


FIGURE 14.17. Plot of the characteristic polynomial for a specific algorithm. Large dot indicates single root for the m th order; circled dot indicates double root.

1. To detect the fundamental frequency ($m = 1$), the characteristic polynomial should have a root at $z = -m$ but none at $z = +m$.
2. If the characteristic polynomial has roots at $z = \pm m$ for $m \neq 1$, then the algorithm is insensitive to the m th-harmonic component.
3. In the case when only one $z = -m$ or $z = m$ is a double root of the characteristic polynomial but not the other, then that harmonic component can be detected, but the algorithm still is insensitive to miscalibration of the phase shifter.
4. If both $z = -m$ and $z = m$ are double roots of the characteristic polynomial, then the algorithm is insensitive to the harmonic component (m), as well as being insensitive to miscalibration of the phase shift at that case.

In 1996, Surrel also showed a compact way to display the features of the characteristic polynomials. In this representation shown in Figure 14.17, the phase steps of the algorithm are plotted as lines along the unit circle, where the angle of the line to the origin corresponds to the phase step. A big dot at the intersection of the line with the unit circle corresponds to a root of that specific harmonic. If a circle is drawn around the dot, it indicates the presence of a double root for that location. Figure 14.18 shows the plots of many examples for the characteristic polynomial for the algorithms presented in this chapter. In addition, where a thick line is drawn from the center, it indicates an overlap, where the phase steps include step values beyond one full period of 2π .

14.7.9. Summary of Algorithms

In this section, we examined a number of different PSI algorithms. They all require a series of interferograms to be recorded as the reference phase is shifted. The wavefront phase is then calculated at each measurement site as a function of the

intensities measured at that site. The result of these calculations must then go through a phase unwrapping process before the final wavefront map is produced. The algorithms are valid for both the phase-step and integrating-bucket data collection methods. In Section 14.9, we will examine characteristics of these algorithms with respect to error sources to aid in the selection of algorithms most appropriate for a particular application. The graphical comparison of many of these algorithms in Figure 14.18 is obtained by plotting the average reference phase angles required for each of the recorded interferograms.

Because of the nature of the arctangent function, there is no difference in shifting the algorithms multiples of π :

$$\tan(\phi + n \cdot \pi) = \tan(\phi) \quad (14.129)$$

Malacara et al. (2005) also summarized the effects of shifting the algorithms by $\pi/2$ and $\pi/4$. Shifting all phases by a constant is inconsequential, but can change the appearance of the algorithm. Replacing ϕ by $\phi + \frac{\pi}{2}$ yields:

$$\tan\left(\phi + n \cdot \pi + \frac{\pi}{2}\right) = \frac{\sin\left(\phi + \frac{\pi}{2}\right)}{\cos\left(\phi + \frac{\pi}{2}\right)} = \frac{\cos(\phi)}{-\sin(\phi)} = \frac{D}{-N} \quad (14.130)$$

which corresponds to a simple exchange of the sampling functions, and a change in the sign of one of them, but the sampling point is now located at phases displaced $\pm\pi/2$.

Replacing ϕ by $\phi + \frac{\pi}{4}$ yields:

$$\tan\left(\phi + n \cdot \pi + \frac{\pi}{4}\right) = \frac{\cos \phi + \sin \phi}{\cos \phi - \sin \phi} = \frac{D + N}{D - N} \quad (14.131)$$

These and other transformations can result in mathematical or computational simplifications that may improve efficiency or understanding for particular applications. Malacara et al. (2005) showed one algorithm, for which the reference sampling functions are orthogonal for all frequencies but amplitudes are unequal, changes into another algorithm for which the sampling reference functions have equal amplitudes at all frequencies but are orthogonal only at some. The following tables (14.7, 14.8) summarizes the presented algorithms and their properties along with their $\pi/4$ shifted counterparts.

Tables 14.7 and 14.8 give an overview of the discussed algorithms and also give some recommendations around error behavior. If looking for an algorithm especially targeted to suppress one specific error, like a significant vibration frequency the Malacara et al. (2005) book might be a good start, and also the papers discussed, which describe how to develop these algorithms, might be a required literature to fine tune the algorithm to the specific needs.

Table 14.9 gives an overview of the harmonic sensitivities of the presented algorithms. The harmonic sensitivities were derived by using the Fourier theory discussed and the amplitudes at the specific harmonic frequencies are shown.

TABLE 14.7. Overview of the presented algorithm up to 6 buckets and their properties.

N	Algorithm		Phase-Steps:	Family					Reference	Error behavior				
	$\tan \phi =$	$\tan\left(\phi + \frac{\pi}{4}\right) =$		N bucket	Avg A	Avg B	N + 1 Type A	N + 1 Type B		WDFT	linear phase-shift	vibration	vibration turbulence	detector non linearity
3	$\frac{\sqrt{3}(I_2-I_3)}{I_2+I_3-2I_1}$	$\frac{2I_1-I_2-\sqrt{3}(I_2-I_3)+I_3}{2I_1-I_2+\sqrt{3}(I_2-I_3)-I_3}$	$\delta_i = (i-1)\frac{2\pi}{3}$	X						Bruning, et al. 1974				
	$\frac{I_1-2I_2+I_3}{I_1-I_3}$	$\frac{I_3-I_2}{I_1-I_2}$	$\delta_i = (i-1)\frac{\pi}{2}$							Wyant, 1984b				
	$\frac{I_2-I_4}{I_3-I_1}$	$\frac{I_1-I_2-I_3+I_4}{I_1+I_2-I_3-I_4}$	$\delta_i = (i-1)\frac{\pi}{2}$	X						Bruning et al., 1974				good
4	$\frac{3I_2-(I_1+I_3+I_4)}{(I_1+I_2+I_4)-3I_3}$	$\frac{2(I_3-I_2)}{I_1-I_2-I_3+I_4}$	$\delta_i = (i-1)\frac{\pi}{2}$			X				Schwider et al., 1993	good			
	$\frac{\sqrt{3}(I_2-I_3)}{I_2+I_3-I_1-I_4}$	$\frac{I_1-I_2-\sqrt{3}(I_2-I_3)-I_3+I_4}{I_1-I_2-\sqrt{3}(I_2-I_3)-I_3+I_4}$	$\delta_i = (i-1)\frac{2\pi}{3}$				X			Larkin and Oreb, 1992				
5	$\frac{I_1-3I_2+3I_3-I_4}{\sqrt{3}(I_1-I_2-I_3+I_4)}$	$\frac{I_1-3I_2+3I_3+I_4+\sqrt{3}(I_1-I_2-I_3+I_4)}{-I_1+3I_2-3I_3-I_4+\sqrt{3}(I_1-I_2-I_3+I_4)}$	$\delta_i = (i-1)\frac{2\pi}{3}$					X		Surrel, 1993				
	$\frac{2(I_4-I_2)}{I_1-2I_3+I_5}$	$\frac{I_1-2I_2-2I_3+2I_4+I_5}{I_1+2I_2-2I_3-2I_4+I_5}$	$\delta_i = (i-1)\frac{\pi}{2}$		X		X			Schwider et al., 1983, Hariharan 1987	good	good	good	good
	$\frac{I_1-4I_2+4I_4-I_5}{I_1+2I_2-6I_3+2I_4+I_5}$	$\frac{I_1-I_2-3I_3+3I_4}{3I_2-3I_3-I_4+I_5}$	$\delta_i = (i-1)\frac{\pi}{2}$			X			X	Schmit and Creath, 1995				
	$\frac{3I_1-6I_2+4I_3-2I_4+I_5}{2I_1+2I_2-3I_3+2I_4-2I_5}$	$\frac{5I_1-3I_2+I_3-I_5}{-I_1+7I_2-7I_3+4I_4-3I_5}$	$\delta_i = (i-1)\frac{\pi}{2}$							Bi, 2004				
6	$\frac{I_1-2I_2+2I_4-I_5}{2(I_2-2I_3+I_4)}$	$\frac{-I_1+4I_3-4I_4+I_5}{I_1+4I_2-4I_3+I_5}$	$\delta_i = (i-1)\frac{\pi}{2}$							Surrel, 1997b				
	$\frac{-\sqrt{3}(I_2+I_3-I_5-I_6)}{2I_1+I_2-I_3-2I_4-I_5+I_6}$	$\frac{2I_1+I_2-I_3-2I_4-I_5+I_6-\sqrt{3}(I_2+I_3-I_5-I_6)}{2I_1+I_2-I_3-2I_4-I_5+I_6+\sqrt{3}(-I_2-I_3+I_5+I_6)}$	$\delta_i = (i-1)\frac{\pi}{3}$	X						Bruning et al., 1974				
	$\frac{-3I_2+4I_4-I_6}{I_1-4I_3+3I_5}$	$\frac{I_1-3I_2-4I_3+4I_4+3I_5-I_6}{I_1+3I_2-4I_3-4I_4+3I_5+I_6}$	$\delta_i = (i-1)\frac{\pi}{2}$				X			Schmit and Creath, 1995	good	good	good	
	$\frac{4I_4+I_5-I_1-3I_2-I_6}{I_1-I_2-4I_3+3I_5+I_6}$	$\frac{-2(I_2+I_3-I_4-I_5)}{I_1+I_2-2I_3-2I_4+I_5+I_6}$	$\delta_i = (i-1)\frac{\pi}{2}$							Zhao 1997, Surrel 1996, Malacara-Doblado, 2000	good			good
	$\frac{I_1-4I_2+2I_3+2I_4+I_5-2I_6}{2I_1-I_2-2(I_3+I_4-2I_5)-I_6}$	$\frac{3I_1-5I_2+5I_3-3I_6}{I_1+3I_2-4I_3-4I_4+3I_5+I_6}$	$\delta_i = (i-1)\frac{\pi}{2}$							Onodera, 2003				

TABLE 14.8. Overview of algorithms of 7 to 11 buckets and their properties.

N	Algorithm $\tan \phi =$	Phase-Steps:	Family						Error Behavior				
			N bucket	Avg A	Avg B	M Average	N + 1 Type A	N + 1 Type B	WDFT	Reference	linear phase-shift	vibration	vibration turbulence
7	$\frac{\sqrt{3}(-I_2 - I_3 + I_5 + I_6)}{I_1 + I_2 - I_3 - 2I_4 - I_5 + I_6 + I_7}$	$\delta_i = (i - 1)\frac{\pi}{3}$					X		Larkin, and Oreb, 1992		good		good
	$\tan \phi \frac{-I_1 - 3(I_2 + I_3 - I_5 - I_6) + I_7}{\sqrt{3}(I_1 + I_2 - I_3 - 2I_4 - I_5 + I_6 + I_7)}$	$\delta_i = (i - 1)\frac{\pi}{3}$						X	Larkin and Oreb, 1992, Surre, 1993		good		very good
	$\tan \phi \frac{7(I_3 - I_5) - (I_1 - I_7)}{8I_4 - 2I_2 - 2I_6}$	$\delta_i = (i - 3)\frac{\pi}{2}$							de Groot, 1994			good	
	$\tan \phi \frac{I_1 - 3I_3 + 3I_5 - I_7}{I_1 - 3I_3 + 3I_5 - I_7}$	$\delta_i = (i - 1)\frac{\pi}{2}$						X	Hibino, 1995			good	good
8	$\tan \phi \frac{I_2 - I_4 + I_6 - I_8}{I_1 - I_3 + I_5 - I_7}$	$\delta_i = (i - 1)\frac{\pi}{2}$					X		Schmit, 1996				
	$\tan \phi \frac{2I_2 - 4I_4 + 3I_6 - I_8}{I_1 - 3I_3 + 4I_5 - 2I_7}$	$\delta_i = (i - 1)\frac{\pi}{2}$					X			good			
	$\tan \phi \frac{3I_2 - 7I_4 + 5I_6 - I_8}{I_1 - 7I_3 + 5I_5 - 3I_7}$	$\delta_i = (i - 1)\frac{\pi}{2}$					X			good	good		
	$\tan \phi \frac{4I_2 - 11I_4 + 8I_6 - I_8}{I_1 - 8I_3 + 11I_5 - 4I_7}$	$\delta_i = (i - 1)\frac{\pi}{2}$					X			good	good	good	
10	$\tan \phi \frac{5I_2 - 15I_4 + 11I_6 - 8I_8}{I_1 - 11I_3 + 15I_5 - 5I_7}$	$\delta_i = (i - 1)\frac{\pi}{2}$					X			very good			very good
	$\tan \phi \frac{\sqrt{3}(-I_1 - 3I_2 - 3I_3 + I_4 + 6I_5 + 6I_6 + I_7 - 3I_8 - 3I_9 - I_{10})}{I_1 - I_2 - 7I_3 - 11I_4 - 6I_5 + 6I_6 + 11I_7 + 7I_8 + I_9 - I_{10}}$								Hibino, 1995			very good	very good
11	$\tan \phi \frac{(I_1 - I_{11}) - 8(I_3 - I_9) + 15(I_5 - I_7)}{(4I_2 + I_{10}) - 12(I_4 + I_8) + 16I_6}$	$\delta_i = (i - 6)\frac{\pi}{2}$							de Groot, 1994	good			very good

TABLE 14.9. The harmonic sensitivities for the described algorithms are presented below.

Algorithm	Suppresses harmonics										Frames
	2	3	4	5	6	7	8	9	10	N	
Bruning, et al. 1974	-	y	-	-	y	-	-	y	-	-	} 3
Wyant, et al. 1984b	-	-	y	-	-	-	y	-	-	-	
Bruning, et al. 1974	y	-	y	-	y	-	y	-	y	-	} 4
Schwider, 1993	-	-	y	-	-	-	y	-	-	-	
Larkin and Oreb, 1992	-	y	-	-	y	-	-	y	-	-	} 4
Larkin, 1992, Surrel, 1993	-	y	-	-	y	-	-	y	-	-	
Schwider, 1983, Hariharan, 1987	y	y	y	-	y	y	y	-	y	-	} 5
Schmit and Creath, 1995	-	-	y	-	-	-	y	-	-	-	
Bi, 2004	-	-	y	-	-	-	y	-	-	-	} 5
Surrel, 1997b	-	-	y	-	-	-	y	-	-	-	
Schmit and Creath, 1995	y	-	y	-	y	-	y	-	y	-	} 6
Zhao, 1997, Surrel, 1996	y	-	y	-	y	-	y	-	y	-	
Onodera, 2003	-	-	y	-	-	-	y	-	-	-	} 6
Larkin and Oreb, 1992	y	y	y	-	y	-	y	y	y	-	
Larkin, 1992, Surrel, 1993	y	y	y	-	y	-	y	y	y	-	} 7
de Groot, 1994	y	-	y	-	y	-	y	-	y	-	
Hibino, 1995	y	-	y	-	y	-	y	-	y	-	} 7
Schmit and Creath, 1996	y	-	y	-	y	-	y	-	y	-	
Schmit and Creath, 1996	y	y	y	y	y	y	y	y	y	y	} 8
Schmit and Creath, 1996	y	-	y	-	y	-	y	-	y	-	
Schmit and Creath, 1996	y	-	y	-	y	-	y	-	y	-	} 8
Schmit and Creath, 1996	y	-	y	-	y	-	y	-	y	-	
Hibino, 1995	y	y	y	-	y	-	y	y	y	-	} 10
de Groot, 1994	y	-	y	-	y	-	y	-	y	-	

The amount of steps in the design of a phase shift algorithm is principally not limited, however for very large numbers it develops towards and Fourier evaluation with its own advantages and disadvantages (See de Groot, 1997 for a discussion).

14.8. PHASE SHIFT CALIBRATION

An important step in setting up a PSI system is to calibrate the phase shift α between recorded interferograms. Some algorithms are very sensitive to errors in the phase shift, and even the algorithms that are tolerant of phase shift errors require calibration for best performance. If the integrating bucket approach is used, the phase centroid of each integrating period must be calibrated.

A convenient way to calibrate the phase shift is to use the solution for α from the Carré algorithm (Eq. (14.74), which gives half the step size). If a series of four interferograms are recorded with equally spaced steps, the Carré algorithm yields the phase shift at each measurement point.

A simpler expression for the phase shift can be found by using five interferograms recorded with equal steps of α (Schwider et al., 1983; Cheng and Wyant, 1985b). These are the same intensities used in the Hariharan algorithm (Eqs. (14.79)–(14.83), and the solution for the phase shift is

$$\alpha(x, y) = \cos^{-1} \left[\frac{1}{2} \left(\frac{I_5 - I_1}{I_4 - I_2} \right) \right] \quad (14.132)$$

Note that the third interferogram or intensity of the series is not needed for this calculation, and the phase shift can be calculated at each point in the field.

Both of these analytic expressions for the phase shift have a singularities for certain values of the wavefront phase $\phi(x, y)$. To avoid errors, a few tilt fringes are introduced into the interferogram, and data points which cause the denominators of the particular equation to fall below some threshold are eliminated from the analysis. A convenient way to display these calculated results is to show a histogram of the measured phase shifts (Creath, 1988). A sample histogram for a 90° phase shift is shown in Figure 14.19, and the controller of the phase shifter should be adjusted to center this curve on the desired value and to minimize its standard deviation. The method using five frames in Eq. (14.132) is the most commonly used calibration procedure.

The width and shape of the histogram also provides an easy way of monitoring the variation of spatial phase shift. The histogram in Figure 14.19, above, was generated on a Fizeau interferometer with plano objectives. It is very symmetric and has a half width of 4° . The histogram in Figure 14.20 was obtained on the same interferometer, but with a $f/0.75$ reference objective. In this case, the width of the curve has increased to about 15° , and it is skewed to lower values of phase shift. The skew is due to the high f -number rays that make a large angle with the optical axis. These rays see a phase shift that has been reduced by the cosine of this angle, and this situation points out the need for an algorithm that is insensitive to errors in the phase shift.

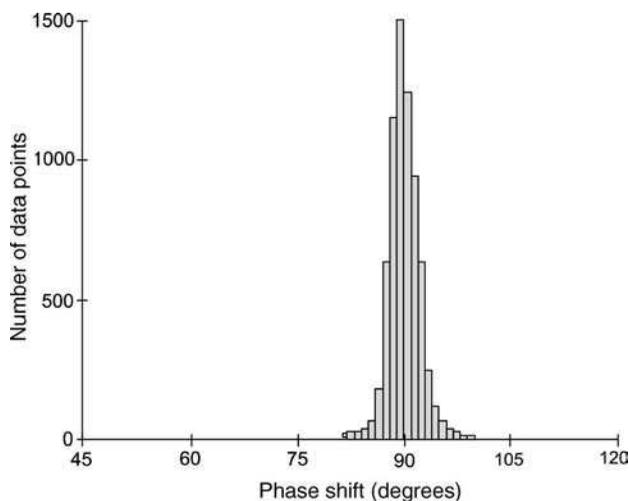


FIGURE 14.19. A histogram showing the distribution of the measured phase shifts.

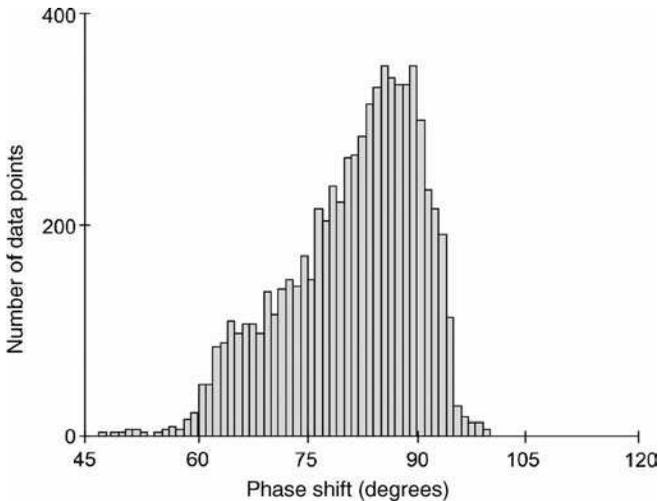


FIGURE 14.20. A histogram of the measured phase shifts obtained with a fast reference objective on a laser Fizeau interferometer.

A few simpler but less accurate methods for calibrating the phase shift involve the visual comparison of a series of interferograms. If four interferograms collected with a 90° phase shift are displayed (as described by Eq. (14.79)–(14.83) with $\alpha = \pi/2$, and as shown in Figure 14.6), the first and fifth frames are 2π out of phase and should appear identical. Complementary intensity patterns should appear in the first and third, second and fourth, and third and fifth because of the π phase shift between each pair. The addition of a few tilt fringes makes the analysis easier. This intensity comparison can also be done by plotting a linear slice across each of these interferograms (Cheng and Wyant, 1985b). For these five interferograms, the curves from the first and the fifth should lie on top of each other. This is demonstrated in Figure 14.21a as a properly adjusted 90° phase shift and in Figure 14.21b, which shows the error that results for an 97° phase shift. These techniques can be expanded to other values of phase shift. If $N + 1$ interferograms are collected with a phase shift of $2\pi/N$ (the N -step algorithm in Eq. (14.67), the first and the last of these interferograms should overlap. With either of these two techniques, the phase shift is calibrated by changing the gain on the PZT driver until the visual match is obtained. An accuracy of plus or minus a few degrees is possible.

Another method for calibrating the phase shift is to use a separate interferometer to monitor the position of the reference mirror as it translates (Hayes, 1984). The detected intensity is then used to control the phase shift controller. This reference interferometer can also be a phase shift interferometer. An alternative solution is to use the measured phase positions with the generalized least-squares algorithm (Section 14.8.2). This results in an algorithm that adapts to the actual phase shifts for which the data was collected (Seligson et al., 1984).

This type of reference monitoring system is very useful for detecting and correcting nonlinearities in the phase shift. The nonlinearity in even good PZTs can range

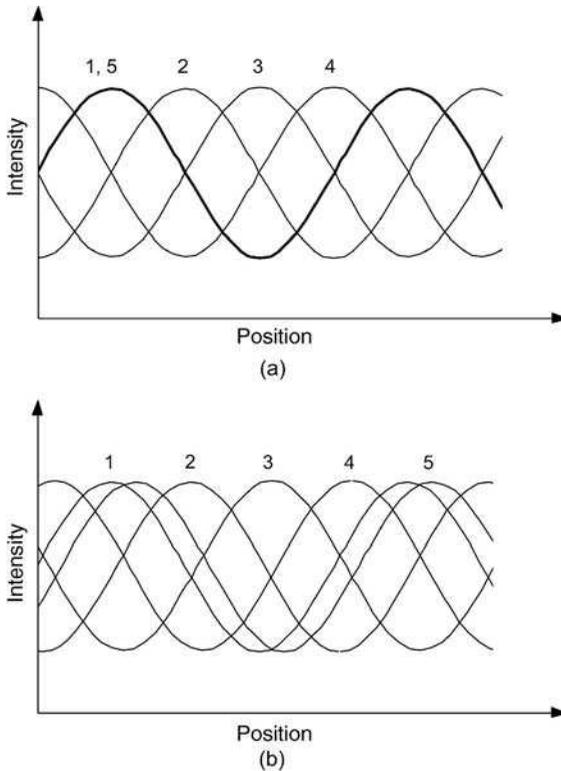


FIGURE 14.21. Intensity traces across five phase-shifted interferograms: (a) 90° phase shift and (b) 97° phase shift (From Cheng and Wyant, 1985b).

from less than 1% to a few percent, and they often exhibit hysteresis. The effect of these nonlinearities will be examined in the next section. The other calibration procedures provide a linear correction for the phase shift, but do not correct for nonlinearities. By knowing the actual travel of the reference mirror as a function of drive signal or voltage, the controller signal that will produce a linear phase shift can be generated. In practice, this is usually done through a digital look-up table that is converted into the analog drive signal.

14.9. ERROR SOURCES

There are numerous sources of error that can affect the accuracy of phase measurements determined by the various PSI algorithms. Some of the PSI algorithms are more sensitive to a particular error source than others, while some sources of error are fundamental and affect the accuracy of all algorithms. Creath (1986) conducted a thorough simulation of errors for most of the common algorithms. Error contributions specific to each algorithm have been conveniently tabulated in a recent reference and

include those attributable to the laser source, the phase shifter, the detector, and the environment (van Wingerden et al., 1991). PSI algorithms produce a phase measurement that is a relative OPD measurement. Often what is desired is an absolute measurement of an optical surface or the OPD of an optical system. Many factors inhibit a direct measurement of the desired quantity. What follows is a discussion of how to identify and minimize the influence of the various sources of error.

It is important to emphasize the difference between precision, repeatability, and accuracy. A system that possesses very small *random* errors has high precision. A system that gives the same result for consecutive measurements has high repeatability. High repeatability, however, does not imply either high accuracy or high precision, but high accuracy is not possible without high repeatability. A system with very small *systematic* errors can have high accuracy because random errors can be reduced by averaging. Generally we try to minimize both types of error, but in all cases the error contributions must be understood if improvements in overall accuracy are to be attained. Some of the largest sources of error in PSI are systematic, and though quite well understood, are often difficult to eliminate. The use of high-performance microprocessors has contributed to the reduction of many component errors through the efficient and inexpensive manipulations of large quantities of data. More improvement can be expected.

Error sources generally fall into three categories: (1) those associated with the data acquisition process, (2) environmental effects such as vibration and air turbulence, and (3) those associated with defects in optical and mechanical design and fabrication. The data acquisition process includes errors in the phase shift process, nonlinearities in the detection system, amplitude and frequency stability of the source and quantization errors obtained in the analog-to-digital conversion process. We begin with the discussion of errors in the first category since these are the most controllable and tend to be quite algorithm dependent.

14.9.1. Phase Shift Errors

All PSI algorithms rely on shifting the phase in a known manner in an interferometer as part of the data acquisition process. The expected variation of intensity at any point in the interference pattern, under ideal conditions, is purely sinusoidal with a linear change in path length of the phase shifting arm (Eq. (14.6)):

$$I(x, y, t) = I'(x, y) + I''(x, y) \cos[\phi(x, y) + \delta(t)] \quad (14.133)$$

If the increment and total range of the phase shift $\delta(t)$ during data acquisition is not the predicted subdivision of 2π , then the measured intensity samples do not precisely map to one complete period of a sinusoidal signal. In other words, a precisely sinusoidal signal is expected, and the acquired data points are fit to a sinusoidal signal in a least-squares or Fourier series sense, with that assumption. For example, if the reference phase shifter for an N -step algorithm does not shift each step by the expected $2\pi/N$ increment in OPD, but instead is off by some error ε_r at each step, the least-squares algorithm will try to fit a set of intensities representing a

sinusoidal signal of period $2\pi + \varepsilon'_r$ to a sinusoidal function of period $2p$, where ε'_r represents the deviation from a full 2π period. A sinusoidal signal of period $2\pi + \varepsilon'_r$ can be represented by a Fourier series of sinusoids at the fundamental frequency $\nu = 1/2\pi$, together with all higher harmonics, $2m\nu$, where m is an integer. There is a corresponding error that is imparted to the measured phase under these circumstances that shows up most severely in the first harmonic of the measured phase $\phi'(x, y)$. Schwider (1989) has shown that for a small error ε_r in the reference phase $\phi_r(x, y)$, there is a corresponding error in the measured phase $\phi'(x, y)$ from the true phase $\phi(x, y)$ of

$$\Delta\phi(x, y) = \phi'(x, y) - \phi(x, y). \quad (14.134)$$

A detailed calculation of the phase error is quite lengthy, but for the general case of an N -step least-squares algorithm, we have

$$\Delta\phi(x, y) \approx a + b \cos[2\phi(x, y)] + c \sin[2\phi(x, y)] \quad (14.135)$$

The coefficients a , b , and c are summations that can be interpreted as Fourier coefficients for the spatial variation of the interference pattern at twice the spatial variation of the true phase $\phi(x, y)$. A higher-order approximation would show higher-order harmonics with correspondingly smaller coefficients.

The error in Eq. (14.135) can be minimized in a number of practical cases. The most notable of these is the Schwider et al. (1983) and Hariharan (1987) algorithm (Section 14.7.4) in which 5 samples are taken with a phase shift of $\pi/2$ between samples. In this case, Eq. (14.135) reduces to Eq. (14.95) and is repeated here:

$$\Delta\phi(x, y) \approx (\varepsilon_r^2/4) \sin[2\phi(x, y)] \quad (14.136)$$

The Carré algorithm shows no measurement error for a linear error in the phase shift increment since this algorithm analytically determines the phase steps taken during the measurements. This algorithm was developed for long baseline interferometric distance measurements in a coordinate measurement tool (Carré 1966).

The Schwider et al. (1983) and Hariharan (1987) algorithm is quite insensitive to other error types in addition to the linear calibration error in the phase shift step. Higher-order nonlinearities in the phase shifter are more difficult to predict analytically, but quite simple to simulate numerically. Hariharan has shown that deviations in linearity of 1% from the nominal $\pi/2$ phase shift steps cause a maximum error of less than 0.005° in the measured value of the phase ϕ . Hysteresis errors of the magnitude give rise to comparable errors.

Linear drift in the phase shifter is minimized with this algorithm, but in fact, can be nearly eliminated by a simple additional operation. The effect of such drifts, if small, can nearly be eliminated by acquiring a second data set in the reverse order and averaging with the first. Hariharan et al. (1987a) have shown in this case that a drift of 4° over two sample periods (corresponding to values of $-178^\circ, -88.5^\circ, 1^\circ, 90.5^\circ$ and 180° in the first set and $180^\circ, 89.5^\circ, -1^\circ, -91.5^\circ$ and -182° in the second set) contributes an error of only 0.001° when the two sets are averaged. In practice, drift or instability in a particular part of the interferometer may be indistinguishable to the algorithm from the drift in the phase shifter since both can appear as changes in OPD over the same field.

In spite of the relative insensitivity of the Schwider et al. (1983) and Hariharan (1987) algorithm to phase shift errors, it is important to calibrate the phase shift steps. A simple calculation, which is convenient for this algorithm, was proposed by Schwider et al. (1983) in which the phase step α is calculated from the following equation:

$$\cos(\alpha) = \frac{I_5 - I_1}{2I_4 - I_2} \quad (14.137)$$

The use of this equation is restricted because of the singularity when $I_4 - I_2 = 0$, as discussed in Section 14.1. In-depth analyses of the effects of phase shift errors are covered by Kinnstaeter et al. (1988), Ohyama et al. (1988) and by Freischlad and Koliopoulos (1990).

In practice, it is important to choose the algorithm which is most insensitive to phase shift errors for each particular application. This results from the fact that most phase errors are induced by changes in OPD in the interferometer cavity due, and depending on the application may be dominated by drifts, turbulence, vibrations, and not by incorrect movements of the reference mirror (i.e., phase shifter) or a combination thereof.

Figure 14.22 shows a numerical simulation of the error sensitivity of linear phase shift error for algorithms presented before. For algorithms using up to 4 steps, the algorithm presented by Schwider et al. (1993) has the best compensation behavior as shown in the top chart. When using up to 6 phase steps, the algorithm introduced by Schmit and Creath (1995) using 6 intensities has improved behavior over the best and very common 5 bucket algorithm by Schwider et al. (1983) and Hariharan (1987).

Figure 14.23 summarizes the behavior of phase-shift algorithms with 7 to 11 frames. Here the averaged algorithms presented by Schmit and Greath (1996) and de Groot (1994) using an optimized sampling window yield the best performance. In general this also shows that the more frames are available, the better an algorithm can be tailored to reduce error sources.

14.9.2. Detector Nonlinearities

Nonlinearity of the detector represents an error that can become important when the dynamic range or contrast of the fringes is very high. Correspondingly, when contrast is low, detector nonlinearity is less significant. Kinnstaeter et al. (1988) have shown that with the four-step algorithm ($\alpha = \pi/2$), phase measurement is free of error not only for a linear detector, but even when the detector has second-order nonlinearities. This is not the case for the three-step algorithm. Forth-order detector nonlinearities, however, contribute errors that excite the third harmonic of the spatial variation of phase $\phi(x, y)$ for the four-step algorithm. More phase steps in the PSI algorithm are effective in reducing higher-order nonlinearities in the detector, but these errors are generally of no consequence. Modern silicon array sensors operate linearly over dynamic ranges of nearly 100:1 at room temperature. The range can be easily extended by at least another order of magnitude with thermoelectric or cryogenic cooling. This is not usually needed except in cases where light levels are very low or the pixel integration time, by requirement, is very short.

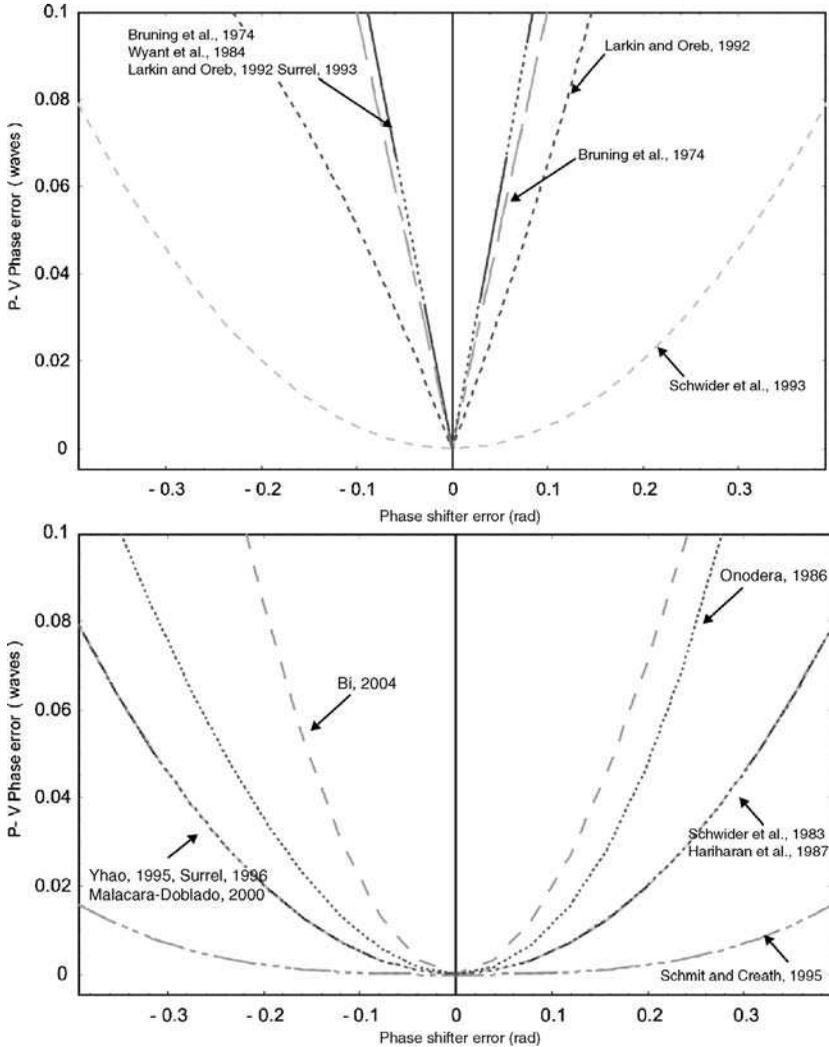


FIGURE 14.22. Linear phase shifter error for algorithms with up to 4 frames presented in the top chart. Phase error for algorithms containing 5 or 6 intensity samples presented in the bottom chart.

Detector nonlinearity in CCDs is more likely to be encountered at high intensities when the diode array is operating at or near saturation. This is easily remedied by insertion of an appropriate attenuator. The integrating CCD-type detector has the greatest dynamic range if it can be operated just below saturation. This is achieved by selecting the proper laser power, optical attenuation, and integration time (frame rate).

Figure 14.24 shows different detector nonlinearities illustrated by different gamma values. The terminology of gamma correction is common in modern video applications.

$$I_{detected} = I_{real}^\gamma \tag{14.138}$$

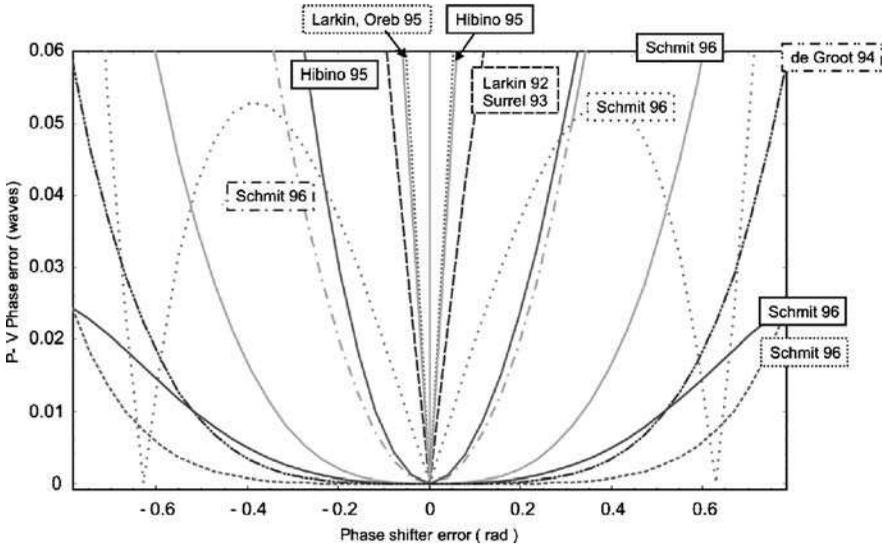


FIGURE 14.23. Simulated linear phase shifter error for algorithms presented using 7 buckets or more.

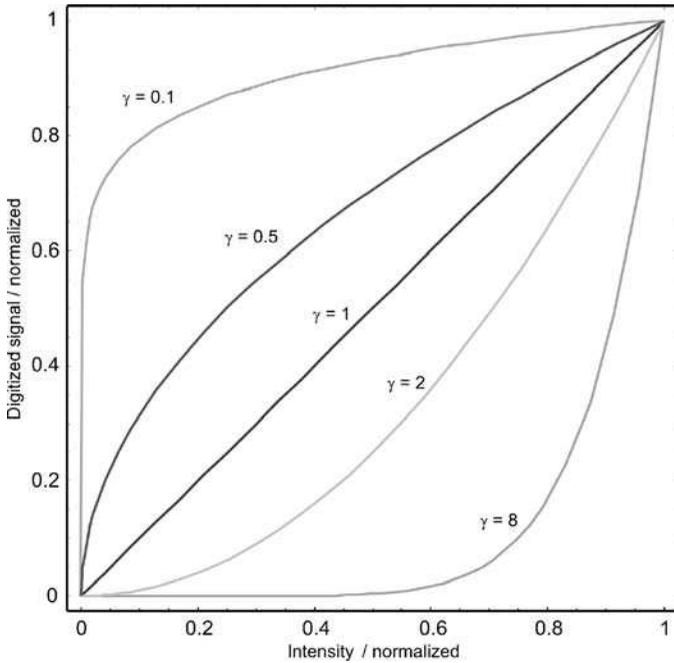


FIGURE 14.24. Example gamma transfer functions for several cases. Note that a low gamma value is equivalent of a detector operating near saturation and a high gamma value corresponds to intensities partly below the detection limit.

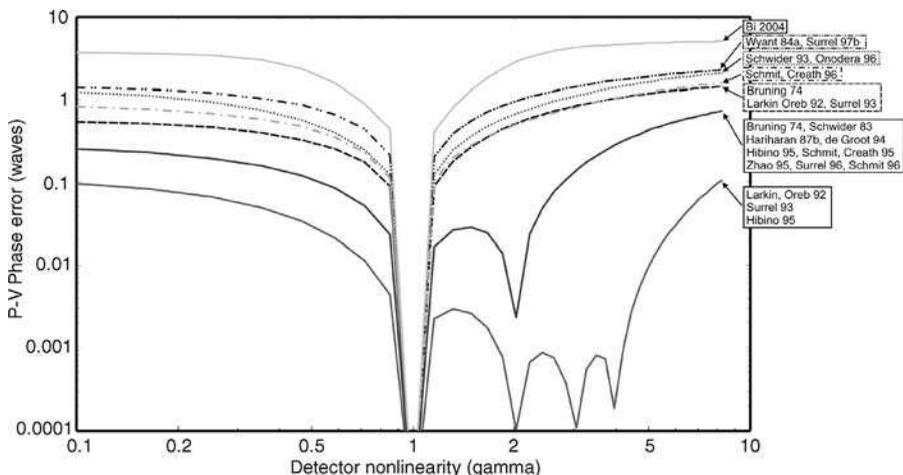


FIGURE 14.25. P-V wavefront error introduced by detector nonlinearities for specific gamma transfer functions. The chart shows all 3 types of algorithms that can be distinguished.

Eq. (14.138) illustrates the relations between the actual intensity at the detector location I_{real} and the detected intensity $I_{detected}$. Both intensities are normalized between 0 and 1 for Eq. (14.138) to be valid. As can be seen in Figure 14.24, a low gamma value is equivalent of a detector operating near saturation and a high gamma value corresponds to intensities below the detectors threshold. Figure 14.25 shows the sensitivities of different PSI algorithms to different simulated gamma values for detector nonlinearities.

14.9.3. Source Stability

Stability of the interferometer source is important both in frequency and amplitude. Frequency instability, particularly with lasers, can give rise to corresponding instabilities in the fringe pattern and reduce the accuracy of the measured wavefront. This is potentially more troublesome for interferometers that operate at large differences in path length in the two arms of an interferometer. The long coherence length of the He-Ne laser may cause one to neglect this factor, but for high accuracy requirements this must be considered. For a path length difference in the interferometer of ΔL and a desired measurement accuracy of k wavelength, the frequency stability $\Delta\nu$ of the sources satisfies the inequality

$$\Delta\nu < \frac{c \cdot k}{\Delta L}, \tag{14.139}$$

where c is the velocity of light. This is derived from the assumption that we should be able to determine the phase to within one wave if the coherence length of the laser matches the path length difference. Thus, a path difference of 3m requires a source stability of 1MHz for $< 0.01 \lambda$ accuracy. While large OPDs present a problem for

sources with poor frequency stability, this suggests an alternate method of shifting the fringes if the source frequency or wavelength is an easily controllable parameter in the system. This has been exploited with laser diode sources and is covered in greater detail in the Section 14.16.4.

Intensity fluctuation of the source, if random, can be mitigated by averaging since all points in the field vary together under most circumstances. The effect of intensity fluctuations on phase measurement depends on the algorithm used. In the case of the N -step least-squares algorithm, it has been shown that the standard deviation in the measured phase σ_ϕ is given by (Bruning 1978; Koliopoulos 1981)

$$\sigma_\phi = \frac{1}{\sqrt{NS}} \quad (14.140)$$

where S is the signal-to-noise ratio of the detector system. This and other noise sources have been analyzed in great detail for most of the algorithms by Freischlad and Koliopoulos (1990) and Brophy (1990). Intensity fluctuations have the same general behavior as small amplitude vibrations, which are covered in Section 14.9.5. In the presence of systematic and slowly varying intensity fluctuations, a supplementary detector can be used to monitor the variations, and either the ratio of the results at the detector or the signal may be used in feedback to stabilize the source. The latter is usually preferable.

The ultimate limitation to precision in two-beam interferometry is ascribed to photon shot noise at the detector. Tarbeyev (1986) estimated the precision limit at 6×10^{-5} nm.

14.9.4. Quantization Errors

PSI has become a practical technique because of inexpensive digital computing. The first step is the conversion of analog intensity signals to digital information. A source of error at this stage is quantization error of the video signal. Since conversion is accomplished with an analog-to-digital converter, the accuracy of this conversion process depends upon the number of bits in the digital word transferred to the computer. Common video converters digitize the analog input signal into an 8- to 16-bit word, meaning that there are $2^8 = 256$ to $2^{16} = 65536$ discrete quantization levels in the digital word. The effect of quantization errors was first discussed by Koliopoulos (1981) for a three-step algorithm and later generalized by Brophy (1990) who rigorously derived the error for most of the common algorithms along with the special case of a 13-step algorithm with phase shift steps of 75° . The simulations are complicated by the fact that quantization error is not a statistical error. Brophy has shown that the standard deviation of the phase due to Q quantization levels, ϕ_4 , for a four-step algorithm ($a = \pi/2$) is approximately

$$\sigma_{\phi_4} \approx \frac{1}{\sqrt{3Q}} \quad (14.141)$$

and

$$\sigma_{\phi_N} \approx \frac{1}{\sqrt{3NQ}} \quad (14.142)$$

can be inferred for an N -step algorithm. The 13-step algorithm shows about a factor of two in the improvement in error as a result of reduced correlation between steps. Expressions (14.138) and (14.139) assume that the fringe modulation spans the full dynamic range of quantization levels. For 1024 quantization levels, the root mean square (RMS) phase error of the four-step algorithm is < 0.0001 waves. This is completely negligible under most circumstances, suggesting that an 8-bit converter is sufficient. In practice, when fringe modulation decreases, fewer bits of dynamic range are used, which has the effect of decreasing Q and thus increasing the error. Today, using high-resolution images, digitized with more bits per pixel, is no longer an issue because camera interfaces, transfer speeds as well as computer memory have caught up to the increasing requirements.

14.9.5. Vibration Errors

Mechanical stability is paramount in interferometry. The presence of vibration can be detrimental to the point of completely obliterating fringes as seen by the detector. Yet vibration, in some instances, cannot be eliminated. Special techniques have been developed to cope with these situations. Examples are stroboscopic illumination and holographic methods where the data acquisition system can freeze the vibrating fringes so that some of the same analytical methods can be deployed to analyze the fringes. A condition of large amplitude vibrations could manifest itself by motions in two or three dimensions of the test structure or interferometer, while other motions may exist at some other fundamental vibrational mode. This type of vibrational analysis will not be covered here, but can be found in other references (Powel and Stetson, 1965; Hariharan et al., 1987b; Hariharan, 1984).

The usual precautions to minimize vibrations in interferometry are assumed, such as the use of passive or active vibration isolations mounts, acoustical vibration and damping, as well as generally good structural practice in the design of the interferometer and the table to which it is mounted. Location of the interferometer in a mechanically quiet area, far from rotating machinery, is also crucial.

The presence of a small amount of vibration usually manifests itself in a manner equivalent to perturbations of the phase shift increment and can thus be analyzed with familiar methods. Eq. (14.135) indicates the general behavior of error in the measured phase with error in the phase shift increment. This behavior is characterized by error terms at twice the spatial frequency of the fringe pattern. To first order, those algorithms that have the greatest immunity to errors in the phase shift increment have the greatest immunity to vibration and drift of small amplitude. When the vibration frequency is high compared to the pixel frame rate,

the amplitude fluctuations tend to average out, and there is little correlation in the vibrational component at each pixel location. As one might expect, the effect of vibration induced intensity fluctuations depends on the PSI algorithm used. The standard deviation of the phase introduced by small vibrations, averaged over all possible values of phase ϕ_1 , is related to the standard deviation of the intensity fluctuation σ_1 through the simple equation (Brophy 1990):

$$\sigma_{\phi_f} = \frac{\sqrt{k}\sigma_I}{\gamma I_0} \quad (14.143)$$

Here γ is the fringe modulation, I_0 is the mean intensity, and k is a constant that depends on the PSI algorithm used. The parameter k decreases roughly in inverse proportion to the number of phase steps in the algorithm. For a four-step algorithm, $k = 1/2$, and for an N -step algorithm, $k = 2/N$. The variance in phase (in radians) is always less than the variance of uncorrelated frame-to-frame intensity noise, even if that intensity noise is introduced by small vibrations or intensity variations of the source.

In 1995 and 1996 de Groot published articles especially addressing vibration in phase shifting interferometry. In his 1995 article he derived first order approximations resulting in a phase-transfers function, that multiplied by the noise spectrum is able to predict the response of an interferometer for various forms of vibration.

In their 1996 article, de Groot and Deck presented numerical simulations of vibration in phase-shifting interferometry and their effect on the root-mean-square (RMS) measurement error. Here they used the RMS error for a given amplitude A and frequency ν :

$$E = \frac{1}{2\pi} \left(\int_{-\pi}^{\pi} \int_{-\pi}^{\pi} \{ \Delta\varphi(\phi, \alpha) - \Delta\varphi_{average}(\alpha) \}^2 d\phi d\alpha \right)^{1/2} \quad (14.144)$$

with the phase error $\Delta\varphi = \varphi - \phi$ including the actual phase ϕ and the calculated phase $\varphi = \tan^{-1}(T) + \text{const}$, where

$$T = \frac{\sum_{j=1}^J s_j \bar{g}_j}{\sum_{j=1}^J c_j \bar{g}_j} \quad (14.145)$$

is the algorithm used and \bar{g}_j represents the intensities. In the case of the integrating-bucket method, which is preferred for use when trying to reduce vibration sensitivity, de Groot and Deck (1996) showed that

$$\bar{g}_j = \frac{1}{\tau} \int_{-\tau/2}^{\tau/2} g(\phi, t_j + t') dt' \quad (14.146)$$

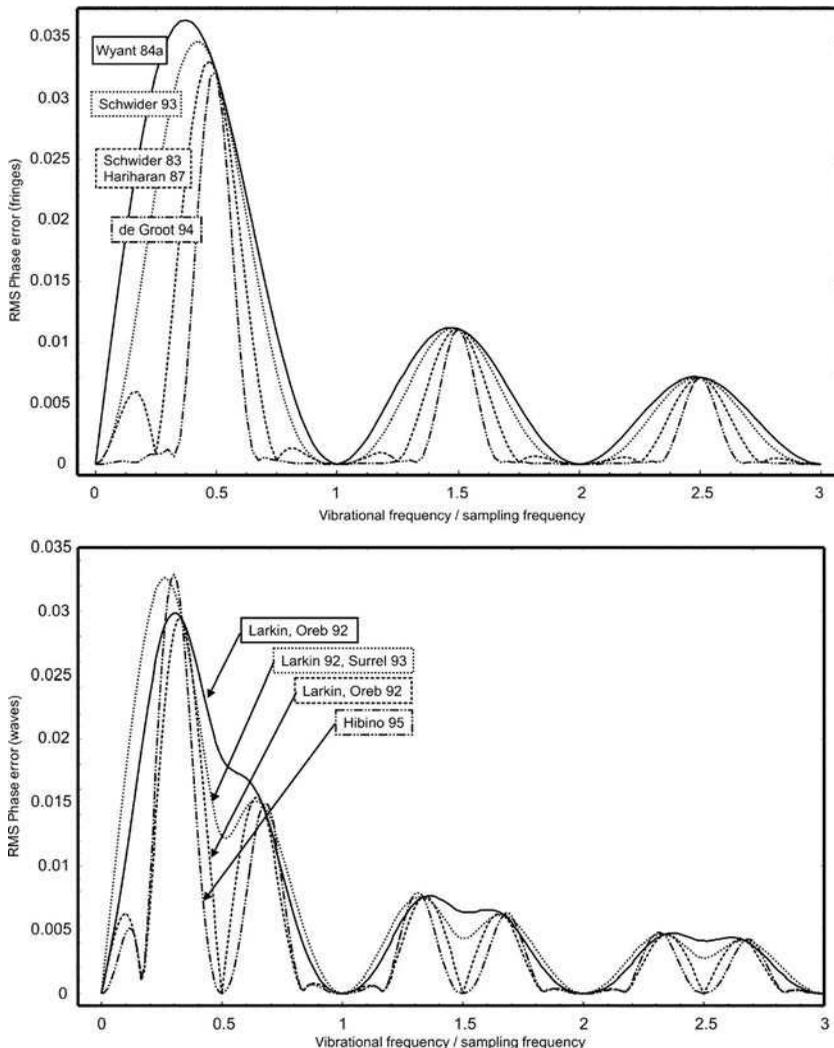


FIGURE 14.26. Numerical simulations of the RMS wavefront error introduced by various algorithms vs. normalized vibration frequency.

Figure 14.26 shows numerical simulations using de Groot’s (1996) method for various algorithms presented in section 14.7. Two types of general improved behavior under vibration conditions can be observed. They are the algorithms in the upper chart in Figure 14.26; narrow the peaks and flatten the valleys for the odd harmonics of the ratio of the sampling frequency to the vibrations frequency.

There are also those algorithms that show reduced sensitivity at the odd harmonics but introduce sidelobes between the even and odd harmonics.

There are significant differences between the two types of algorithms here, particularly at frequencies below the actual sampling frequency as shown in Figure 14.27.

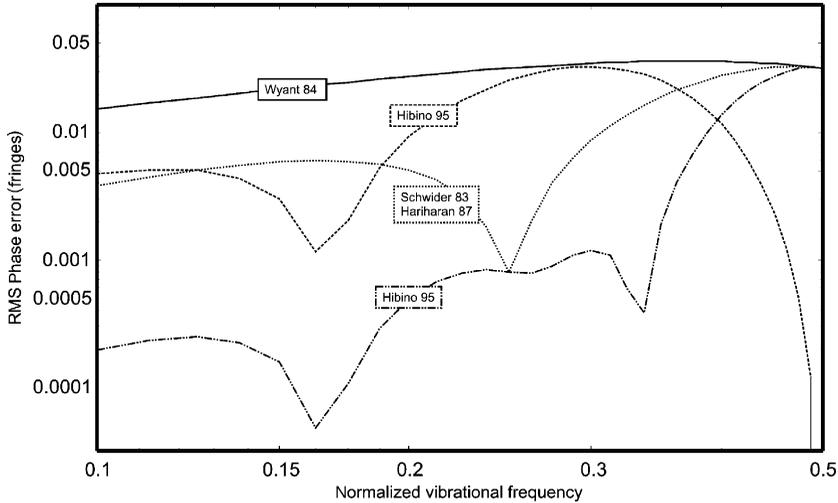


FIGURE 14.27. Plot of numerical simulations for low vibration frequencies. Note the double logarithmic scale.

14.9.6. Air Turbulence

Air currents and air turbulence are difficult source of error, particularly, when high accuracy measurements are required. The character at any particular point in the fringe field tends to be a slowly varying phase error with reduced correlation to neighboring points at distances greater than a few millimeters to centimeters (Rosenbluth and Bobroff, 1990). The same problem plagued early holographers until they learned how to properly “mix” the ambient air. Proper mixing of the air is achieved by creating enough airflow to prevent stagnation or layering, without generating excessive turbulence.

Air currents and turbulence are not unrelated to thermal gradients, which exist as a result of isolated heat sources and sinks. When a thermal gradient exists in any of the structural parts of the interferometer or part under test, time-dependent changes are bound to be there in the optical path over the wavefront aperture. Sufficient time must be allotted for thermalization and the instrumentation of the part. This can be hours, when measurement accuracies at the $\lambda/100$ level are required. The particular measurement strategy, however, must be tailored to take into account the time constants of the potential error sources and the time of the data acquisition algorithms. Figure 14.27 shows numerical simulations of low-frequency vibration sensitivity of various algorithms using de Groot’s (1996) method.

14.9.7. Extraneous Fringes and Other Coherent Effects

The long coherent length of the laser source allows interference with very large differences in path length in the test and reference arms in a Twyman–Green or Fizeau interferometer. While this may provide mechanical convenience in some cases, reflections from surfaces within the coherence length can interfere with one

another and confuse the intended measurement. Extraneous fringes introduce systematic measurement errors that, in some cases, can be eliminated.

Consider first the effect of three interfering beams. Let $e^{i\phi(x,y)}$ be the test arm field, $e^{i\phi_r(x,y)}$ the reference arm field, and $q \cdot e^{i\eta(x,y)}$ be an extraneous reflection field with fractional amplitude q of the primary interfering beams. When these three beams interfere, the resulting intensity is

$$I_r = 2 + q^2 + 2 \cos[\phi(x,y) - \phi_r(x,y)] + 2q \cos[\eta(x,y) - \phi_r(x,y)] + 2q \cos[\phi(x,y) - \eta(x,y)] \quad (14.147)$$

The true phase $\phi(x,y)$ is distorted to $\phi'(x,y)$ by the error $\Delta\phi(x,y)$, which can be shown to be (Schwider et al., 1983)

$$\Delta\phi(x,y) = \tan^{-1} \left\{ \frac{q \sin[\eta(x,y) - \phi(x,y)]}{1 + q \cos[\eta(x,y) - \phi(x,y)]} \right\} \quad (14.148)$$

The error $\Delta\phi(x,y)$ is seen to be dependent on the true phase $\phi(x,y)$. This reduces the value reported by Bruning et al. (1974) when $\phi(x,y) = 0$. If the test arm phase is shifted by π by incorporating an additional phase shifter and another data set is acquired and averaged with the first, a strong decrease in the amplitude of the disturbing wavefront results (Schwider et al., 1983).

Ai and Wyant (1988) proposed a modification of this procedure that is easier to implement, particularly in the case of the four-step method. In their method four intensities are taken using the four-step bucket, and the individual intensities are stored. Following that, the test beam is blocked just prior to the test surface, and four new intensities are acquired and subtracted from the first set. The phase is calculated from the resulting intensities, which removes the effect of the spurious reflection without error. These analyses apply to cases where the reference and test surfaces have relatively low reflectivity and are matched. The situation becomes much more difficult if the test surface is highly reflective. Hariharan (1987) has shown that the three-step algorithm is much more sensitive to spurious reflections than the four-step algorithm.

Other forms of coherent noise, such as dust and scratches on optical surfaces, inhomogeneities and imperfections in the optical elements and coatings, can cause troublesome interference. A moving diffuser close to the source point can also average out some coherent noise. Scrupulous cleaning of the optics can further reduce the scattered light, improve contrast, and reduce artifacts. Judicious use of quarter-waveplates in a polarization interferometer can minimize unwanted surface reflections and allow balancing of reference and test arm intensities (Bruning and Herriott, 1970).

14.9.8. Interferometer Optical Errors

The overall optical quality of the interferometer and its components impacts the accuracy of the measurement due to wavefront shear. This is a result of the fact that rays from an imperfect wavefront do not retrace themselves even when reflected from a perfectly flat or spherical surface. When the rays do not retrace themselves they

shear. The measurement error introduced by wavefront shear becomes greater with increased wavefront slope errors in the interferometer. Because of inevitable residual aberration, the best results are obtained with fringes that are nulled as completely as possible, by minimizing focus and tilt fringes.

It is difficult to give general formulas that quantify the effect of optical errors, but larger slopes give rise to greater error, in many instances, at quadratic or faster rates. The precise optical design of the interferometer must be ray traced to accurately characterize the effect. Any effect that introduces more slope for a given shear or more shear for a given slope will introduce greater error. Selberg (1987, 1990a) has calculated some of these errors for a Twyman–Green interferometer for several particular configurations. He introduced various amounts of tilt, power and spherical aberration in a simulated test surface and tabulated the calculated measurement error. Table 14.10 gives a summary of the results for 1 wave and 10 waves of peak-to-valley (P-V) error in tilt, plus and minus power errors (without tilt) and spherical aberration of the surface (without tilt and focus error) for test parts approximately 100 mm in diameter. Note that for a spherical surface, tilt and focus errors due to set-up can be removed by adjustment, whereas spherical aberration cannot. The simulation shows that these errors can seldom be combined analytically due to the nonlinearities. Specific situation must be simulated in their entirety, particularly when aberrations are large. Surface errors that return rays away from the axis introduce larger errors as indicated by the result for convex power error in Table 14.10. It is important that the reference wavefront be reasonably similar in radius to the wavefront under test. The error introduced by an imperfect reference or interferometer is dependent on both the aberrations and the disparity in radius of the test and reference wavefronts.

Optical errors resulting from design, fabrication or alignment imperfections are inevitable. To attain the highest accuracies, we must find ways for minimizing or compensating these errors. If the errors can be characterized in one form or other, the interferometer can be calibrated. This important process depends on the interferometer configuration.

Geometric distortion in the interferometer imaging or viewing system will also introduce errors that are proportional to the slope of the wavefront under test. This error can result from slope errors in the test surface or as a result of defocus. For example a detector imaging system that has 5% distortion at the edge of the field will introduce 0.01 waves of error for one wave of defocus (Truax, 1988). This emphasizes the need to null the test fringes as much as possible prior to taking data.

Another source of wavefront measurement error occurs when the aperture or surface being tested is not imaged at the detector in sharp focus. This tends to create strong slope errors at the edge of the test aperture due to Fresnel diffraction.

TABLE 14.10. P-V Measurement errors due to test conditions.

Setup for test part error type	P-V Ray mapping error in waves for 1 fringe error	P-V Ray mapping error in waves for 10 fringe error
Tilt	<0.0001	0.007
Concave power	<0.0001	0.007
Convex power	<0.0001	0.017
Spherical aberration	0.0004	0.089

14.10. DETECTORS AND SPATIAL SAMPLING

In most PSI systems in use today, solid-state detector arrays are used to collect the required intensity frames. These detectors are chosen because the response at each pixel shows excellent linearity with intensity, there is no image lag between measured frames, and there is no geometrical distortion introduced by the sensor in the recorded interference pattern. The intensity measured at each discrete pixel site can be digitized and stored in computer memory for the PSI calculation. Tube-type sensors, such as vidicons, have been successfully used in PSI systems, but the system performance is usually degraded by detector nonlinearities, image lag and distortion.

14.10.1. Solid State Sensors

Solid-state sensors can be classified by their geometry either as area arrays or as linear arrays. Area arrays are used most commonly for interferometry and permit the measurement of a two-dimensional section of a surface or wavefront. The system spatial resolution is related to the number of pixels along each dimension of the sensor, and commonly available sensors have resolutions appropriate for television applications; typically about 500 by 500 pixels or less. Some newer sensors designed for machine vision applications or high-definition television have dimensions of about 1000 by 1000 or even 4000 by 4000 pixels. Linear arrays, on the contrary, measure only a one-dimensional trace across the part, but make up for this disadvantage by providing a higher spatial resolution along this line. Linear sensors containing over 7000 pixels are available today, so that measurements with extremely high spatial resolution can be obtained. The amount of data from these large linear arrays is small and easily handled when compared to the 100,000 or more pixels from a low-resolution area array. Because of advances in semiconductor fabrication, we can expect to see the data density of available sensors continuing to grow.

There are many different configurations and technologies used in the manufacture of solid-state sensors including frame-transfer and interline-transfer charge coupled devices (CCDs), charge injection devices (CIDs), and photodiode arrays. It is beyond the scope of this chapter to fully review this technology (see Hall, 1980 or Janesick, 2001), but any of these sensors can be used for PSI. For practical considerations, the major differences between cameras with these various solid-state sensors are resolution, sensitivity to light, dynamic range, and data output format (Prettyjohns, 1984). The resolution is related to the number of pixels on the sensor, and the sensitivity is a function of both the quantum efficiency and geometry (the percentage of the imaging area that is occupied by photo-sensitive pixels as opposed to storage and transfer registers and control lines). There are also differences in camera performance as measured by the signal-to-noise ratio of the output.

For area arrays, there are two primary formats for data output: interlaced and line sequential. Sensors designed for television applications produce a video signal where each video frame is composed of two interlaced video fields. The odd numbered lines of the frame are output during the first field, and even numbered lines occur in the second field. A common US-video standard requires that there be 242 active lines per field or

484 active lines per frame, and the frame and field rates are fixed at 1/30 s and 1/60 s. This interlaced signal can be a problem for PSI systems as the two fields are integrated at different times and at different locations. While the use of a series of video frames to record the measured interferograms is common with phase step data collection, it does not work with the integrating bucket strategy since the average phase shift δ_i is different for the two fields in each frame. One approach to accommodate either of the data collection schemes is to use only the odd or even fields of the video signal to record the series of interferograms and to ignore the other field. This results in 242 usable lines of data. Some video cameras provide options where only one of the two fields is output in a repeat field mode. Cameras using the line sequential output, such as those produced by EG&G Reticon and others, do not have an interlaced output, and every frame can be used. These cameras usually consist of a square array of pixels (128 by 128 and 256 by 256 are common) and can have an adjustable frame rate. A square array is more appropriate than the 4:3 television format, since for most applications, it is important to retain the geometric integrity of the interferogram and to have equal resolution in x and y . Because of all the complexities involved in designing a camera, no single sensor technology or configuration has emerged as best for this application.

14.10.2. Spatial Sampling

Regardless of the type of sensor used in the interferometer, certain aspects of the data collection must be carefully considered for their impact on the measured intensity data. All of these sensors consist of an array of pixels with a given active area and spacing. The interferogram is imaged onto the sensor, and the intensity pattern is averaged not only over the integration time of the sensor, but also spatially over this active area. As the sensor is read out, the analog voltage of the video signal corresponding to each pixel is digitized and stored in the computer memory. One number is assigned for each pixel. This process is repeated as the phase shift is varied to collect the required number of frames. For a PSI system built around a tube-type sensor without defined pixels, the effective pixel area is defined by the integration period of the analog to digital converter in the horizontal direction and the width of the scan lines in the vertical direction. The digitization process spatially samples the interferogram. The algorithms and implications associated with this spatial sampling process on the PSI system performance will be studied in the remainder of this section and in Section 14.13. In a real sense, the PSI process samples the interference pattern in three dimensions; time and the two spatial dimensions.

We will model the sensor as a rectangular array of rectangular pixels, as shown in Figure 14.28. The pixels have dimensions and spacings of a by b and x_s by y_s , respectively. The signal measured at a specific pixel (mx_s, ny_s) , where m and n are integers, is found by integrating the time-averaged interferogram intensity $I_i(x, y)$ over the active area of the pixel:

$$\hat{I}_i(mx_s, ny_s) = \int \int_{-\infty}^{\infty} I_i(x, y) \text{rect}\left(\frac{\alpha - mx_s}{a}, \frac{\beta - ny_s}{b}\right) dx d\beta \quad (14.149)$$

This the result is for the i th intensity frame where the rect function represents the active area of the pixel, and α and β are variables of integration. The fully sampled

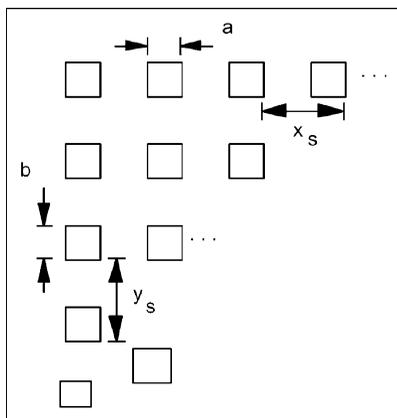


FIGURE 14.28. Illustration of a typical sensor geometry.

interferogram produced by the sensor is stored in a specific memory location and summoned over all pixels:

$$I_i^s(x, y) = \sum_m \sum_n \hat{I}_i(mx_s, ny_s) \delta(x - mx_s, y - ny_s) \tag{14.150}$$

where $\delta(x, y)$ is the two-dimensional Kronecker delta function. With some manipulation, Eqs. (14.108) and (14.109) are combined to yield the following expression for the *i*th sampled interferogram:

$$I_i^s(x, y) = \left[I_i(x, y) ** \text{rect}\left(\frac{x}{a}, \frac{y}{b}\right) \right] \text{comb}\left(\frac{x}{x_s}, \frac{y}{y_s}\right) \tag{14.151}$$

where ** indicates two-dimensional convolution and the comb function is an array of δ -functions positioned coincident to the pixels. Interpreting this expression, the convolution of the intensity with the rect function averages the intensity over the active area of each pixel, and the comb function yields each average to each pixel.

Since the interferogram intensity pattern is averaged over the pixel active area, it is reasonable to conclude that this process will result in a reduction of the modulation of the digitized interference pattern. This effect is completely analogous to the reduction in modulation that occurs with the temporal averaging that occurs in the integrating bucket data collection scheme referred in Equation 14.34. With point detectors, there is no loss of modulation because there is no averaging, but as the active areas get larger, the recorded modulation decreases. It should also be clear that the modulation reduction is a function of the spatial frequency of the fringes being recorded. For a given size pixel, greater loss will occur for higher frequencies than for lower frequencies; the higher the frequency, the greater the variation in intensity that occurs within the pixel. In fact, when a full fringe period exactly fits within the active area of a pixel (or the fringe frequency equals $1/a$) the recorded modulation is zero; the spatial average over the active area results in a constant that is independent of the position of the fringe relative to the pixel. This situation is equivalent to the temporal condition of $\Delta = 2\pi$ for the integrating bucket.

It is useful to relate the reduction of modulation with increasing fringe spatial frequency to spatial averaging in the frequency domain. The frequency-space representation of the spatially sampled interferogram is obtained through a Fourier transform of Eq. (14.110):

$$\tilde{I}_i^s(\xi, \eta) = [\tilde{I}_i(\xi, \eta) \text{sinc}(a\xi, b\eta)]^{**} \text{comb}(x_s\xi, y_s\eta) \tag{14.152}$$

where ξ and η are the spatial frequency coordinates, \sim indicates the Fourier transform, and the sinc function is the two-dimensional equivalent of the function defined in Eq. (14.35):

$$\text{sinc}(\alpha, \beta) = \frac{\sin(\pi\alpha)}{\pi\alpha} \frac{\sin(\pi\beta)}{\pi\beta} \tag{14.153}$$

Equation (14.111) states that the sampled spectrum is formed by first multiplying the spectrum of the time-averaged interferogram by the sinc function corresponding to the pixels. This filtered spectrum is then replicated at each multiple of the two-dimensional sampling frequency ($1/x_s, 1/y_s$). Since the average over the pixel represented by the sinc function serves to low-pass filter the interferogram intensity incident on the sensor, a convenient name for this term is the pixel MTF:

$$\text{Pixel MTF} = \text{sinc}(a\xi, b\eta) \tag{14.154}$$

The reduction in recorded signal modulation due to the non-zero size of the pixels can be found by evaluating this function at the fringe frequency. As mentioned above, the first zero of this function occurs at a spatial frequency equal to $1/a$ (along one dimension and scaled to the plane of the sensor). The pixel MTFs for sensors with pixels with a width to pitch ratios of 50% and 100% are shown in Figure 14.29. If non-rectangular pixels are used, these expressions can be modified, and functions other than a rect and a sinc will appear.

The other major consequence of a sampled imaging system is its limitation on resolution. The limiting resolution of a sampled system is the Nyquist frequency

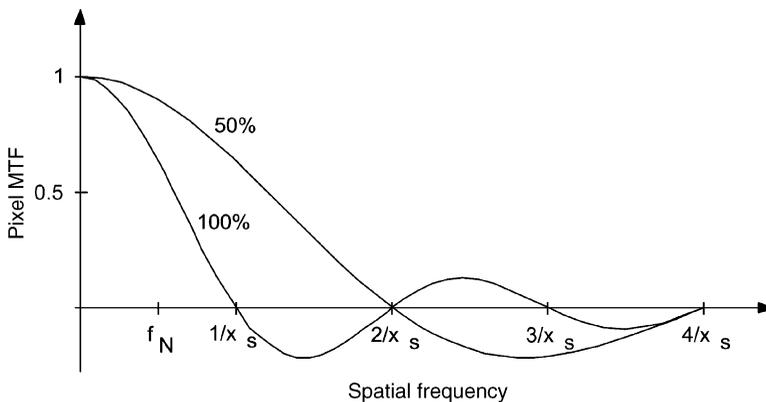


FIGURE 14.29. The pixel MTFs for sensors with width-to-pitch ratios of 50 and 100%.

which is defined as half the sampling frequency:

$$f_N = \text{Nyquist frequency} \equiv \frac{1}{2x_s} \tag{14.155}$$

where this is measured along the x axis (Gaskill, 1978). The vertical Nyquist frequency is often different and is dependent on the vertical pixel pitch. The reasoning behind this limit is that to measure a fringe of a particular frequency or period, at least two samples per period must be measured, i.e. we must be capable of measuring the high and low points on a fringe. If the fringe frequency exceeds the Nyquist frequency, the aliasing results of this sampled interferogram are not interpretable by standard PSI techniques. Aliasing and methods to get around this limitation are discussed in Section 14.13.

14.11. QUALITY FUNCTIONS

There is one more operation that must be performed to the calculated phase $\phi(x, y)$ before it is ready to be displayed and evaluated. We must correct for the discontinuities that occur in the phase calculation as a result of the arctangent, since the standard arctangent is defined only over the limited range of angles, $-\pi/2$ to $\pi/2$. Regardless of the actual value of the phase, only values of the phase within this limited result from Eq. (14.27), or its equivalent for other algorithms. This limitation might appear to restrict our ability to measure OPDs longer than a half a wavelength. Fortunately, there is sufficient information in the calculation to remedy the situation and provide a usable extended measurement range. In addition, we know that the wavefront or surface is in fact continuous and can extend over a much larger range.

The first correction to be made to the calculated phase is to extend the calculation range from 0 to 2π . This is possible because the signs of the sine and cosine are known independently of the sign of the tangent. Eq. (14.24) and (14.25), for the four step algorithm, are directly proportional to the sine and the cosine. Similar relationships exist for the other algorithms that are discussed in this chapter. Table 14.11 gives the formulas needed to convert the arctangent results to values between 0 and 2π as a function of the values of the sine and the cosine, and this process is shown graphically in Figure 14.6. The result of this correction is to produce the wavefront phase modulo 2π ; every time the actual phase equals a multiple of 2π , the calculated

TABLE 14.11. Modulo 2π phase correction.

Sine	Cosine	Corrected phase $\phi(x, y)$	Phase range
0	+	0	0
+	+	$\phi(x, y)$	0 to $\pi/2$
+	0	$\pi/2$	$\pi/2$
+	-	$\phi(x, y) + \pi$	$\pi/2$ to π
0	-	π	π
-	-	$\phi(x, y) + \pi$	π to $3\pi/2$
-	0	$3\pi/2$	$3\pi/2$
-	+	$\phi(x, y) + 2\pi$	$3\pi/2$ to 2π

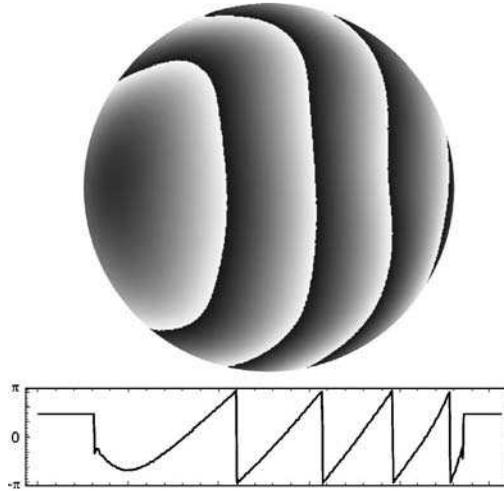


FIGURE 14.30. Raw or wrapped phase calculated on the 4 interferograms shown in Figure 14.6.

value returns to a value of zero. Usually computers already support this calculation by the use of an $\arctan(y/x)$ function sometimes designated as a \tan^{-1} . For simplicity, this corrected phase is here referred to as the raw phase, wrapped phase or the phase modulo 2π . Note that Table 14.11 is used only if the values returned by the arctangent lie between $-\pi/2$ and $\pi/2$. Figure 14.30 graphically shows the wrapped phase map calculated from the four interferograms shown in Figure 14.6.

The previously discussed methods to evaluate the phase can only be used if there is actual modulation of the intensity pattern. If detector pixels are outside the interferometer aperture or if they are covered by dust or are otherwise disabled, they deliver invalid data. This usually shows up as a constant value or very low modulation for that particular pixel. Where all intensity values I_i have approx. the same values, the numerator and denominator would become very small numbers and approach the singularity of the arctan function where the phase is undefined.

$$\tan \phi = \frac{N}{D} \quad \arctan\left(\frac{0}{0}\right) = \text{undefined} \quad (14.156)$$

To prevent this several metrics are used to identify valid and invalid pixels for evaluation and assign a quality function to each evaluated pixel. This is an important step before performing the unwrapping process on the raw phase.

14.11.1. Modulation

In Eqs. 14.29 and 14.65 we introduced the per pixel modulation or visibility $\gamma(x, y)$, which is a very good quality indicator for the data associated with each pixel,

$$\gamma(x, y) = \frac{I''(x, y)}{I'(x, y)} = \frac{2}{\sum I_i(x, y)} \sqrt{\left[\sum I_i(x, y) \cos(\delta_i)\right]^2 + \left[\sum I_i(x, y) \sin(\delta_i)\right]^2} \quad (14.157)$$

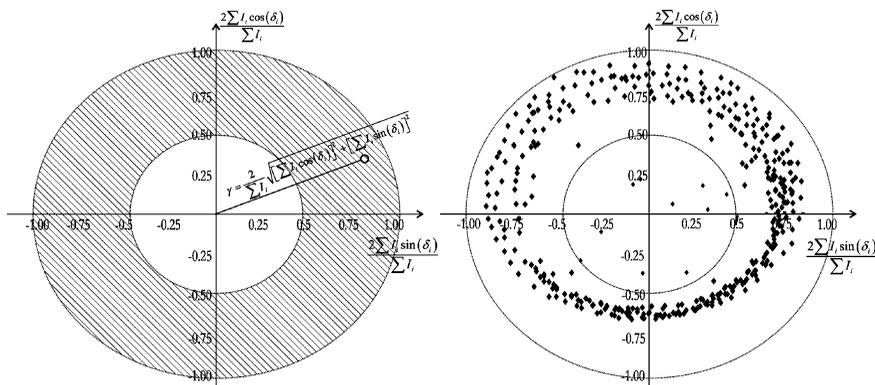


FIGURE 14.31. Modulation or Visibility plotted as the radius in a polar-coordinate system.

If the raw data is plotted as a vector in polar coordinates, the raw phase for each pixel is the vector angle and the modulation is the vector length. Figure 14.31 shows a schematic view of the polar coordinate system on the left and a real measured data set on the right. Note the marked thresholds and the deformation of the circle on which the measurement points are plotted. This is due to a miscalibration in the phase-shifter for this data set. Here, like phase evaluation, modulation evaluation also suffers from error sources like phase-shifter miscalibration, harmonic errors, and errors in the intensity linearity that can deform an otherwise ideal circle.

In 1994, Schwider showed that the modulation is sensitive to these same errors and that second-, and fourth-order harmonics to the phase occur when errors are present.

$$\widehat{\delta}_i = \delta_i + \varepsilon_i \Rightarrow \widehat{\gamma} = \gamma \frac{\sqrt{(1 + A^2 + C^2 + S^2) + (2S - 2AC) \cos 2\phi - (2C - 2AS) \sin 2\phi}}{1 + \gamma c \sin \phi - \gamma s \cos \phi}$$

$$A = \frac{1}{N} \sum_i \varepsilon_i, S = \frac{1}{N} \sum_i \varepsilon_i \sin 2\delta_i, C = \frac{1}{N} \sum_i \varepsilon_i \cos 2\delta_i, s = \frac{1}{N} \sum_i \varepsilon_i \sin \delta_i, c = \frac{1}{N} \sum_i \varepsilon_i \cos \delta_i$$

(14.158)

Schwider also demonstrated that depending on the detector used, the modulation can be utilized to determine under-sampled pixels as the modulation decreases. The modulation decrease depends on the fill factor of the detector. If the preferable fill factor of 1 is used, modulation decreases by 36% at the Nyquist limit.

Kinnstaeter et al. 1988 demonstrated that the use of Lissajous plots as in Figure 14.31 are a very powerful graphical debugging tool quickly identifying error sources and the quality of a measurement.

14.11.2. Residues

One condition often used to determine noise and to identify critical conditions for the unwrapping process is the evaluation of residues in the wrapped phase map

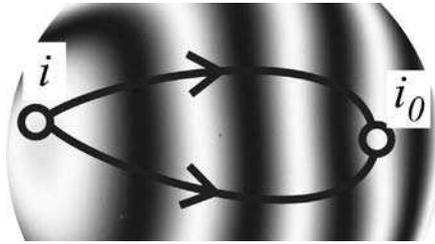


FIGURE 14.32. Different paths in a continuous phase map.

(Goldstein et al., 1988; Huntley, 1989). The basic idea behind the residue approach is that between two arbitrary points, i_0 and i , in a continuous interferogram pattern, the amount of fringes does not depend on the path taken.

$$\phi(i) = \int_C \nabla\phi \, di + \phi(i_0) \tag{14.159}$$

Figure 14.32 tries to illustrate the different paths along a continuous interferogram.

$$\oint \phi(r)dr = 0 \tag{14.160}$$

Equation (14.160) is the main criterion used to detect residues. Every closed loop on a continuous interferogram has to result in zero phase jumps. If it is different, then a local residuum is present.

Figure 14.33 illustrates a function with a singularity in the center. If the amount of phase discontinuities is counted going left or right around the center, it is different by one.

$$\oint \phi(i)di = \pm 2\pi \cdot \sum \text{enclosed residues} \tag{14.161}$$

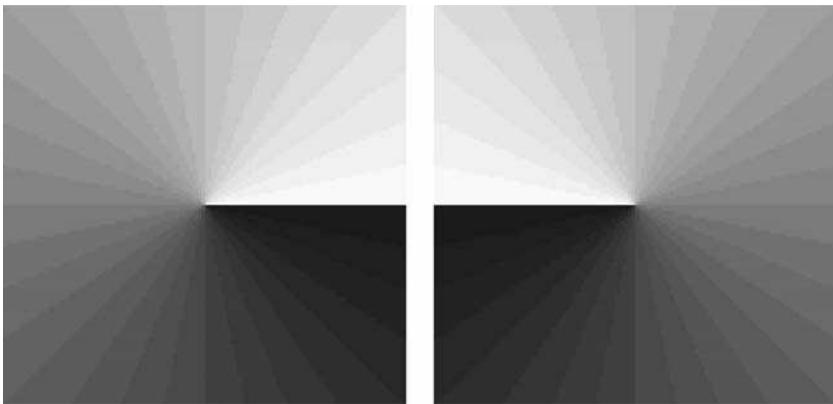


FIGURE 14.33. A negative (ccw, left) and positive (cw, right) residuum pictured.

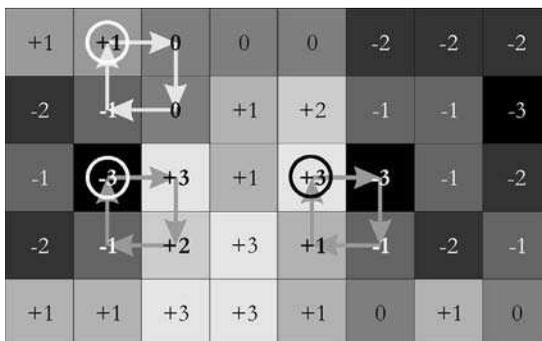


FIGURE 14.34. Different closed loops in a simulated wrapped phase map. The numbers are corresponding to the phase value in rad. A phase discontinuity is present if the difference between neighbors is larger than π : no singularity (left top) positive singularity (left bottom) and negative singularity (right bottom).

As described in Eq. (14.161), the closed loop unwrapping results in either zero or the sum of residues enclosed in the path chosen. The residues can be positive or negative depending if there is a positive or negative singularity present. As Goldstein et al. and Huntley point out, residues in a phase map occur in pairs and can be correlated to each other.

Being able to correlate two residues with each other allows one to block the path in between. As pictured in Figure 14.33, since a branch-cut prohibits a path between the two residues (only a path around them is permitted), a continuous phase condition in Eq. (14.160) does in fact exist.

Figure 14.35 shows a real measurement containing a localized residue. In this case the lens under test contained optical activity, and due to the polarized beam passing through the lens and interferometer, the detected interferogram saw a loss in contrast at the location of the residue and an inversion of the fringes after the residuum. Still, the phase in the interferogram can be evaluated if the residue is isolated.

Figure 14.36 shows on the top a very noise wrapped phase map containing many residues. Phase maps like this can happen if a large amount of coherent noise is present as in speckle interferometry. However, even for complicated cases such as this, methods have been developed to reduce the amount of residues through filtering.

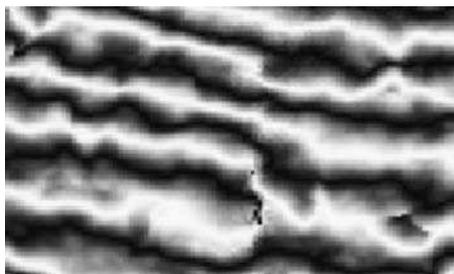


FIGURE 14.35. Measured phase map containing residues.

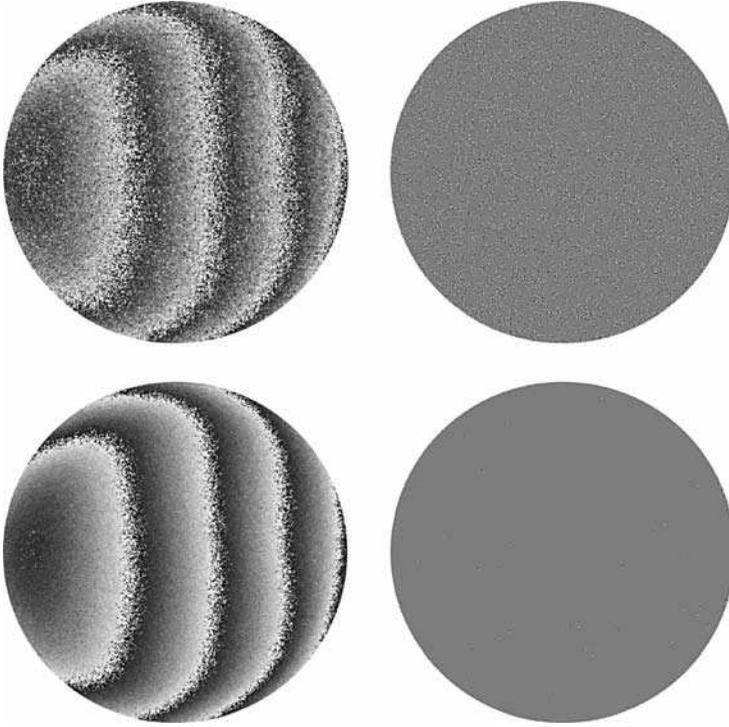


FIGURE 14.36. Noisy wrapped phase images and associated residues (top). Filtered interferogram and reduced number of residues (bottom).

14.11.3. Filtering

A noisy wrapped phase map as shown in Figure 14.36 can benefit greatly from filtering before the quality functions are calculated. However, if filtering is performed on the wrapped phase data itself, discontinuities would be spatially smeared.

There is however a simple workaround to the discontinuity problem which is apparent from the definition of the tangent function.

Since the tangent function is a discontinuous function of the calculated phase ϕ from the quotient of two numerically determined functions that we know are everywhere continuous with ϕ , it is most appropriate to independently filter the numerator and denominator before making the quotient.

$$\tan \phi = \frac{\sin \phi}{\cos \phi} \approx \frac{\text{filter}(\sin \phi)}{\text{filter}(\cos \phi)} \quad (14.162)$$

Using this technique does not affect the phase jumps in the later calculated signal, that would otherwise be polluted by low-pass filtering. Several approaches for the best filter to be used, depending on the application, have been published by

Bone et al. (1986), Huntley (1989), Vikhagen, 1990, Goncalves et al. (1993) and Seara et al. (1998).

Figure 14.36 bottom shows, a low-pass filtered version of the top Figure. The right handed side shows how dramatically the filtering process reduces the number of residues and hence the complexity of the later unwrapping process also. Bushman et al. (1993), Pritt (1994), and Friedlander and Frances (1996) have proposed methods to filter and unwrap the phase simultaneously.

Filtering can also be done, after the continuous phase map has been calculated; however, a priori knowledge of the interferogram can be helpful in the filtering process.

14.12. PHASE UNWRAPPING

The final step in the wavefront reconstruction process is to remove the 2π discontinuities that are present in the raw phase data that has been generated. This process is referred to as phase unwrapping, phase integration or phase continuity, and it converts the modulo 2π phase data to a continuous representation of the wavefront under test. Whenever a large discontinuity occurs in the reconstruction, 2π or multiples of 2π are added to the adjoining data to remove the discontinuity. This process is diagramed for one-dimensional data in Figure 14.37. The effects of phase unwrapping operation in two dimensions are shown in Figure 14.8.

14.12.1. Unwrapping in One Dimension

The basics of the unwrapping process are described in Itoh 1982. In principal the unwrapped phase is calculated by adding a multiple of 2π to the wrapped phase Eq. (14.115). The process of the unwrapping is to determine the unknown piston function $p(i)$ of the multiples of 2π .

$$\phi_{unwrapped}(i) = \phi_{wrapped}(i) + p(i) \cdot 2\pi \quad (14.163)$$

An aid to determine $p(i)$ is to examine the phase differences along the path i , where $i - 1$ corresponds to the pixel index previous to i . As can be seen in Figure 14.38, $\Delta\phi$ has a distinct spike at every phase discontinuity.

$$\Delta\phi_{wrapped}(i) = \phi_{wrapped}(i) - \phi_{wrapped}(i - 1) \quad (14.164)$$

Using $\Delta\phi$ as a help, the function $p(i)$ can be defined recursively in Eq. (14.117) by increasing or decreasing the piston based on the direction of the spike in $\Delta\phi$. The threshold for when $\Delta\phi$ shows a phase discontinuity or just a strong phase gradient is usually set to π because at a fringe frequency at the Nyquist limit, the phase differences are exactly π . Therefore, everything below the Nyquist limit is determined to be a phase gradient, and everything above is identified as a phase discontinuity and needs unwrapping. When the recorded interferograms satisfy the

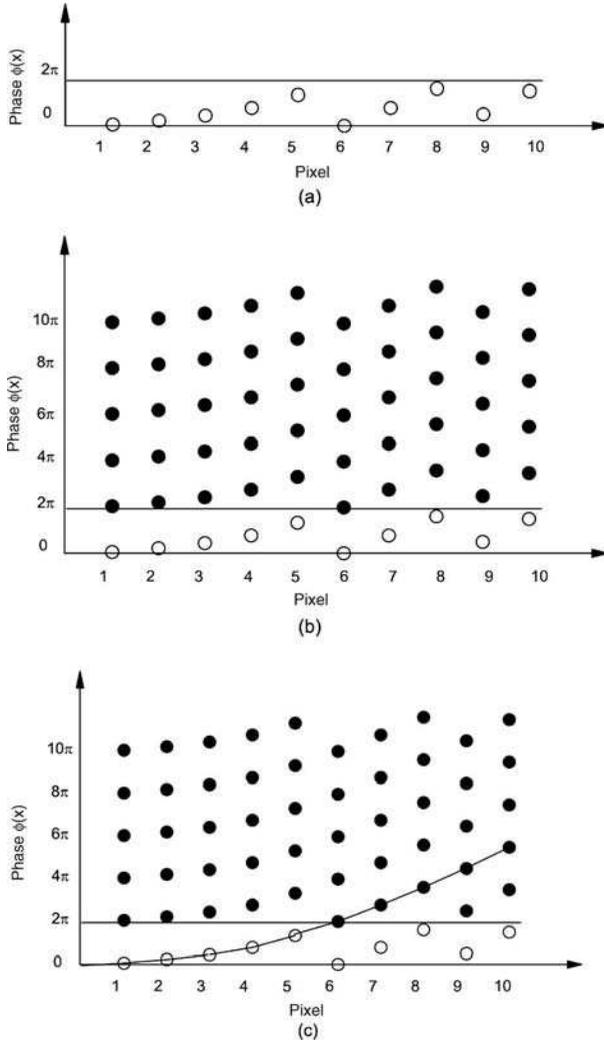


FIGURE 14.37. The PSI phase unwrapping process: (a) the wavefront data modulo 2π at each pixel; (b) all possible solutions for the wavefront phase; and (c) the reconstructed wavefront.

Nyquist criteria, namely there are at least two pixels per fringe period, the phase unwrapping process is straight forward. Having at least two pixels per fringe implies that the wavefront phase changes by no more than π per pixel spacing. Therefore this criteria is used to reconstruct the missing piston term.

$$p(i) = \begin{cases} p(i-1) - 1 & \Delta\phi(i) < \pi \\ p(i-1) & -\pi < \Delta\phi(i) < \pi \\ p(i-1) + 1 & \Delta\phi(i) < -\pi \end{cases} \quad (14.165)$$

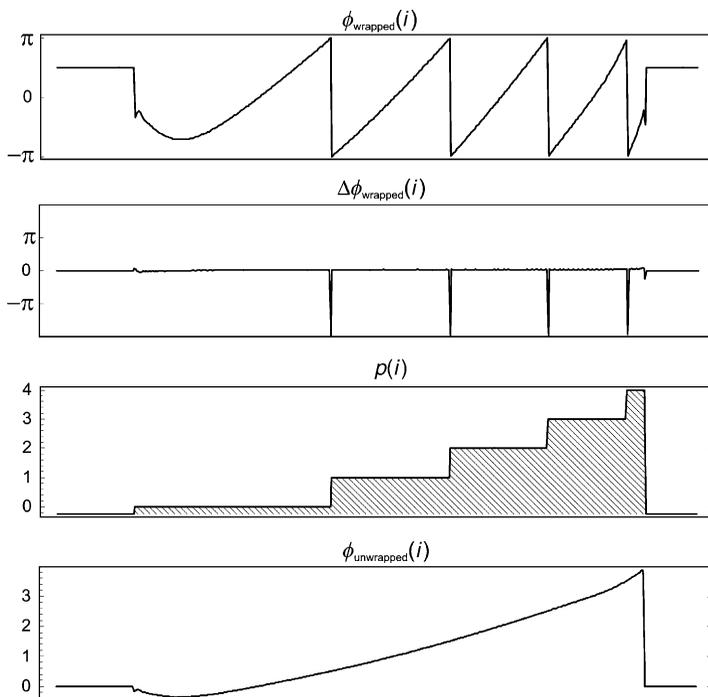


FIGURE 14.38. Illustration of the unwrapping process and the different functions in one dimension. The phase is shown in radians, while the piston function and the unwrapped phase are displayed in multiples of a fringes or 2π .

Starting at some location in the wavefront, normally near the center of the interferogram, we require that the phase between any two adjacent pixels does not change by more than π .

If the phase difference calculated for two pixels exceeds π , then 2π or multiples of 2π must be added to or subtracted from the calculated value of the second pixel until this condition is met. The entire wavefront map is calculated by working outward from the starting location therefore setting $p_{\text{start}} = 0$.

14.12.2. 2-D Phase Unwrapping

While the one-dimensional unwrapping process leaves little room for variation, the two dimensional problem is significantly more complex. In one dimension, if there is no valid data along the path, it is not possible to unwrap without additional knowledge. In two dimensions, there are multiple paths to unwrap the data, and the results achieved may depend on the path taken. When dealing with noisy or corrupted data, the unwrapping “problem” becomes quite significant and many publications deal with solutions to specific applications.

After Itoh 1982, several publications by Goldstein et al. (1988), Huntley (1989), Huntley et al. (1995), Ghiglia and Eichel (1994), Ghiglia and Romero (1996), Flynn

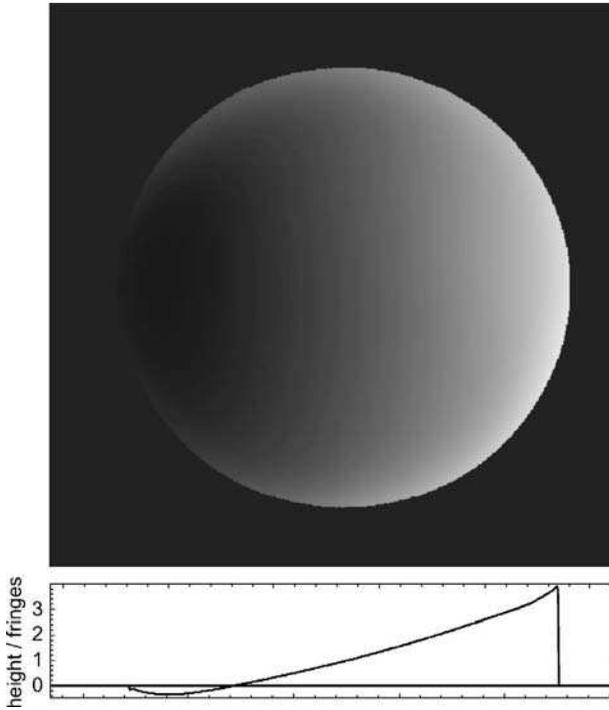


FIGURE 14.39. Unwrapped phase map of the sample shown earlier.

(1996), and Chen et al. (2000) deal with the unwrapping problem and make use of the increased computing power available today to make these evaluations.

We give here a brief overview of the two-dimensional unwrapping and invite the reader to review Ghiglia's and Pritt's book "Two dimensional phase-unwrapping" (Wiley, NY 1998), for specific examples and more in-depth study.

In principal the families of algorithms can be divided into path-following algorithms and path-independent algorithms.

14.12.3. Path-Following Algorithms

The straight forward approach to phase unwrapping presented by Oppenheim and Schafer, 1975 consists of simply using a one-dimensional description and applying it to the two dimensional problem. Figure 14.40 shows one possible solution. By starting from a known valid starting point, a one dimensional unwrapping is performed progressive towards the edge of the aperture. Then orthogonal to this unwrapped line, further one-dimensional lines are unwrapped toward the edges of the aperture. Other scan patterns have been proposed such as spirals or even fill algorithms can be employed to get around masked and invalid pixels. For well-behaved, low noise data, this method works quite well. However, if there are noisy or invalid pixels in the path, all pixels thereafter retain the error introduced, resulting in

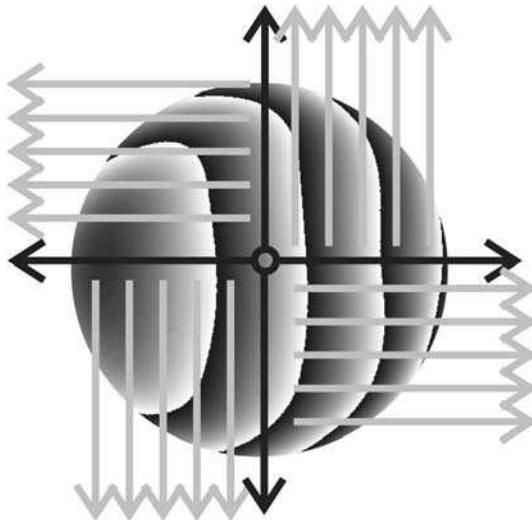


FIGURE 14.40. Straight forward adaptation of the one-dimensional approach into two dimensions.

an erroneous piston function. This method applied to two dimensions is usually fast but quite vulnerable to noise-propagation. Therefore, many unwrapping algorithms have been proposed to improve noise-immunity or reduce error-propagation.

Cut-Based Algorithms. The concept of branch-cuts was first introduced in section 14.11.2. Introduced by Goldstein et al. (1988) and later refined by Huntley (1989), these algorithms use the information obtained by identifying positive and negative residues and correlating them to each other. By introducing branch cut's, they prohibit unwrapping paths between these residues, forcing the algorithms to find an alternate path. Determining and correlating residues to generate the proper unwrapping paths does take additional computing effort but generally delivers better results than the simpler one-dimensional methods.

Quality Guided Path. Ching et al. (1992), Quiroga and Bernabeu (1994), Quiroga et al. (1995), and Herráez, 1996 present algorithms that focus on how the unwrapping path is chosen and have introduced criteria for path propagation. These algorithms constrain path propagation based on one or multiple metrics that give a measure of the reliability of paths chosen (Figure 14.31). For example, the modulation γ can be used to determine a path that focuses on the pixels with the highest contrast first and eventually unwrapping pixels with lowest modulation last. Another metric proposed is the phase gradient $\Delta\phi$, which diverts unwrapping over the phase jump with the largest phase difference first assuming it is the most likely one to be a phase jump and not being a residue or noise related.

The quality of the result depends significantly on the metric used to determine the reliability of these values.

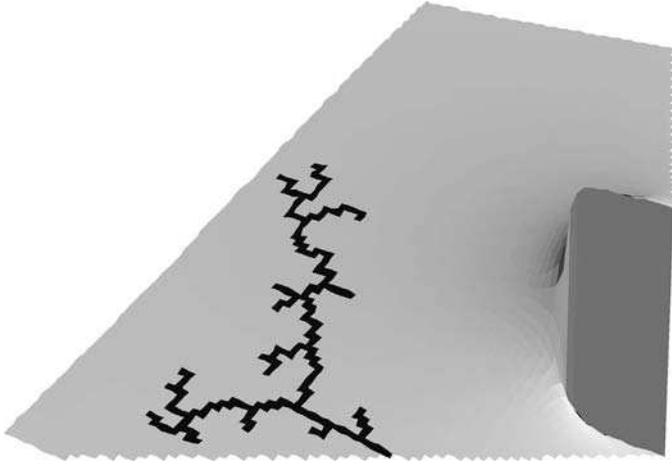


FIGURE 14.41. Quality guided path algorithm propagating along a minimum slope of the wavefront.

14.12.4. Path Independent Methods

In contrast to the methods on the path, there are other approaches that avoid choosing a path entirely and instead treat the measured phase map and unwrapping process on the data set as a whole.

Partitioning Algorithms. Gierloff 1987 introduced the concept of “unwrapping by regions”. Here as we see in Figure 14.42, the wrapped-phase map is first clustered into regions separated by the phase jumps. The algorithms then tries to determine the offset of whole regions to each other and tries to find the best balance to eliminate all the jumps in between them.

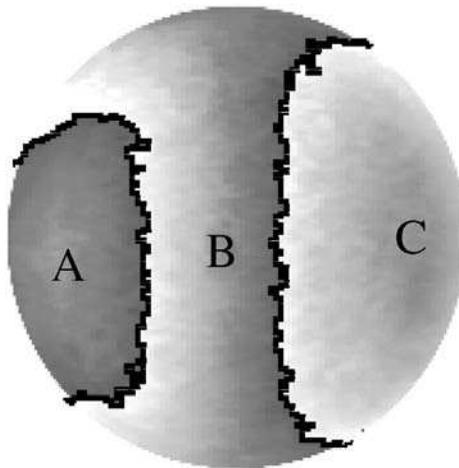


FIGURE 14.42. Wrapped phase map with three regions A,B, and C separated by the phase discontinuity.

Similar algorithms, like the one from Herráez 2002, use image decomposition methods to define and unwrap the regions.

Another partitioning approach is to divide the wrapped phase-map into smaller grids of smaller areas that are unwrapped individually as shown by Towers et al. (1989), Herráez, 1996 and Baldi (2001). The smaller unwrapped tiles are then combined and unwrapped with respect to each other.

Cellular Automata Algorithms. Ghiglia et al. 1987 proposed a cellular automata method based on simple neighborhood rules. The method tries to iteratively minimize the discontinuities in the phase map by adding or subtracting multiples of 2π to the connected pixels. While quite good in minimizing the amount of overall discontinuities, invalid pixels and noise are still able to propagate; however, the amount is minimized depending on the rules. On standard RISC/CISC type microprocessors, the cellular automata method is quite long in execution time, because its parallel approach can not be exploited. However, there are also cell-processors planned today that would allow a much faster implementation of that technique.

Temporal Phase Unwrapping. Huntley and Saldner (1993) proposed a technique that would avoid the actual unwrapping by looking at a dynamic signal, such as these that might be present in mechanical deformations. Starting with zero deformation, each new step introduces deformation that is not large enough to contain discontinuities. While the deformation in the final result might produce discontinuities, each small dynamic step would not. In this way, the final result is reached by splitting into small incremental steps. This process of introducing small deformations to the phase change with minimal error propagation makes up the incremental unwrapping process.

14.13. ASPHERES AND EXTENDED RANGE PSI TECHNIQUES

From a testing point of view, the prime characteristic of an aspheric wavefront or surface is that it has a large departure from a best-fit reference sphere, and as a result, an interferogram made without some sort of aspheric null contains many fringes (see also Chapter 12). The Nyquist limit implies that we must have at least two pixels per fringe which sets a limit to the amount of asphericity that can be measured with a PSI system. The maximum wavefront slope measureable without aliasing is thus limited to π per pixel. The spatial fringe frequency is proportional to the wavefront slope, and even weakly aspheric surfaces often violate this condition when tested against a spherical reference surface with instrumentation available today. Typically, we are limited to testing surfaces with no more than 10–20 waves of asphericity. A precise number is not predictable since the maximum fringe frequency is related to slope, not the peak-to-valley departure for a reference sphere.

There are typically four approaches available to test aspherics with PSI: (1) Use a null optic (refractive, refractive or holographic) to reduce the amount of asphericity in the wavefront; (2) Use of a detector array with greater pixel density so that the

sensor Nyquist frequency is increased; (3) Use of a longer wavelength (such as $10.6\ \mu\text{m}$ from a CO_2 laser) to rescale the wavefront departure to the longer wavelength; or (4) Testing sub-apertures of the wavefront with a small enough segmentation to sufficiently resolve the individual subapertures. None of these approaches is entirely satisfactory as each may involve trade-offs which often place long lead times on the design of the test, requires additional fabrication, increases the difficulty of using and calibrating the instrument, decreases its precision, or greatly increases the instrument cost. In order to test aspheric surfaces without resorting to one of these four approaches, a PSI system must be capable of handling fringe densities above the Nyquist frequency, i.e., those that are aliased. Two useful techniques that have been proposed and demonstrated for analyzing large departure surfaces or aliased interferograms are sub-Nyquist interferometry and two-wavelength PSI.

14.13.1. Aliasing

Before describing either of these extended range techniques, it is useful to discuss aliasing in more detail. Aliasing is the property of a sampled imaging system which causes high-frequency content in the input to be lost or obscured and displayed at lower spatial frequencies, that is, at frequencies less than the Nyquist frequency. This property can be graphically demonstrated in Figure 14.43 where two different spatial frequency inputs are sampled. For illustration, infinitely small pixels are assumed and the dots represent sampled values at the pixels. In Figure 14.43(a), the fringe frequency is two-third the Nyquist frequency, or the sampling rate is three pixels per fringe or cycle. The sampled output is clearly at the same frequency as the input. The fringe frequency in Figure 14.43(b) is four-third the Nyquist frequency, with only three pixels for every two fringes. There are insufficient pixels to resolve this frequency, but the recorded samples are exactly the same as for the lower frequency measured in (a). Since the detected outputs in (a) and (b) are identical, the displayed output in (b) occurs at a frequency of $2f_N/3$ instead of $4f_N/3$. The effect of aliasing on the displayed fringe pattern is to replace fringes of spatial frequency greater than the Nyquist frequency with lower frequency fringes and corrupts the normal interpretation of the fringe pattern.

More importantly for PSI, aliasing is an indication that the wavefront phase is changing at a rate greater than π per pixel. Under this condition, the phase unwrapping operation described so far in Section 14.12 must fail as this process restricts the wavefront change to π per pixel.

This failure is predicted by the Whittaker–Shannon sampling theorem which states that if a scene is bandlimited to within the Nyquist frequency of the sensor, the scene can be recovered without error from the sampled image. The default PSI phase unwrapping operation is based upon the assumption that this bandlimited requirement is met. It is important to note that this theorem does not state anything about the possibility of image recovery when this condition is not met. If the input scene or fringe pattern is not bandlimited to the Nyquist frequency, some additional information or *a priori* knowledge about the wavefront or surface under test is required to properly recover and unwrap the phase.

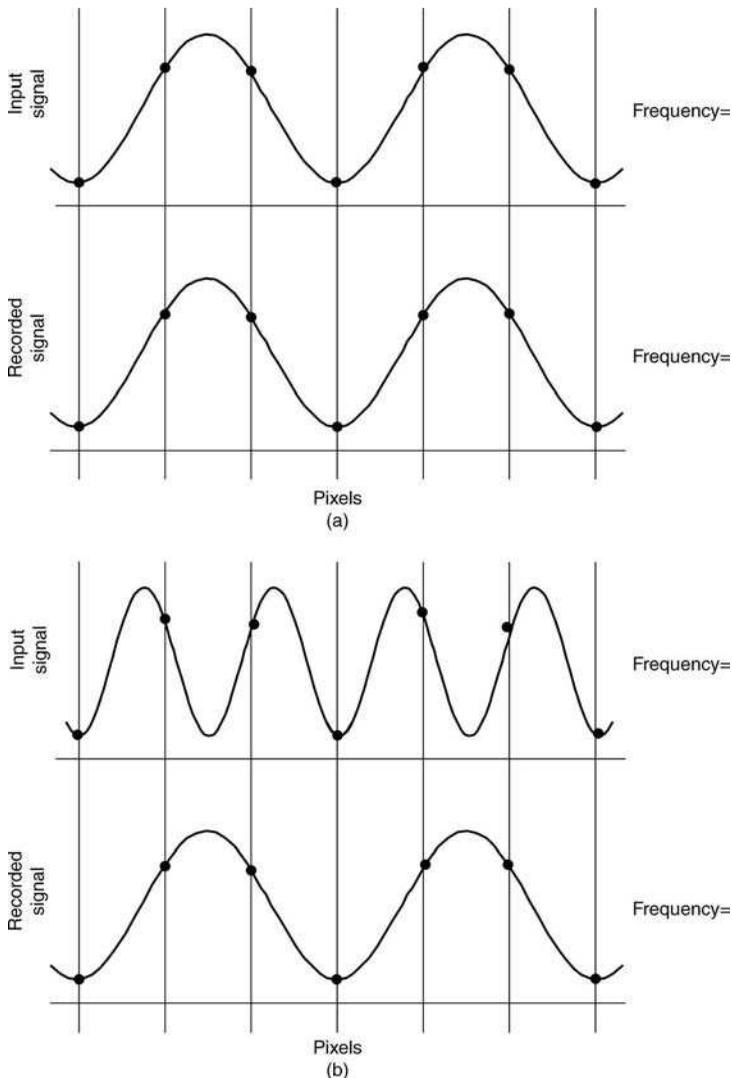


FIGURE 14.43. Sampling and aliasing in one dimension: (a) the input frequency is less than the Nyquist frequency and (b) the input frequency exceeds the Nyquist frequency.

14.13.2. Sub-Nyquist Interferometry

Sub-Nyquist interferometry (SNI) is a data collection and analysis method that is capable of greatly extending the measurement range of PSI through the use of *a priori* information (Greivenkamp, 1987b; Greivenkamp, 1990). SNI applied to aspheric testing makes use of the simple assumption that the wavefront or surface under test is smooth and continuous, and therefore has continuous derivatives. This

additional information allows the analysis to interpret fringes that occur at frequencies well above the Nyquist frequency and are significantly undersampled. The differences between SNI and PSI occur during the phase unwrapping of the modulo 2π data. Instead of using the usual wavefront height constraints of PSI, SNI requires that the derivatives of the reconstructed wavefront do not exhibit large changes from pixel to pixel. This slope continuity constraint limits the change of the wavefront slope to π per pixel, and large changes of the wavefront height between pixels are permitted. The appropriate number of 2π 's are added to each pixel to satisfy this condition, and there is only a single solution at each pixel that produces this result. The slope continuity constraint correctly reconstructs the wavefront from the aliased data until the second derivative of the actual wavefront exceeds the limit imposed by the constraint. When this situation arises, further correction is possible by requiring that the second, or even higher-order, derivative be continuous, and adding more 2π 's. This procedure can be continued until a more fundamental limit, as explained below, is reached. In practice however, first derivative or slope continuity is very effective in improving the measurement range of PSI, and higher orders are usually not needed.

The SNI phase unwrapping process for aspherics is graphically demonstrated in Figure 14.44 using the connect-the-dot pictures similar to those in Section 14.12. Figure 14.44(a) shows all of the possible solutions to the arctangent at each pixel, and the open circles indicate the phase modulo 2π . The asphere used to generate these points is shown as the dotted line, and it goes through one point at each pixel. The object of the reconstruction is to determine these correct dots. The conventional PSI reconstruction of these data is given in Figure 14.44(b), and this reconstruction fails at pixel 5 where the correct dot is not the closest dot; the wavefront changes by more than π per pixel. The reconstruction shows a large change in wavefront slope. The SNI reconstruction in Figure 14.44 (c) gives the correct result by applying the slope continuity constraint. The proper dot is chosen by extrapolating a line from the previous two dots into the next pixel, and the dot that is closest to this line gives the solution with the smallest change in wavefront slope. Any other choice would violate the slope continuity condition. Higher-order constraints operate by extrapolating a curve of the appropriate order that has been fit through a number of already selected dots. For example, second derivative continuity uses a quadratic fit through the previous three points.

The fundamental limit to the measurement range of an SNI system is in the ability of the sensor to respond to the high frequency fringes; the pixel MTF. For good results in the arctangent, the measured data modulation must be high, and the sensor must be able to respond to fringes well beyond the Nyquist frequency. From Eq. (14.154), we see that the extent of the pixel MTF is increased by using a sensor with a small pixel width to pitch ratio. This ratio for standard sensors is usually 50–100%, and the sparse array sensor that is needed for this application should have a ratio closer to 10%. The pixels on this sensor approximate point detectors, and the first zero of the pixel MTF is at a frequency 20 times the Nyquist frequency. A comparison to this pixel MTF to that of a standard sensor is shown in Figure 14.45. The maximum wavefront slope that can be measured with this sensor and SNI is up to 20

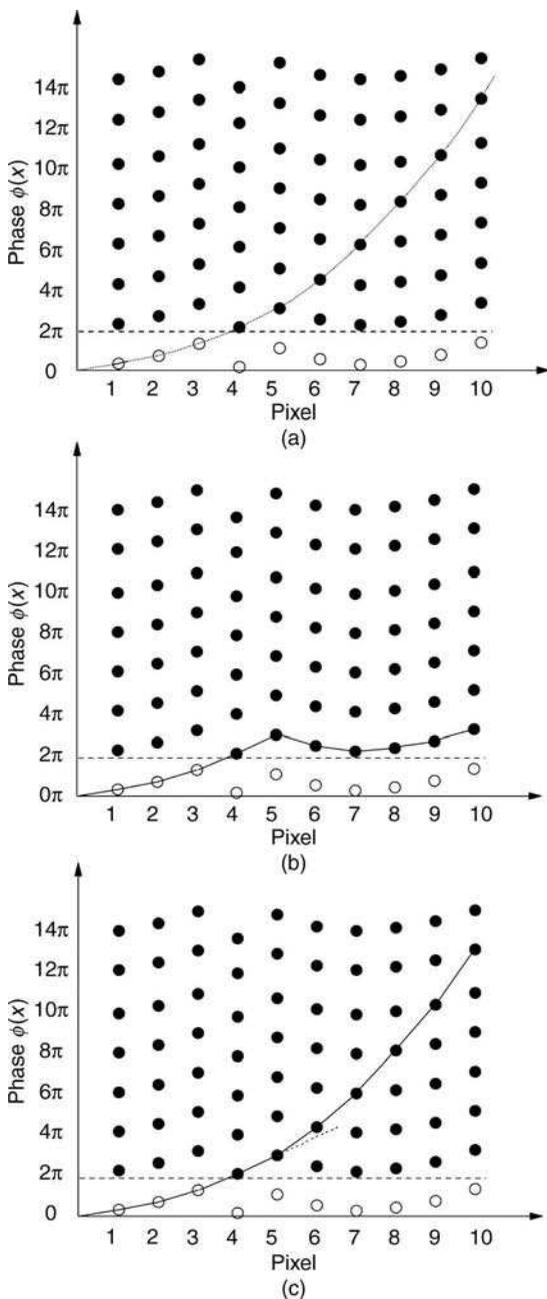


FIGURE 14.44. The SNI reconstruction process: (a) the possible solutions for the wavefront phase at each pixel; (b) the standard PSI reconstruction of this data; and (c) the SNI reconstruction using slope continuity (from Greivenkamp, 1987b).

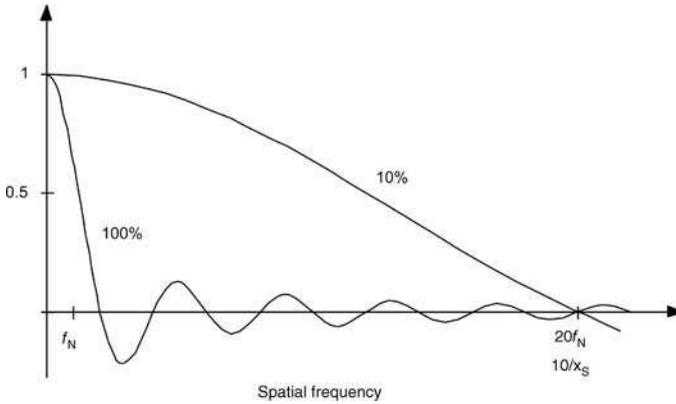


FIGURE 14.45. The pixel MTF of a sparse array sensor compared to that of a standard sensor.

times that which can be measured with PSI and the same number of pixels. The measurement range, which is dependent on the functional form of the wavefront, is increased by an even larger factor.

Some interferograms and evaluation are displayed in Figure 14.44. The interferogram of this wavefront is shown in Figure 14.46(a), and the undersampled interferogram as seen by a detector is shown in Figure 14.46(b) and is heavily aliased. The ring patterns surrounding the central ring pattern are aliased fringes. The center of each of these additional patterns occur when the actual fringe frequency equals a multiple of the sensor Nyquist frequency. As expected, the PSI reconstruction of this interference pattern in Figure 14.46(c) shows a good reconstruction only in the central unaliased region. The aliased fringes are improperly interpreted. The two-dimensional SNI result in Figure 14.46(d) smoothly reconstructs the quadratic wavefront. The horizontal streaks in this result are due to the simplicity of the algorithm used; the phase unwrapping algorithm has not been taught to jump over bad data points.

SNI allows the interpretation of very complex fringe patterns, and the primary issue that needs to be addressed before SNI can be practically implemented is calibration. The testing of aspheric surfaces without null optics can lead to errors.

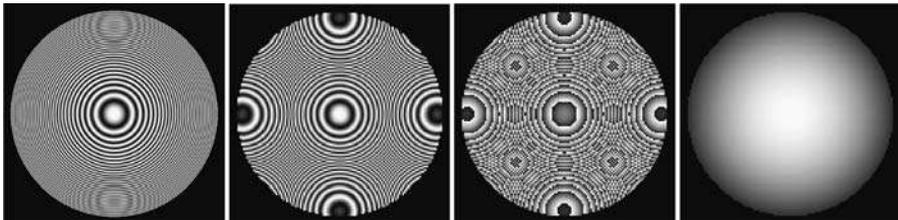


FIGURE 14.46. (a) The sampled interferogram of a defocused wavefront; (b) undersampled interferogram as seen by the detector; (c) the standard PSI reconstruction of this data; and (d) the SNI reconstruction using slope continuity.

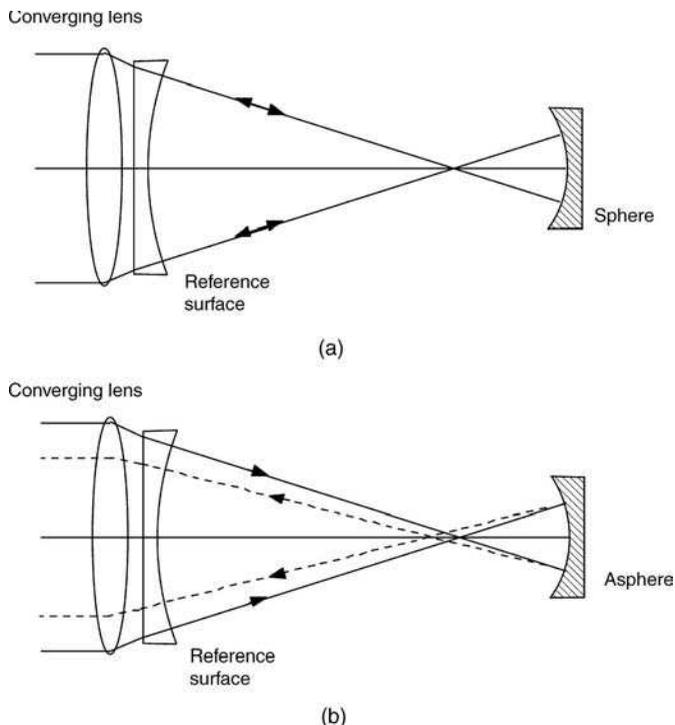


FIGURE 14.47. The need for calibration and raytracing: (a) testing a spherical surface at its center of curvature; (b) the same configuration with an aspheric surface.

In a null configuration, such as testing a sphere at its center of curvature in a laser Fizeau interferometer (Fig. 14.47(a)), the rays reflected by the test surface follow the same path through the converging lens and the interferometer as the rays reflected by the reference surface. Both the reference and test wavefronts see the same amount of additional aberration, and this aberration cancels. This is not the case in testing aspheric surfaces (Fig. 14.46(b)). These two ray paths can be very different, producing different system aberrations in the two interfering beams. A simple way to think about this is that different zones on the aspheric surface use the converging lens at different conjugates and different apertures, and it is only corrected at one conjugate. The wavefront difference measured in the interferogram is not simply the difference between the test and the reference wavefronts. An aspheric measurement system requires that the entire interferometric system, including all of the interferometer optics, be raytraced for the asphere under test.

14.13.3. Two Wavelength PSI

The additional information that two wavelength PSI (TWPSI) uses to extend the PSI measurement range beyond the Nyquist frequency is a separate measurement of the part at a different wavelength (Wyant et al., 1984a; Cheng and Wyant, 1984; Fercher

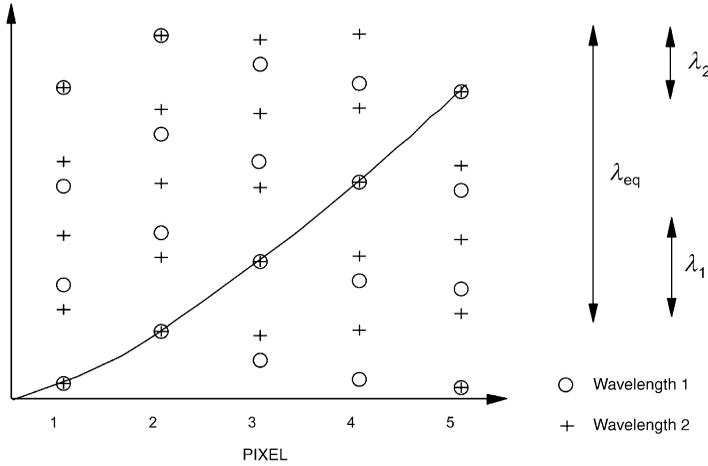


FIGURE 14.48. Two wavelength PSI.

et al. 1985; Creath et al., 1985; Creath and Wyant, 1986; Wyant and Creath, 1989). Two complete sets of interferometric data are collected, and the phase modulo 2π at each wavelength is computed. Since a phase of 2π is a different OPD at each wavelength, only one choice of wavefront deformation will satisfy both data sets. TWPSI can also be interpreted by the connect-the-dot representation (Fig. 14.48). The vertical scale is now OPD instead of phase so that it is not wavelength dependent. Each set of data produces its own column of possible solutions at each pixel, and the spacing of the dots is the wavelength. The locations where the two sets of dots fall on top of each other indicate possible correct solutions for the wavefront; many individual dots are now eliminated.

At each pixel, there are multiple locations of coincidence, and the spacing between these points is the equivalent wavelength:

$$\lambda_{eq} = \frac{\lambda_1 \lambda_2}{|\lambda_1 - \lambda_2|} \tag{14.166}$$

where λ_1 and λ_2 are the two measurement wavelengths. Phase unwrapping of the common points is now done using standard PSI techniques at the equivalent wavelength. Since the equivalent wavelength is much longer than the actual wavelengths, large wavefront slopes can be handled before the algorithm breaks down. Two wavelength operation of an interferometer for testing aspheres requires optical correction of the optical system at both wavelengths as well as raytracing at these wavelengths for the asphere under test. It is also possible to extend this technique to more than two wavelengths (Cheng and Wyant, 1985a). More will be said about these multiple wavelength techniques in Chapter 17.

There is an important similarity between SNI and TWPSI for measuring aspheric surfaces: Both depend on being able to measure aliased fringes, and therefore require sparse array sensors with small pixel width to pitch ratios. They are both limited by

the pixel MTF. A practical problem that has hindered the implementation of TWPSI for this application is chromatic aberration in the reference optics and the interferometer. These optics must be achromatized at both wavelengths.

The application for measuring steps or discontinuities greater than a quarter wave in surface height (π in wavefront phase) requires different solutions than those needed for aspheres. Instead of high-frequency fringes, there is a break in the fringe pattern. Single wavelength techniques lose track of the fringe order over the step and do not produce the correct result. Multiple wavelength techniques, on the contrary, allow the fringe order to be determined and can measure step heights well in excess of a quarter wave. TWPSI has been successfully applied to this problem using interference microscopes (Creath, 1987a), and this application will also be discussed in Chapter 17. SNI has been used for discontinuous surfaces measured at a single wavelength, but the requirement is added that the step height be known to $\pm\lambda/4$ (Greivenkamp et al., 1989). This additional information allows the correct number of 2π 's to be added to the arctangent data. Since neither of these techniques for measuring steps relies on detecting aliased fringes, there are no special sensor considerations.

14.13.4. Subaperture Stitching

A number of other approaches have been applied to the problem of aspheric testing (Wyant, 1987). These tests aim to reduce the number of fringes that are analyzed in a single interferogram. Two of these methods are shearing interferometry and subaperture testing.

The attractive feature of using shearing interferometry for aspheric testing is that the sensitivity of the test can be varied by changing the shear distance (see Chapter 4). In this way, the number of fringes is reduced. However, to measure an asymmetric wavefront, two sets of PSI data with orthogonal shears must be collected. Configurations based upon lateral shear, radial shear and the Ronchi test have been demonstrated (Hariharan et al., 1984; Seligson et al., 1984; Yatagai and Kanou, 1984; Yatagai, 1984; Kothiyal and Delisle, 1985b; Kanoh, 1986; Omura and Yatagai, 1984). The wavefront slope is measured with shearing interferometry, and the actual wavefront or surface then is found by integration. One practical difficulty is that noise in the measurement tends to propagate across the reconstructed wavefront as a result of this integration; a bad pixel can influence large portions of the wavefront map. When the wavefront is measured directly, the effects of a bad pixel are localized at that pixel.

The basic idea behind subaperture testing of aspherics is to divide the wavefront up into small sections; the wavefront departure in each subaperture is within the measurement range of the instrumentation. The maximum fringe frequency is kept below the Nyquist frequency of the sensor. The problem is then reduced to fitting all of these separate measurements, which can contain different amounts of tilt, piston and sometimes defocus, back into a complete map of the aspheric wavefront or surface. (Kim, 1982; Thunen and Kwon, 1982; Jensen et al., 1984; Negro, 1984; Stuhlinger, 1986; Liu et al., 1988) An example of a set of subapertures is shown in Figure 14.49. The overall wavefront is represented by a polynomial expansion, and

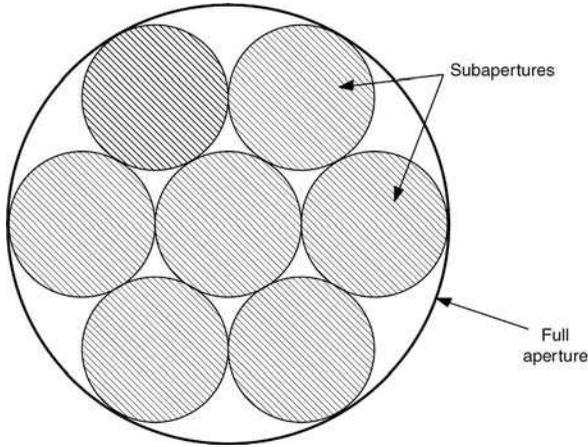


FIGURE 14.49. One configuration for subaperture testing.

the subaperture data is analyzed to determine the expansion coefficients. The Zernike polynomials are the usual choice, and a limited number of terms are used. This technique provides the overall aspheric wavefront shape. Small or localized errors will not appear in the final polynomial fit, and must be determined from the subaperture data. It is important that the location of each subaperture within aperture be precisely known.

Another subaperture technique has been developed to expand the field of view of interferometric surface profiling microscopes (Cochran and Wyant, 1986; Cochran and Creath, 1987, 1988). A series of partial overlapping linear traces of a surface are collected. These measurements are then combined into a single trace by adjusting the tilt, piston and position to minimize the difference between the traces in the overlap region. Stitching techniques have been implemented into professional aspheric testing solution that allow to measure a variety of low and high departure aspheres by selecting different geometries and grids for stitching (Tricard et al. 2006).

14.14. OTHER ANALYSIS METHODS

The bulk of this chapter has concentrated on the phase-step and integrating-bucket algorithms since they are the most commonly used PSI techniques. A number of other PSI algorithms and systems have been developed. In this section, we briefly review some of these special methods. Some are included for historical significance and others are useful in particular situations.

14.14.1. Zero Crossing Analysis

Zero crossing analysis is an electronic technique that directly measures the time delay between the intensity signals at different locations in the interferogram. The

relative wavefront phase at the two points can then be calculated (Crane, 1969; Moore, 1973; Wyant and Shagam, 1978). Referring back to Figure 14.1, the intensity at any location in the interferogram varies sinusoidally, and the temporal phase at a point is proportional to the wavefront phase. One point in the interferogram is designated as a reference point, and the time delay between the signal at this point and another point is measured. These values can be converted to the wavefront phase modulo 2π by normalizing with the period of the reference signal. The location of the zero crossings is a convenient measure of the time lag. Since the intensity never goes negative, a reference intensity level (such as the average intensity $I'(x, y)$) is established for determining the zero crossing locations. To measure a two-dimensional wavefront, this time delay relative to the reference point must be measured at a large number of points.

14.14.2. Synchronous Detection

Synchronous detection is an early technique that is the clear precursor to the phase stepping techniques in use today. It is based upon the well known methods of radio communication theory, where the phase of an unknown signal is recovered by correlation of this signal with a sinusoidal signal of the same frequency (Bruning et al., 1974; Bruning, 1978). For a linear reference phase shift, the intensity pattern in the interferogram is (from Eq. (14.6)):

$$I(x, y, t) = I'(x, y) + I''(x, y) \cos[\phi(x, y) + \alpha t] \quad (14.167)$$

where α is the rate of change of the reference phase. Since we are controlling the reference phase, α is well known, and the sinusoidal and cosinusoidal signals with this same frequency can be generated. This cosinusoidal signal is correlated with the interferogram intensity at each point:

$$\langle I_1(x, y) \rangle = \lim_{T \rightarrow \infty} \frac{1}{T} \int_0^T I(x, y, T) \cos(\alpha t) dt \quad (14.168)$$

which reduces to

$$\langle I_1(x, y) \rangle = I''(x, y) \cos[\phi(x, y)] \quad (14.169)$$

In a similar manner, the correlation of the intensity with a sinusoidal signal produces

$$\langle I_2(x, y) \rangle = -I''(x, y) \sin[\phi(x, y)] \quad (14.170)$$

As with all of the phase step techniques we have examined, the wavefront phase modulo 2π can be determined through an arctangent of the ratio of these two results:

$$\phi(x, y) = \tan^{-1} \left[\frac{-\langle I_2(x, y) \rangle}{\langle I_1(x, y) \rangle} \right] \quad (14.171)$$

This analysis also provides to measure the noise immunity of this technique and for PSI. Additive noise included in the interferogram intensity is not related to the frequency α corresponding to the phase shift, and the correlation operation averages these noise terms to zero.

In a similar manner, this correlation can be performed with temporally sampled data. When N measurements of the interferogram intensity are made during a 2π change in the reference phase,

$$\delta_i = i2\pi/N; \quad i = 1, \dots, N \quad (14.172)$$

Elementary Fourier series analysis can be used to solve for the unknown wavefront phase:

$$\phi(x, y) = \tan^{-1} \left[\frac{-\sum I_i \sin(\delta_i)}{\sum I_i \cos(\delta_i)} \right] \quad (14.173)$$

where the summations run from 1 to N . This result is identical to the result obtained using a least-squares fit of the intensity data to a sinusoidal function of the proper frequency (Eq. 14.66).

14.14.3. Heterodyne Interferometry

Heterodyne or AC interferometry is the general term used to describe interferometers that produce a temporal phase shift by introducing two different optical frequencies into the two arms of the interferometer. Several methods exist for generating this frequency difference, the most common being moving gratings and two-frequency lasers (see Section 14.4). With any of these approaches, the interferogram intensity is modulated at difference in frequency (Eq. (14.9)–(14.12)). The primary applications for heterodyne interferometry are distance measuring interferometry, measuring system dynamics, and surface profilometry (Crane, 1969; Koliopoulos 1980; Som-margren, 1981; Barnes, 1987).

Digital heterodyne interferometry (DHI) has been developed to address the need to perform high-speed analysis of dynamic systems (Massie and Nelson, 1978; Massie, 1980, 1987; Massie et al., 1979; Mottier, 1979; Massie et al. 1981; Massie et al. 1983; Evans, 1983). Two significant applications of this technology are deformable mirrors and turbulent flow field testing. To obtain meaningful PSI data, the phase shifts must occur at a rate that is faster than the changes that are being observed. Acousto-optic Bragg cells are the most convenient method to obtain the frequency shifts, and a frequency difference of up to 1 MHz has been demonstrated. The interferogram intensity will oscillate at this frequency, and the wavefront phase is encoded in the phase difference of these oscillations at different measuring points. Because of the frequencies involved, a high-speed detector is needed; a standard video camera will not work. The conceptionally simplest method to collect the data is to use an array of discrete detectors. Each will produce an output at 1 MHz with a different phase. One detector is defined as the reference detector, and the phase of the

other signals are measured relative to this signal. A phase detection method, such as the zero crossing method described above, measures these phases, and a wavefront map modulo 2π is produced. A high frame rate image tube, called an image dissector tube, can also be used as the detector. This camera functions as a rapidly scannable photomultiplier tube. A repeatability of better than in excess of $\lambda/100$ RMS has been demonstrated with DHI systems.

14.14.4. Phase Lock Interferometry

Phase lock interferometry involves applying a small sinusoidal oscillation to the reference mirror (Johnson and Moore, 1977; Johnson et al., 1977, 1979; Moore et al., 1978; Matthews, et al., 1986). The average reference phase is also allowed to vary so that

$$\delta(t) = \delta + a \sin(\omega t) \quad (14.174)$$

where a is the amplitude of the oscillation and is much less than a wavelength. The resulting interferogram intensity is

$$I(x, y, t) = I'(x, y) + I''(x, y) \cos[\phi(x, y) + \delta + a \sin(\omega t)] \quad (14.175)$$

A series expansion of this intensity produces

$$\begin{aligned} I(x, y, t) = I'(x, y) + I''(x, y) \{ & \cos[\phi(x, y) + \delta] [J_0(a) + 2J_2(a) \\ & \cdot \cos(2\omega t) + \dots] - \sin[\phi(x, y) + \delta] \\ & \cdot [2J_1(a) \sin(\omega t) + 2J_3(a) \sin(3\omega t) + \dots] \} \end{aligned} \quad (14.176)$$

where J_n is a Bessel function of order n . In addition to the fundamental frequency of the oscillation, higher order terms are also present. This signal is then filtered to allow only the fundamental frequency ω :

$$I_f(x, y, t) = -2I''(x, y)J_1(a) \sin[\phi(x, y) + \delta] \sin(\omega t) \quad (14.177)$$

There are a couple of interesting features to this signal. First, if the average reference phase δ is varied so that $\phi(x, y) + \delta$ equals a multiple of π , $I_f(x, y, t)$ goes to zero, and there is no intensity modulation due to the reference mirror oscillation or dither. Referring back to Figure 14.1, this situation corresponds to sitting on either a maximum or a minimum of this curve; the small change in phase due to the dither does not result in any change of intensity. The second thing to notice about $I_f(x, y, t)$ is that its magnitude is proportional to the slope of the interference pattern, and for small values of $\phi(x, y) + \delta$, it is actually proportional to the phase distance from a fringe maximum. The sign of the slope is encoded in the phase of this signal relative to the drive signal; this allows the instrument to find the maximum intensity instead of the minimum.

The interferometer operates by measuring the filtered intensity signal $I_f(x, y, t)$ at a location in the interferogram with a detector. The magnitude and phase of this signal are used to produce a feedback signal that tells the phase shift controller how to change the average phase shift δ to null out this signal. When a null is achieved, the sum $\phi(x, y) + \delta$ is equal to zero modulo 2π ; we are sitting on one of the fringe maxima. Since we know δ , the value of wavefront phase $\phi(x, y)$ modulo 2π has been determined. As the wavefront phase changes, the feedback signal is used to vary the average phase shift to maintain the null condition. When the required average phase shift exceeds $\pm 2\pi$, an up/down counter is incremented or decremented, and δ is reset to zero. This reduces the travel requirements on PZT producing the phase shift, and counts the number of fringes. The wavefront phase is then given by the number of 2π 's that have been counted plus the current value of δ . To measure an area, the detector is scanned across the wavefront. Since the scan rate can be varied, interferograms with large numbers of fringes can be measured with phase lock interferometry, and repeatabilities on the order of $\lambda/100$ have been demonstrated.

14.14.5. Spatial Synchronous and Fourier Methods

While strictly not speaking of PSI techniques, the types of interferogram analysis that are known as spatial synchronous detection and Fourier analysis have some similarity to PSI. Both are fringe analysis methods that look at a single interferogram, and are therefore can be used when turbulence and vibration are present. Both of these techniques are more precise than standard interferogram analysis as it is not necessary to determine fringe centers; the data is also collected on a regular grid providing compatibility with computer analysis.

Spatial synchronous detection operates in the spatial domain by multiplying the measured interferogram with a reference pattern (Ichioka and Inuiya, 1972; Womak, 1984a; Toyooka and Tominaga, 1984). Expressions for the unknown interferogram and the reference pattern are

$$I(x, y) = I'(x, y) + I''(x, y) \cos[\phi(x, y) + 2\pi f_0 x] \quad (14.178)$$

and

$$R(x, y) = \cos(2\pi f_0 x) \quad (14.179)$$

where a significant amount of wavefront tilt has been added to the interferogram, and f_0 is the fringe frequency due to the tilt. $R(x, y)$ is a uniform fringe pattern (tilt fringes) that closely matches the interferogram and is implemented either physically or computationally. The product of these two pattern is

$$I(x, y)R(x, y) = I'(x, y) \cos(2\pi f_0 x) + [I''(x, y)/2] \cos[\phi(x, y) + 4\pi f_0 x] + [I''(x, y)/2] \cos[\phi(x, y)] \quad (14.180)$$

The third term contains low spatial frequencies that can be filtered from the other two terms to produce:

$$S_1(x, y) = [I''(x, y)/2] \cos[\phi(x, y)] \quad (14.181)$$

Similarly, a second filtered signal results from the multiplication of the interferogram with a reference pattern equal to $\sin(2\pi f_0 x)$:

$$S_2(x, y) = [I''(x, y)/2] \sin[\phi(x, y)] \quad (14.182)$$

The wavefront phase difference modulo 2π can then be computed from the ratio of these two signals:

$$\phi(x, y) = \tan^{-1} \left[\frac{S_2(x, y)}{S_1(x, y)} \right] \quad (14.183)$$

The similarity between this algorithm and the many PSI algorithms, especially the synchronous detection algorithm in Eq. (14.171), should be apparent. Similar systems have been built that image the interferogram onto a segmented detector, and the individual outputs are sinusoidally weighted to determine the phase (DeCou, 1974; Mertz 1983a,b, 1989).

The Fourier analysis methods provides a virtually identical analysis, but the computations are done in the spatial frequency domain instead of the spatial domain (Takeda et al., 1982; Macy, 1983; Nugent, 1895; Bone et al., 1986; Kreis, 1986; Roddier and Roddier, 1987). An interferogram pattern with tilt as described by Eq. (14.178) can be rewritten as

$$I(x, y) = I'(x, y) + c(x, y)e^{i2\pi f_0 x} + c^*(x, y)e^{-i2\pi f_0 x} \quad (14.184)$$

where

$$c(x, y) = [I''(x, y)/2]e^{i\phi(x, y)} \quad (14.185)$$

and * indicates the complex conjugate. A one-dimensional Fourier transform of $I(x, y)$ produces

$$\tilde{I}(\xi, y) = \tilde{I}'(\xi, y) + \tilde{c}(\xi - f_0, y) + \tilde{c}^*(\xi + f_0, y) \quad (14.186)$$

where ξ is the spatial frequency coordinate and \sim indicates a Fourier Transform. This function is a trimodal function with peaks at $-f_0$, f_0 , and the origin as sketched in Figure 14.50(a). The component of this spectrum centered at f_0 can be recovered without the carrier by first bandpass filtering and then shifting the isolated spectrum back to the origin (Figure 14.50(b)). This results in the function $\tilde{c}(\xi, y)$. An inverse

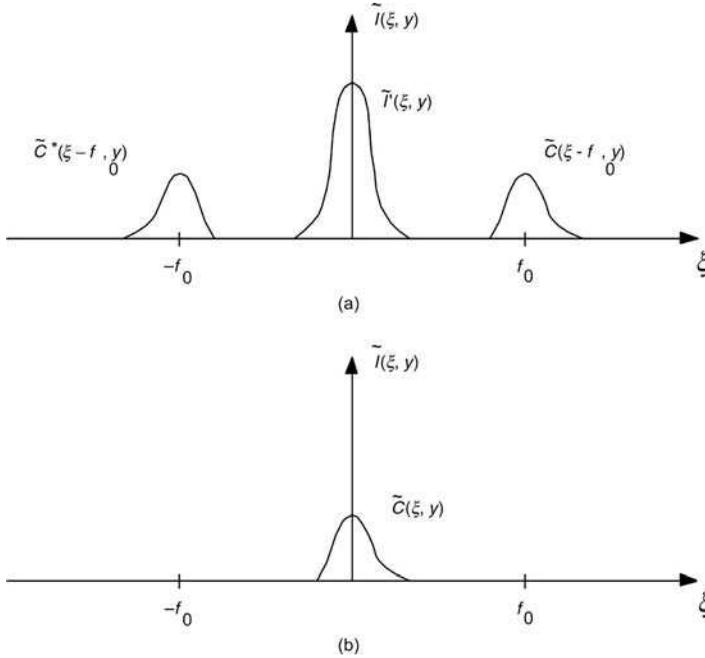


FIGURE 14.50. Fourier fringe analysis: (a) the spectrum of the interferogram and (b) the processed spectrum.

Fourier transform produces $c(x, y)$ as described in Eq. (14.185). The wavefront phase modulo 2π can then be determined from an arctangent:

$$\phi(x, y) = \tan^{-1} \left\{ \frac{\text{Im}[c(x, y)]}{\text{Re}[c(x, y)]} \right\} \quad (14.187)$$

where Re and Im refer to the real and imaginary part of the function.

Both of these spatial techniques suffer from a problem when the wavefront under test has a large deviation from the reference wavefront.

The three components of the interferogram spectrum (Eqs. (14.180) or (14.186)) overlap; and it is not possible to completely separate them through filtering. Various filters have been suggested to optimize the output of these techniques (Womak, 1984a,b).

14.15. COMPUTER PROCESSING AND OUTPUT

Now that a complete representation of the wavefront phase is stored in the computer memory, we can use the computer for the analysis and display of this data. A number of different analyses can be performed to produce results that are appropriate for the application. The initial analysis often starts by fitting the measured wavefront to a polynomial expansion, usually the Zernike polynomials (Kim and Shannon, 1987)

(See Chapter 13). The terms in the expansion corresponding to tilt and focus can then be subtracted from the measured wavefront to leave only the wavefront error as measured from a best-fit reference sphere (for a flat, only tilt is subtracted as we are interested in the radius of curvature of the wavefront). Higher order aberrations, such as spherical, coma, and astigmatism, can also be subtracted.

The PSF and the MTF due to the measured wavefront errors and diffraction can be easily computed from the wavefront data. We start by expressing the wavefront in the form of Eq. (14.3) which includes an amplitude term:

$$w(x, y) = a(x/d, y/d)e^{i\phi(x, y)} \quad (14.188)$$

The amplitude term $a(x, y)$ now also defines the exit pupil diameter d . The amplitude PSF is found through a fast Fourier transform (FFT) of this wavefront function, and the intensity PSF results from the squared modulus of this result (Goodman, 1968). Since the wavefront represents only the error from a spherical reference surface, this answer must be scaled to the image plane with the wavelength and the system $f/\#$. The pupil diameter is already contained in the expression for the wavefront so that the scaling is done by replacing the spatial frequency variables ξ and η with $x/\lambda f$ and $y/\lambda f$, respectively:

$$\text{PSF} = |\tilde{w}(\xi, \eta)|^2 \Big|_{\xi=x/\lambda f; \eta=y/\lambda f} \quad (14.189)$$

where \sim indicates a Fourier transform, and f is the focal distance. The result of this computation for an unaberrated wavefront with a circular pupil is an Airy pattern, and aberrations will distort the PSF. The system MTF contribution of the tested wavefront or surface is found through an FFT of the PSF. This result is properly scaled to the $f/\#$ of the optical system, and the diffraction-limited cutoff frequency of $1/(\lambda f/\#)$ results.

There are times when the system parameters (pupil diameter, focal length and wavelength) are not known or defined at the time of the test, and it is useful to express the results in the arbitrary units. An example might be the test of a flat mirror to be used in a multi-element system; we are interested in the influences of the flat mirror on the system performance, and assume that the aberrations of that surface propagate through the system to the final converging wavefront. We may also not be testing the element at the final system wavelength. It is the parameters of this final wavefront that are used to scale the PSF and MTF due to the element. An arbitrary unit that is sometimes used is waves/radius, which can be obtained by setting the exit pupil radius, the focal distance, and the wavelength all equal to one in the above calculations. The output dimensions in waves/radius can then be scaled at a later time to a physical distance by multiplying by $\lambda f/r = 2\lambda f/\#$, where r is the pupil radius.

Examples of interferometer output for a tested surface are shown in Figure 14.51, where the OPD and the resulting PSF and MTF are displayed. A number of different plotting options, including contour maps and isometric plots, are available. Some other examples of analyses that can be performed on the wavefront data are one-dimensional traces, surface slope plots, Strehl ratio computations, encircled energy distributions, surface height histograms and geometrical spot diagrams. The choice

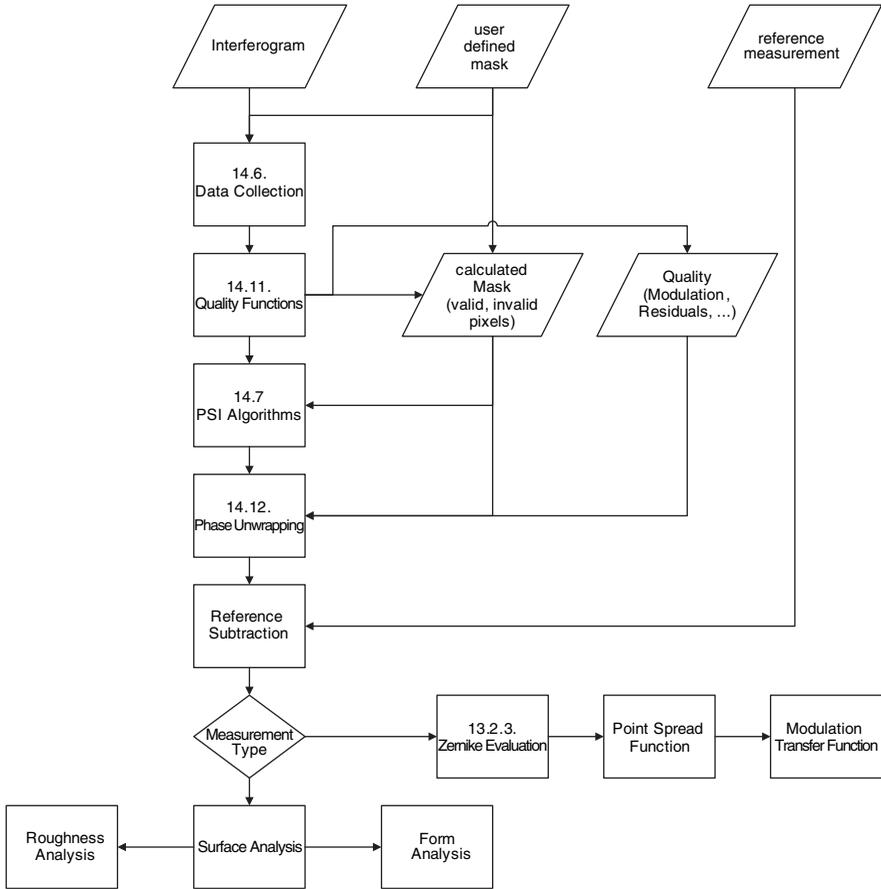


FIGURE 14.51. Examples of the data-flow inside a PSI system.

of the output can be customized to the requirements of the user (Truax, 1986; Creath, 1987b). Other features that can be built into the software are production menus that lead the operator through the test, and then provide a pass/fail display based upon certain predetermined criteria.

Recently, interfaces between interferometer analysis software and optical design programs have become available. While the exploration of this capability is just beginning, it has been used for several applications (Figoski et al., 1989; Stephenson, 1989; Willey and Patchin, 1989). The overall system design can be reoptimized based on the actual parts that have been manufactured and tested; this is especially useful for systems that are built around a difficult to manufacture element. In a similar manner, interferometric data can be used to predict the actual system performance for comparison to the design requirements. The alignment of optical systems has also made use of this capability. The wavefront from either the entire optical system or a subsystem is measured and compared to the wavefront predicted

by the design program. Based on the differences between these two wavefronts, the design program computes in an iterative fashion the changes needed to bring the system into specification. A fourth, and more general, application of this capability is to create reference aspheric wavefronts in the computer software. The design software calculates the desired wavefront, and the interferometer then subtracts this calculated wavefront from the measurement to determine the wavefront errors. This process is almost exactly analogous to the use of computer generated holograms, except that the “hologram” always resides in computer memory.

14.16. IMPLEMENTATION AND APPLICATIONS

In this section, we discuss a number of practical considerations for PSI and also examine some particular applications of PSI. A generic PSI system consists of four things: an interferometer, a device for changing the reference phase, a detector or detector array, and a computer. The detector array provides the interface between the wavefront and the computer, and the computer provides control of the interferometer and phase shifter and performs the necessary computations. The advances in detectors and computers have been the enabling technologies that have driven the dramatic performance improvements and acceptance of PSI. There is sufficient computing power in personal computers to handle most PSI applications.

14.16.1. Commercial Instrumentation

Commercial instrumentation utilizing PSI technology is available for measuring surface and wavefront shape, surface profile, and distance. PSI-based surface profiling microscopes are widely used for quantitative examination of small surface features and microroughness of both optical and non-optical components. Repeatabilities of a few Angstroms RMS can be obtained, and these profiling instruments are fully discussed in Chapter 17.

PSI interferometers have been marketed for measuring surfaces and wavefronts based on all three of the optical configurations shown in Figure 14.2; Twyman–Green, Mach–Zehnder and laser Fizeau (Chapter 1 and 2). All of these systems employ video cameras to record the interferograms; the phase shifts are provided by moving the reference objective with PZT transducers, and a variety of reference objectives and flats are available to match the part under test (Smythe et al., 1987; Creath, 1987b). A prominent feature of these systems is the flexibility of the user interface provided by the computer. As discussed in Section 14.15, a large number of graphics displays can be produced, and the processing can be customized to meet the needs of the user.

A Mach–Zehnder interferometer configuration that is available for testing the quality of optical wavefronts with PSI is shown in Figure 14.52 (Hayes and Lange, 1983; Creath, 1987b). The output of a source, such as a laser diode, is collimated as it enters the system. A spherical reference wavefront is provided by the spatial filter in the reference arm. A similar set of optics, without the pinhole, is used in the test arm to match the optical path length in the two arms. The beams are recombined and

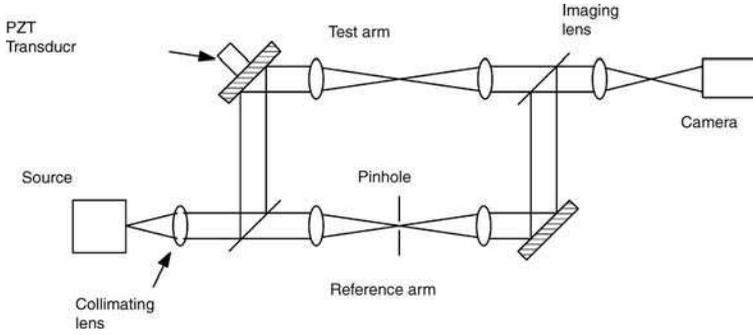


FIGURE 14.52. Mach-Zehnder PSI interferometer for measuring the wavefront quality of sources.

imaged onto the camera. Both the intensity and phase distributions of the wavefront can be measured.

An interesting system that can be used for testing the shape or flatness of nonoptical parts has been developed based on grazing-incidence interferometry (Synborski, 1978). The system is diagrammed in Figure 14.53(a). The component under test is placed in close proximity to the hypotenuse of a large glass prism which serves as the reference surface for the interferometer. An air gap is maintained.

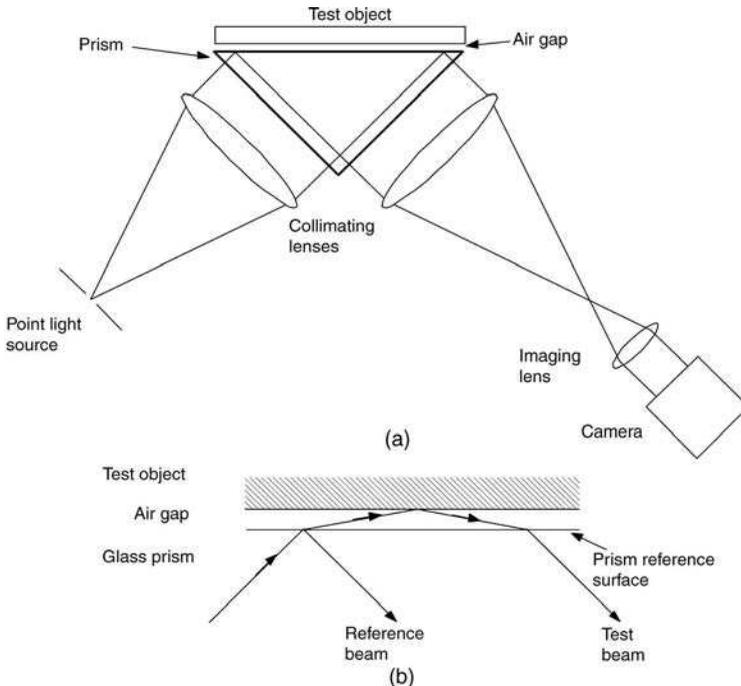


FIGURE 14.53. Grazing incidence PSI interferometer: (a) overall system layout and (b) a closeup of the interface at the air gap.

A close-up of this interface is shown in Figure 14.53(b). Because of refraction, the test beam is incident on the part at a large angle. For a 41.7° angle of incidence and an index of 1.5, the angle of incidence at the test surface is 86° . This is close to the critical angle of the prism. There are two beneficial consequences to this configuration. First, the sensitivity of the test is reduced by the cosine of the angle of incidence. At 86° , the sensitivity is 7% of the normal incidence sensitivity; a fringe represents 7 waves of surface height change instead of a half wave. It is as if the surface is tested with a longer wavelength. The second benefit is that almost everything, including a ground surface, is highly reflective at this angle of incidence, and a variety of different materials and samples can be used. If the angle of the light in the prism is slightly changed, the optical path length in the air gap will change much more than the path in the glass. This tilt when used produces the phase shift between the reference and test beams. Systems are available with up to a 200 mm aperture, and have been applied to testing silicon wafers and machined or ground parts. The clear drawback to a system such as this is the need for the large precision glass prism.

Distance measuring interferometers (DMIs) are capable of providing position readout with a resolution of better than a thousandth of a wavelength (less than a nanometer), and have found extensive use in the area precision machine control (Dukes and Gordon, 1970; Steinmetz et al., 1987; Smythe et al., 1987; Steinmetz, 1990). Most DMIs use heterodyne techniques (Section 14.14.3), and a basic system is diagrammed in Figure 14.54. The two output frequencies from the source are orthogonally polarized, and the polarizing beamsplitter splits the frequencies into the reference and test arms of the interferometer. The two corner cubes return these two beams to the detector, where the polarizer at 45° allows them to interfere. When the movable corner cube is stationary, a beat frequency equal to the difference in the two optical frequencies is obtained; this frequency is identical to a similar reference signal generated in the laser head. A motion of this corner cube introduces a Doppler frequency shift in the light returning from the test arm, and the measured beat frequency changes. The change in the length of the test arm is determined by comparison of the measured beat frequency with the reference beat frequency. Some specific applications for DMIs are microlithographic step and repeat cameras, single-point diamond turning machines and precision photoplotters. Additional information about DMIs is contained in Chapter 18.

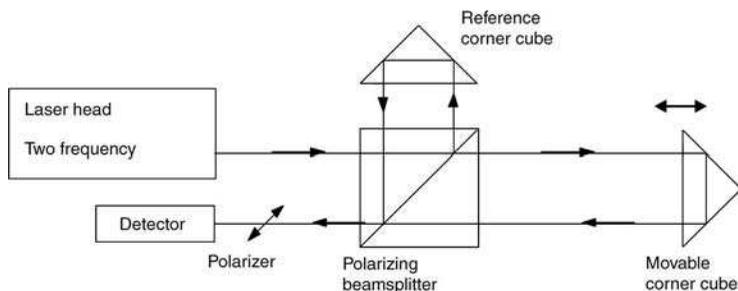


FIGURE 14.54. Distance measuring interferometer.

14.16.2. Interferometer Configurations

A number of grating lateral shearing interferometers that use PSI techniques have been demonstrated (Wyant, 1975; Hardy et al., 1977; Stump, 1979; Koliopoulos, 1980; Hardy and MacGovern, 1987; Schreiber and Schwider, 1997). Translation of the grating produces a frequency shift between the diffracted orders, resulting in a linear phase shift between the interfering beams. More information on grating interferometers can be found in Chapter 4. The wavefront is focused onto the grating, and three or more diffraction orders are produced. The period of the grating is chosen so that there is overlap, and therefore interference, of the diffracted orders. When the grating is translated, the diffracted orders are Doppler shifted (see Section 14.4); the plus one order increases, the minus one order decreases, and the zero order is unchanged. A beat frequency appears in the output that can be analyzed by either heterodyne or integrating bucket techniques to determine the wavefront phase. A convenient grating configuration is to use a radial grating (starburst pattern) that is spun about its center (Stevenson, 1970). With this grating, the shear distance can be changed by moving the grating center relative to the focus spot.

PSI has also been applied to the double-frequency grating lateral shear interferometer (Wyant, 1973). In this system, the wavefront is focused onto a grating with two distinct line spacings. Two diffraction patterns are produced, one from each set of grating lines. The shear between the two interfering first orders is determined by the difference in the two grating frequencies, and the average frequency can be chosen to separate these diffracted orders from the undiffracted light. Translation of the grating through the focus will produce different Doppler shifts for each of the two diffracted beams, and a relative phase shift is produced. The resulting interferogram can be evaluated by PSI techniques.

A very compacted shearing interferometer configuration (Schreiber and Schwider, 1997) uses two Ronchi phase gratings in series. The phase delay of both gratings is chosen so that the zero orders are minimized. Combining the first orders from the first grating with the opposite first orders of the second grating results in two on axis beams sheared to each other. The amount of shear can be adjusted by the grating distance. Besides it being on axis and easy to adjust nature, the combination of positive and negative orders in this configuration results in on-axis beams independent of the wavelength making it perfect for use with light sources using a broader bandwidth like LED diodes or Excimer laser (Schreiber et al., 2002).

In situations where the test wavefront can change with time, there is a desire to obtain all of the required phase-shifted interferograms simultaneously. There should be no time lag between the measurements. Several of these snapshot interferometers have been demonstrated, and they are also compatible with pulsed light sources.

A second stroboscopic interferometer uses a diffraction grating to introduce a phase shift between the reference and multiple test beams (Kwon, 1984, 1987; Kwon et al., 1987). One implementation of the system is shown in Figure 14.55. The three replicas of the test beam are the diffracted orders from the stationary grating, and the reference beam is provided by the spatial filter. A small lateral shift of the grating will produce a 90° phase shift between the three test beams. A simple way to understand

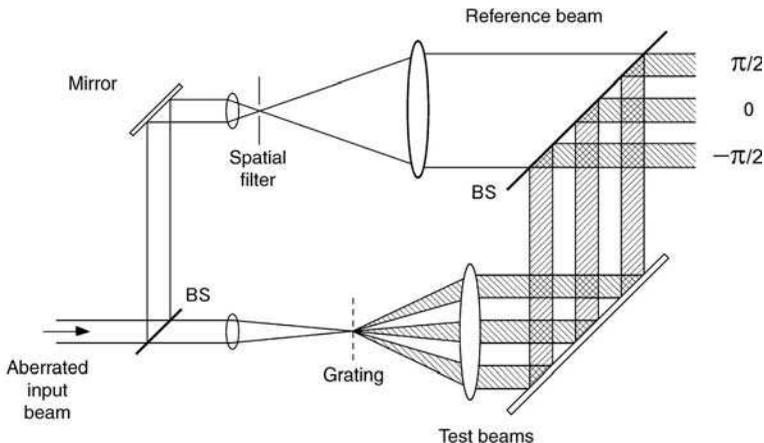


FIGURE 14.55. Stroboscopic PSI interferometer using a diffraction grating (from Kwon, 1987).

this phase shift is that the grating is in the Fourier transform plane of the wavefront; a shift in this plane will result in a linear phase factor in the plane of the interferograms. A shift of a quarter of the grating period will produce the desired phase shift. The three output interferograms can be evaluated with one of the three step algorithms (Eqs. (14.48) or (14.52)).

A third snapshot interferometer (McLaughlin and Horwitz, 1986) uses a phase grating to shear the wavefront under test and to create a carrier fringe frequency in the interferogram. The period of these carrier fringes is adjusted to equal four pixels on the sensor, and each pixel sees a relative phase shift of 90° . The four step algorithm can therefore be used to analyze the sheared wavefront. There is a great deal of similarity between this method and the spatial synchronous detection methods discussed in Section 14.14.5. A crossed grating is used to simultaneously analyze the shear in orthogonal directions.

14.16.3. Absolute Calibration

We generally assume that the accuracy of our interferometric test is only as good as the reference surface since the interferometer compares the test surface with this reference surface. The absolute accuracy of the measurement can be improved if we are able to subtract errors in the results introduced by the reference surface and by the interferometer. A perfect test optic can be used to calibrate the interferometer for a particular configuration, but very often these reference parts do not exist. A procedure has been developed to obtain an absolute calibration of a spherical surface and the interferometer without knowing either in advance (Jensen, 1973; Bruning et al., 1974; Bruning, 1978; Truax, 1988) (See also Sec. 2.5). Since this calibration process involves computations with several wavefronts, it is well suited for PSI systems as the wavefronts are already in computer memory. The three sets of data that are needed for this procedure are shown in Figure 14.56. The first two data sets

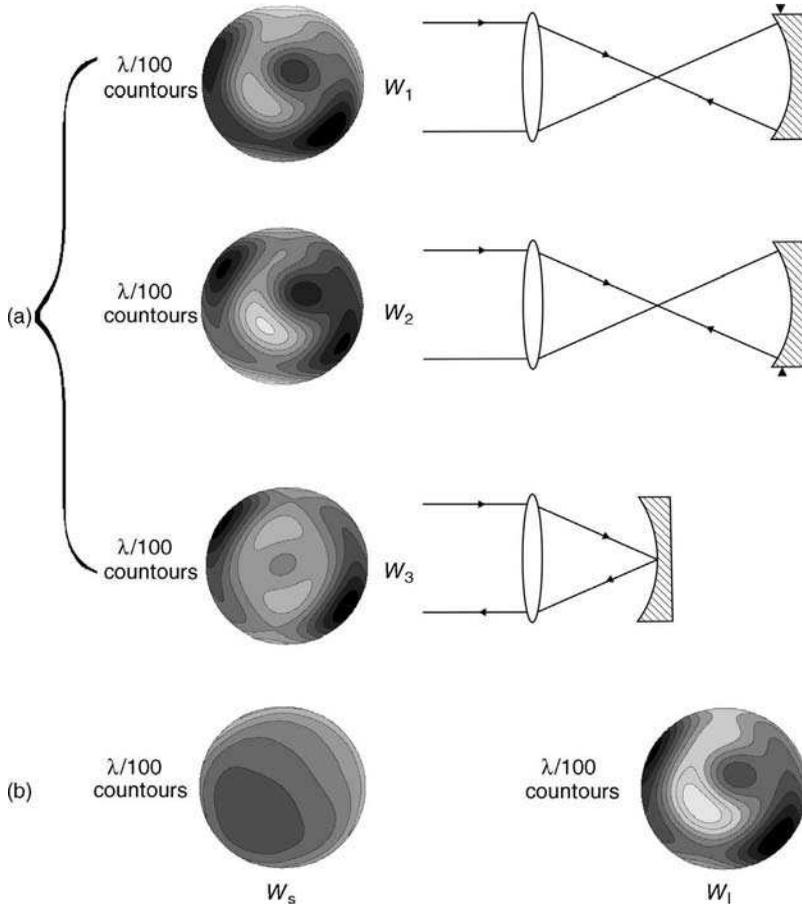


FIGURE 14.56. Absolute calibration of a spherical surface: (a) three raw wavefronts of part with interferometer errors and (b) composite errors of the interferometer and test surface errors.

are tests of the sphere with the test surface rotated by 180° between the measurements. It is imperative that the sphere be precisely rotated about the optic axis. The third test measures the interferometer in a cat's eye reflection with the interferometer focused on the vertex of the sphere. Wavefront maps are obtained in each case. There are three sources of wavefront error in the system: errors in the test surface w_s , errors in the reference arm and reference surface w_R , and errors in the test arm and the diverger lens w_T . Each of these three interferograms can be expressed as a combination of these wavefront errors:

$$W_1 = W_R + W_T + W_S \quad (14.190)$$

$$W_2 = W_R + W_T + W_S^{180^\circ} \quad (14.191)$$

$$W_3 = W_R + \frac{1}{2}(W_T + W_T^{180^\circ}) \quad (14.192)$$

where

$$\begin{aligned} W_R^\alpha &= \text{reference arm wavefront} \\ W_T^\alpha &= \text{reference arm wavefront} \\ W_S^\alpha &= \text{reference arm wavefront} \\ \alpha &= \text{orientation of wavefront (0 or } 180^\circ) \end{aligned}$$

After a cat's eye reflection, the ray paths are flipped about the optic axis, and this results in the rotation of the test arm aberrations in W_3 . The evaluation first averages W_1 and a rotated version of W_2 . The location of the optic axis in the interferogram must be identified for this data rotation. Data-set W_3 is then averaged with a rotated version of itself, and this result is then subtracted from the earlier result to produce the wavefront error due to the test surface:

$$W_S = \frac{1}{2}(W_1 + W_2^{180^\circ} - W_3 - W_3^{180^\circ}) \quad (14.193)$$

The surface height map can be obtained from this wavefront error. The total error in the interferometer, denoted by W_I , is composed of the sum of W_R and W_T . The interferometer errors can be calculated by substituting (14.193) into (14.190), which gives

$$W_I = \frac{1}{2}(W_1 - W_2^{180^\circ} + W_3 + W_3^{180^\circ}) \quad (14.194)$$

Note that both Eqs. (14.193) and (14.194) involve the manipulation of wavefronts that derive from the rotation of the part and the rotation of the wavefront data matrices by 180° . This allows the separation of even and odd errors in the interferometer when we have access to the cat's eye configuration. Figure 14.56b shows the results of such calibration calculations as well as the individual and composite errors in the interferometer and test part. This method is usually employed when it is essential to achieve the very highest accuracies possible. The method is only effective when the is operated in a stable environment and the 0° and 180° test part configurations of W_1 and W_2 match the pixel coordinates (x', y') and $(-x', -y')$ of the detector matrix. This calibration procedure is conveniently carried out with PSIs that can easily manipulate and store wavefronts in memory. Errors incurred by improper alignment of the various surfaces have been examined by Truax (1988) and Ellsner et al. (1989). Generally, lateral alignments are the most critical and need to be held to within one pixel or less.

The special case of flat surfaces can be calibrated by using three surfaces in different orientations. A similar approach has been applied to Fizeau reference calibration (Schulz and Schwider, 1976). The use of Zernike polynomials facilitates the calibration process for flats (Fritz, 1984).

14.16.4. Sources

PSI systems have been built at wavelengths ranging from the ultraviolet to the infrared. The test wavelength is sometimes chosen to test the system at the operating wavelength (Prettyjohns et al. 1985). An improvement in the precision of the test results from using shorter wavelengths; 2π of wavefront phase corresponds to a smaller physical distance. Longer wavelengths reduce the sensitivity of the test, and permit the testing of aspheric wavefronts without nulls. A second advantage to these long wavelengths is that high contrast fringes can be obtained even when testing ground surfaces, and testing can commence during the grinding process. A common IR wavelength that has been used for PSI is the $10.6 \mu\text{m}$ output of a CO_2 laser, and the detector used for these systems is a pyroelectric vidicon (PEV) (Kwon et al., 1980; Stahl et al., 1987; Creath, 1987b; Stahl, 1989). PEV's respond only to time-varying imagery which makes them appropriate for a PSI system. The phase shift provides the intensity variations, but the rate of phase shift must be adjusted to optimize the camera response. A modulation frequency of 7.5 Hz has been found to be an optimum in one system (Stahl et al., 1987). Work in the development of IR solid state sensors has pushed the cutoff wavelength into the mid-IR, about $6 \mu\text{m}$ (Zanio, 1990; Kozlowski et al., 1990). This work in materials such as HgCdTe may eventually provide an alternative to PEV's.

An area that shows great promise for simplifying the operation of PSI systems is to obtain the phase shifts through modulation of the output wavelength of a laser diode (Tatsuno and Tsunoda, 1987; Ishii, et al., 1987; Chen and Murata, 1988; Chen et al., 1988). The output wavelength of a laser diode varies with both the injection current and temperature. The laser diode output is sent to both arms of the interferometer and the phase shift is proportional to the OPD between the two arms:

$$\alpha = \frac{2\pi \text{OPD } \Delta\lambda}{\lambda_0^2} \quad (14.195)$$

where $\Delta\lambda$ is the wavelength change, and λ_0 is the average wavelength. As with the instrumentation for the 2+1 algorithm (Section 14.8.6), this phase shift corresponds to the difference in the number of wavelengths that fit into the OPD at the two wavelengths. The laser wavelength can either be stepped to implement the phase-step PSI approach or ramped for the integrating-bucket approach. This method of phase shift eliminates the need for a mechanical motion of a PZT and the associated high-voltage drive circuit.

A source modulation interferometer has been used to measure both surfaces and the refractive index inhomogeneities of a glass plate from a set phase-shifted interferograms (Okada, Sakuta, Ose and Tsujiuchi, 1990). The plate is put in the test arm of a Twyman–Green interferometer, and fringes are formed between the reference surface and the front surface reflection, the reference and the rear surface reflection, and between the two surface reflections. Since each of these overlapping fringe patterns corresponds to a different OPD, each phase shifts at a different rate, and the analysis is able to separate the three quantities to be measured.

A phase-lock laser diode interferometer (see Section 14.4) has also been demonstrated by modulating the drive current, and therefore the output wavelength, of the diode (Suzuki et al., 1989). Two potential problems with the use of laser diode sources are changes in the output power as the wavelength is varied and the effects of a small wavelength modulation needed to lock the lock the wavelength to a reference cavity (Hariharan, 1989a,b).

14.16.5. Alignment Fiducials

One area that is often overlooked in interferometric testing is the accuracy required in relating the position of a fringe or a point in an interferogram to an actual position in the aperture of the component or system under test. This relationship is accomplished by placing fiducial marks on the test aperture. The image of these marks in the interferogram relates the interferogram coordinates to locations on the part. Precise fiducial location determination allows for the removal of calibration wavefronts, such as those due to the interferometer or reference optics. These alignment aids are important for the absolute calibration and subaperture techniques described above. Fiducials also provide a coordinate system for geometric wavefront manipulations, including flips, rotations and magnification changes, as well as for aberration fits. The importance of the geometrical alignment of a test part with the analysis can be understood by considering the situation of subtracting a stored calibration wavefront from a measured wavefront. Misalignment will shift the error to be subtracted on itself during the analysis. These shifts will result in comatic and/or astigmatic errors.

14.17. FUTURE TRENDS FOR PSI

This chapter has concentrated on the theory of phase-shifting analysis and its application to the area of interferometry. PSI has provided improved repeatability, precision and accuracy, and allowed for a more complete analysis of the wavefront and surface data. Not surprisingly, these same techniques can be applied to almost any system that uses fringes for measurement. Chapter 15 describes the application of PSI analysis to holographic and speckle interferometers (Hariharan, Oreb and Brown, 1982a,b, 1983; Creath, 1985; Hariharan, 1985; Nakadate and Saito, 1985; Thalmann and Dändliker, 1985), and Chapter 16 discusses its use with moiré and fringe projection techniques (Indebetouw, 1978; Moore and Truax, 1979; Shagam, 1983; Bell et al., 1984). Optical profilers (Chapter 17) and distance measuring interferometers (Chapter 18) are two other areas of instrumentation where phase-shifting techniques have been applied with great success.

There are several areas where further improvements to this technology are needed. A partial list includes systems that work in the deep ultraviolet with excimer lasers, systems with improved compatibility with pulsed lasers, improved calibration procedures through ray-tracing the aberrations of the interferometer, the integration of alignment fiducials into the analysis programs, and the testing of aspheric surfaces and wavefronts.

Even with the available technology, aspherics remain difficult to test. There exists a need for a device that could be described as a walk-up aspheric interferometer. Such a system would allow for the testing of an aspheric surface without the need for the preparation of any special reference or null optics; this is the equivalent to the situation that exists for spherical optics today. It remains to be seen if the extended range techniques, such as sub-Nyquist interferometry, will provide a path for reaching this goal. The development of this instrument will almost certainly be coupled with the integration of optical design programs into the interferometer to provide for on-line ray tracing and calibration.

There are some current trends in PSI systems that are certain to remain. The continued influence of the improvements in solid state electronics on PSI will mean larger detector arrays, detector arrays that are sensitive a greater range of wavelengths, and faster and more complete computation. The computer interface will become more flexible and user friendly. The use of interferometers in conjunction with optical design programs will also become important for optimization, alignment and testing of systems. In manufacturing, the computer analysis of interferograms permits the integration of statistical process control with the manufacturing process to improve quality and yield. Interferometric testing will also be applied to a greater variety of nonoptical parts.

It is clear that over the last 30 years, phase shifting interferometry has become established as an integral part of optical testing. Its success can be measured by its widespread use and the increasing number of applications to which it has been applied.

REFERENCES

- Ai C. and J. C. Wyant, "Effect of Spurious Reflection on Phase Shift Interferometry," *Appl. Opt.*, **27**, 3039 (1988).
- Angel J. R. P. and P. L. Wizinowich, "A Method of Phase Shifting in the Presence of Vibration," *European Southern Observatory Conf. Proc.*, **30**, 561 (1988).
- Baldi A., "Two-Dimensional Phase Unwrapping by Quad-Tree Decomposition," *Appl. Opt.*, **40**, 1187-1194 (2001).
- Barnes T. H., "Heterodyne Fizeau Interferometer for Testing Flat Surfaces," *Appl. Opt.*, **26**, 2804 (1987).
- Bell B. W. and C. L. Koliopoulos, "Moire Topography, Sampling Theory Charged-coupled Devices," *Opt. Lett.*, **9**, 171 (1984).
- Bi H., Y. Zhang, K. V. Link, and C. Wen, "Class of 4+1 Phase Algorithm with Error Compensation," *Appl. Opt.*, **43**, 4199-4207 (2004).
- Bone D. A., H. A. Bachor, and R. J. Sandeman, "Fringe Pattern Analysis Using a 2-D Fourier Transform," *Appl. Opt.*, **25**, 1653 (1986).
- Brophy C. P., "Effect of Intensity Error Correlation on the Computed Phase of Phase-shifting Interferometry," *JOSA A*, **7**, 537 (1990).
- Bruning, J.H. and D.R. Herriott, "A Versatile Laser Interferometer," *Appl. Opt.*, **9**, 2180 (1970).

- Bruning J. H., D. R. Herriott, J. E. Gallagher, D. P. Rosenfeld, A. D. White, and D. J. Brangaccio, "Digital Wavefront Measuring Interferometer for Testing Optical Surfaces, Lenses," *Appl. Opt.*, **13**, 2693 (1974).
- Bruning, J. H., "Fringe Scanning Interferometers," in *Optical Shop Testing*, D. Malacara Ed., 1st Edition, Wiley, New York, 1978.
- Bryngdahl O., "Polarization-Type Interference Fringe Shifter," *JOSA*, **62**, 462 (1972a).
- Bryngdahl O., "Image-Transfer Characteristics of Carrier-Frequency Photography," *JOSA*, **62**(6), 807 (1972b).
- Bryngdahl O., "Heterodyne Shearing Interferometers Using Diffractive Filters with Rotational Symmetry," *Opt. Comm.*, **17**, 43 (1976).
- Burgwald G. M. and W. P. Kruger, "An Instant-on Laser for Distance Measurement," *Hewlett-Packard J.*, **21**, 14 (1970).
- Bushman T. W., M. A. Gennert, and R. J. Pryputniewicz, "Phase Unwrapping by Least Squares Error Minimization of Phase Curvature," *Proc. SPIE*, **2003**, 334 (1993).
- Carré P. "Installation et utilisation du comparateur photoélectrique et Interférentiel du Bureau International des Poids et Mesures," *Metrologia*, **1**, 13-23 (1966).
- Chen J. and K. Murata, "Digital Phase Measuring Fizeau Interferometer for Testing of Flat and Spherical Surfaces," *Optik*, **81**, 28 (1988).
- Chen J., Y. Ishii, and K. Murata, "Heterodyne Interferometry With a Frequency-Modulated Laser Diode," *Appl. Opt.*, **27**, 124 (1988).
- Chen X., M. Gramaglia, and J. A. Yeazell, "Phase-Shifting Interferometry With Uncalibrated Phase Shifts," *Appl. Opt.*, **39**, 585-591 (2000).
- Cheng Y. and J. C. Wyant, "Multiple-Wavelength Phase-Shifting Interferometry," *Appl. Opt.*, **24**, 804 (1985a).
- Cheng Y. and J. C. Wyant, "Phase Shifter Calibration in Phase-Shifting Interferometry," *Appl. Opt.*, **24**, 3049 (1985b).
- Cheng Y. and J. C. Wyant, "Two-Wavelength Phase Shifting Interferometry," *Appl. Opt.*, **23**, 4539 (1984).
- Ching N. H., D. Rosenfeld, and M. Braun, "Two Dimensional Phase Unwrapping Using a Minimum Spanning Tree Algorithm," *IEEE Transaction on Image Processing*, **1**, 355-365 (1992).
- Cochran E. R. and J. C. Wyant, "Longscan Surface Profile Measurements Using a Phase-Modulated Mirau Interferometer," *Proc. SPIE*, **680**, 112 (1986).
- Cochran E. R. and K. Creath, "A Method for Extending the Measurement Range of a Two-Dimensional Surface Profiling Instrument," *Proc. SPIE*, **818**, 353 (1987).
- Cochran E. R. and K. Creath, "Combining Multiple-Subaperture, Two-Wavelength Techniques to Extend the Measurement Limits of an Optical Surface Profiler," *Appl. Opt.*, **27**(10), 1960 (1988).
- Crane R., "Interference Phase Measurement," *Appl. Opt.*, **8**, 538 (1969).
- Creath K., "Phase-Shifting Speckle Interferometry," *Appl. Opt.*, **24**, 3053 (1985).
- Creath K., "A Comparison of Phase-Measurement Algorithms," *Proc. SPIE*, **680**, 19 (1986).
- Creath K., "Step Height Measurement Using Two-Wavelength Phase-Shifting Interferometry," *Appl. Opt.*, **26**, 2810 (1987a).
- Creath K., "WYKO Systems for Optical Metrology," *Proc. SPIE*, **816**, 111 (1987b).

- Creath K., "Phase-Measurement Interferometry Techniques," in: *Progress in Optics Vol. XXVI*, E. Wolf Ed., Elsevier Science Publishers, Amsterdam, 349–393, (1988).
- Creath K. and J. C. Wyant, "Direct Phase Measurement of Aspheric Surface Contours," *Proc. SPIE*, **645**, 101 (1986).
- Creath K., Y.-Y. Cheng, and J. C. Wyant, "Contouring Aspheric Surfaces Using Two-Wavelength Phase-Shifting Interferometry," *Opt. Acta*, **32**, 1455 (1985).
- Crosdale F. and R. Palum, "Wavelength Control of a Diode Laser for Distance Measuring Interferometry," *Proc. SPIE* **1219**, 490 (1990).
- de Groot P. J., "Long-Wavelength Laser Diode Interferometer for Surface Flatness Measurements," *Proc. SPIE*, **2248**, 136–140, (1994).
- de Groot P. J., "Vibration in Phase Shifting Interferometry," *JOSA A*, **12**, 354–365 (1995).
- de Groot P. J., "101-Frame Algorithm for Phase-Shifting Interferometry," *Proc. SPIE*, **3098**, 283–292 (1997).
- de Groot P.J., "Measurement of Transport Plates with Wavelength-Tuned Phase-Shifting Interferometry," *Appl. Opt.*, **39**(16), 2658–2663 (2000).
- de Groot P. J. and L. L. Deck, "Numerical Simulations of Vibration in Phase-Shifting Interferometry," *Appl. Opt.*, **35**, 2172 (1996).
- Deck L. L., "Multiple Surface Phase Shifting Interferometry," *Proc. SPIE*, **4451**, 424–431 (2001).
- DeCou A. B., "Interferometric Star Tracking," *Appl. Opt.*, **13**, 414 (1974).
- Dorsey A., R. Hocken, and M. Horowitz, "A Low Cost Laser Interferometer System for Machine Tool Applications," *Prec. Eng.* **5**, 29 (1983).
- Duarte F. J., *Tunable Laser Optics*, Elsevier Acad. Press, London, 2003.
- Dukes J. N. and G. B. Gordon, "A Two-Hundred Foot Yardstick With Graduations Every Microinch," *Hewlett-Packard J*, **21**, 2 (1970).
- Dunn T. J., "Frequency-Scanning Interferometry Advances Precision Component Manufacturing, Assembly," *Photonics Spectra*, **39**, 96 (2005).
- Elssner K.-E., R. Burow, J. Grzanna, and R. Spolaczyk, "Absolute Sphericity Measurement," *Appl. Opt.*, **28**, 4649 (1989).
- Fercher A. F., H. Z. Hu, and U. Vry, "Rough Surface Interferometry With a Two-Wavelength Heterodyne Speckle Interferometer," *Appl. Opt.*, **24**, 2181 (1985).
- Figoski J. W., T. E. Shrode, and G. F. Moore, "Computer-Aided Alignment of a Wide-Field, Three-mirror, Unobscured, High-Resolution Sensor," *Proc. SPIE* **1049**, 166 (1989).
- Flynn T. J. "Consistent 2-D Phase Unwrapping Guided by a Quality Map," *Proceedings of the 1996 International Geoscience and Remote Sensing Symposium*, Lincoln NE, IEEE, 1996, pp. 2057–2059.
- Freischlad K. R. and C. L. Koliopoulos, "Fourier Description of Digital Phase-Measuring Interferometry," *JOSA A*, **7**, 542 (1990).
- Friedlander B. and J. Francos. "Model Based Phase Unwrapping of 2-d Signals," *EEE Transactions on Signal Processing*, **44**, 2999–3007 (1996).
- Fritz B. S., "Absolute Calibration of an Optical Flat," *Opt. Eng.*, **23**, 379 (1984).
- Fukano T. and I. Yamaguchi, "Separation of Measurement of the Refractive Index, the Geometrical Thickness by Use Of A Wavelength-Scanning Interferometer with A Confocal Microscope," *Appl. Opt.*, **38**, 4065–4073 (1999).

- Gaskill J. D., "Linear Systems," *Fourier Transforms and Optics* Wiley, New York, 1978.
- Ghiglia D. C. and P. H. Eichel, "High-Resolution Synthetic Aperture Radar Interferometry. Technology for Precise Terrain Elevation Mapping," *DSP and Multimedia Technology*, 10, 1994.
- Ghiglia D. C. and L. A. Romero, "Minimum Lp-Norm Two-Dimensional Phase Unwrapping," *JOSA A*, 12, 1999–2013, (1996).
- Ghiglia D. C. and M. D. Pritt, *Two Dimensional Phase Unwrapping*, Wiley, 1998.
- Gierloff J. J., "Phase Unwrapping by Regions," *Proc., SPIE*, 818, 2–9 (1987).
- Goldstein R. M., H. A. Zebker, and C. L. Werner, "Satellite Radar Interferometry: Synthetic Apertur Radar Observations," *Radio Science*, 23, 713–720 (1988).
- Goncalves A. A., Jr., R. Seara, and P. B. Uliana, "A New Amplitude Weighted Filtering Technique for Noise Reduction in Images With 2p Phase Jumps," *Proc. SPIE*, 2003, 312–322 (1993).
- Goodman J. W., *Introduction to Fourier Optics*, McGraw-Hill, San Francisco, 1968.
- Greivenkamp J. E., "Generalized Data Reduction for Heterodyne Interferometry," *Opt. Eng.*, 23, 350 (1984).
- Kinnstaetter K., A. W. Lohmann, J. Schwider, and N. Streibl, "Accuracy of Phase Shifting Interferometry," *Appl. Opt.*, 27, 5082 (1988).
- Greivenkamp J. E., "Interferometric Measurements at Eastman Kodak Company," *Proc. SPIE*, 816, 212 (1987a).
- Greivenkamp J. E., "Sub-Nyquist interferometry," *Appl. Opt.*, 26, 5245 (1987b).
- Greivenkamp J. E. and J. H. Bruning, "Phase Shifting Interferometry," in: *Optical Shop Testing*, D. Malacara, Ed., 2nd Edition, Wiley, New York, 1992.
- Greivenkamp J. E., K. G. Sullivan, and R. J. Palum, "Resolving Interferometric Step Height Measurement Ambiguities using a Priori Information," *Proc. SPIE*, 1164, 79 (1989).
- Hall J. A., "Arrays and Charge Coupled Devices," *Applied Optics and Optical Engineering*, vol. VIII, R. R. Shannon and J. C. Wyant, Eds., Academic Press, New York, 1980.
- Hardy J. W., J. E. Lefebvre, and C. L. Koliopoulos, "Real-Time Atmospheric Compensation," *JOSA*, 67, 360 (1977).
- Hardy J. W. and A. J. McGovern, "Shearing Interferometry: A Flexible Technique for Wavefront Measurement," *Proc. SPIE*, 816, 180 (1987).
- Hariharan P., *Optical Holography*, Cambridge, London, 1984, p. 232.
- Hariharan P., "Quasi-Heterodyne Hologram Interferometry," *Opt. Eng.* 24, 632 (1985).
- Hariharan P., "Phase-Stepping Interferometry: Effects of Multiple Reflected Beams," *Appl. Opt.*, 26, 2506 (1987).
- Hariharan P., "Phase-Stepping Interferometry With Laser Diodes: Effects of Changes in Laser Power with Output Wavelength," *Appl. Opt.*, 28, 27 (1989a).
- Hariharan P., "Phase-Stepping Interferometry With Laser Diodes 2: Effects of Laser Wavelength Modulation," *Appl.*, 28, 1749 (1989b).
- Hariharan P., B. F. Oreb, and C. H. Freund, "Stroboscopic Holographic Interferometry: Measurements of Vector Components of a Vibration," *Appl. Opt.*, 26, 3899 (1987a).
- Hariharan P., B. F. Oreb, and T. Eiju, "Digital Phase-Shifting Interferometry: A Simple Error-Compensating Phase Calculation Algorithm," *Appl. Opt.* 26, 2504 (1987b).

- Hariharan P., B. F. Oreb, and N. Brown, "A Digital Phase-Measurement System for Realtime Holographic Interferometry," *Opt. Comm*, **41**, 393 (1982a).
- Hariharan P., B. F. Oreb, and N. Brown, "A Digital System for Real-Time Holographic Stress Analysis," *Proc. SPIE*, **370**, 189 (1982b).
- Hariharan P., B. F. Oreb, and N. Brown, "Real-Time Holographic Interferometry: A Microcomputer System for Measurements of Vector Displacements," *Appl. Opt.*, **22**, 876 (1983).
- Hariharan P. B. F. Oreb, and Z. Wanzhi, "Measurement of Aspheric Surfaces Using a Microcomputer-Controlled Digital Radial-Shear Interferometer," *Opt. Acta*, **31**, 989 (1984).
- Hayes J. B., *Linear Methods of Computer Controlled Optical Figuring*, Ph.D. Dissertation, Optical Sciences Center, University of Arizona, Tucson, AZ, 1984.
- Hayes J. B., *Compact Micromotor Translator*, U.S. Patent 4.884.003 (1989).
- Hayes J. and S. Lange, "A Heterodyne Interferometer for Testing Laser Diodes," *Proc. SPIE*, **429**, 22 (1983).
- Henderson A. D., C. Hoffman, R. Culhane, and D. Viggiano III, "Kilohertz Scanning, All-Fiber Optical Delay Line Using Piezoelectric Actuation," *Proc. SPIE*, **5589**, 99–106 (2004).
- Herráez M. A., D. R. Burton, M. J. Lalor, and D. B. Clegg, "Robust Unwrapper for Two-dimensional Images," *Proc. SPIE*, **2784**, 106 (1996).
- Herráez M. A., D. R. Burton, M. J. Lalor, and M. A. Gdeisat, "Fast Two-dimensional Phase-unwrapping Algorithm Based on Sorting by Reliability Following a Noncontinuous Path," *Appl. Opt.*, **41**, 7437–7444 (2002).
- Hettwer A., J. Kranz, and J. Schwider, "Three Channel Phase-Shifting Interferometer Using Polarization-Optics, a Diffraction Grating," *Opt. Eng.*, **39**, 960 (2000).
- Hibino K., "Susceptibility of Systematic Error-Compensating Algorithms to Random Noise in Phase-Shifting Interferometry," *Appl. Opt.*, **36**, 2084–2093 (1997).
- Hibino K., B. F. Oreb, D. I. Farrant, and K. G. Larkin, "Phase Shifting for Nonsinusoidal Waveforms With Phase-Shift Errors," *J. Opt. Soc. Am. A*, **12**, 761 (1995).
- Hibino K., B. F. Oreb, D. I. Farrant, and K. G. Larkin, "Phase-Shifting Algorithms for Nonlinear, Spatially Nonuniform Phase Shifts," *J. Opt. Soc. Am. A*, **14**, 918 (1997).
- Hibino K., K. G. Larkin, B. F. Oreb, and D. I. Farrant, "Phase-Shifting Algorithms for Nonlinear, Spatially Nonuniform Phase Shifts: Reply to Comment," *J. Opt. Soc. Am. A*, **15**, 1234–1235 (1998).
- Hibino K., R. Hanayama, J. Burke, and B. Oreb, "Tunable Phase-Extraction Formulae for Simultaneous Shape Measurement of Multiple Surfaces With Wavelength-Shifting Interferometry," *Opt. Express*, **12**, 5579–5594 (2004).
- Hu H. Z., "Polarization Heterodyne Interferometry Using a Simple Rotating Analyzer. 1: Theory, Error Analysis," *Appl. Opt.*, **22**, 2052 (1983).
- Huntley J. M., "Noise Immune Phase Unwrapping Algorithm," *Appl. Opt.*, **28**, 3268–3270 (1989).
- Huntley J. M. and H. Saldner, "Temporal Phase Unwrapping Algorithm for Automated Interferogram Analysis," *Appl. Opt.* **32**, 3047–3052 (1993).
- Huntley J. M. and J. R. Buckland, "Characterization of Sources of 2p Phase Discontinuities in Speckle Interferograms," *JOSA A*, **12**, 1990–1996 (1995).
- Ichioka Y. and M. Inulyu, "Direct Phase Detecting System," *Appl. Opt.*, **11** (7), 1507 (1972).
- Indebetouw G., P. Klysubun, T. Kim, and T. Poon, "Imaging Properties of Scanning Holographic Microscopy," *JOSA A*, **17**, 380–390 (2000).

- Ishii Y., J. Chen, and K. Murata, "Digital Phase-Measuring Interferometry With a Tunable Laser Diode," *Opt. Lett.*, **12**, 233 (1987).
- Itoh K., "Analysis of the Phase Unwrapping Problem," *Appl. Opt.*, **21**, 2470 (1982).
- Janesick J. R., *Scientific Charge-Coupled Devices*, SPIE Press, Vol. PM83, 2001.
- Jensen A. E., "Absolute Calibration Method for Laser Twyman-Green Wavefront Testing Interferometers," *JOSA*, **63**, 1313 (1973).
- Jensen S. C., W. W. Chow, and G. N. Lawrence, "Subaperture Testing Approaches: A Comparison," *Appl. Opt.*, **23**, 740 (1984).
- Johnson, G. W., D. C. Leiner, and D. T. Moore, "Phase-locked interferometry," *Proc. SPIE*, **126**, 152 (1977).
- Johnson, G. W., D. C. Leiner, and D. T. Moore, "Phase-Locked Interferometry," *Opt. Eng.*, **18**, 46 (1979).
- Johnson, G. W. and D. T. Moore, "Design and Construction of a Phase-Locked Interference Microscope," *Proc. SPIE*, **103**, 76 (1977).
- Kanoh, T., "Automated Interferometric System for Aspheric Surface Testing," *Proc. SPIE*, **680**, 71 (1986)
- Kim C.-J., "Polynomial Fit of Interferograms," *Appl. Opt.*, **21**, 4521 (1982).
- Kim, Shannon, 1987. Kim, C. J. and R.R. Shannon, "Catalog of Zernike Polynomials", in *Applied Optics and Optical Engineering*, Vol. x, R.R. Shannon and J.C. Wyant, Eds. Academic Pres., San Diego, 1987.
- Kinnstaetter K., A. W. Lohmann, J. Schwider, and N. Streibl, "Accuracy of Phase Shifting Interferometry," *Appl. Opt.*, **27**, 5082 (1988).
- Koliopoulos C. L., *Interferometric Optical Phase Measurement*, Ph.D. Dissertation, University Arizona, Tucson, 1981.
- Koliopoulos C. L., "Radial Grating Lateral Shear Heterodyne Interferometer," *Appl. Opt.*, **19**(9), 1523 (1980).
- Kothiyal M. P. and C. Delisle, "Optical Frequency Shifter for Heterodyne Interferometry Using Counterrotating Wave Plates," *Opt. Lett.*, **9**, 319 (1984).
- Kothiyal M. P. and C. Delisle, "Shearing Interferometer for Phase Shifting Interferometry with Polarization Phase Shifter," **24**, 4439 (1985a).
- Kothiyal M. P. and C. Delisle, "Shearing Interferometer for Phase Shifting Interferometry with Polarization Phase Shifter," *Appl. Opt.*, **24**, 4439 (1985b).
- Kozlowski, L. J., K. Vural, V. H. Johnson, J. K. Chen, R. B. Bailey, D. Bui, M. J. Gubala, and J. R. Teague, "256 × 256 PACE-1 PV HgCdTe Focal Plane Arrays for Medium and Short Wavelength Infrared Applications," *Proc. SPIE*, **1308**, 202 (1990).
- Kreis T., "Digital Holographic Interference-Phase Measurement Using the Fourier-Transform Method," *JOSA A*, **3**, 847 (1986).
- Kuechel M. F. and W. Wiedmann, "In-Process Metrology for Large Astronomical Mirrors," *Proc. SPIE*, **1333**, 280 (1990).
- Kuechel M. F., "Advances in Interferometric Wavefront-Measuring Technology Through the Direct Measuring Interferometry Method," *Proc. SPIE* **1573**, 159 (1992a).
- Kuechel M. F., "Advanced Interferometry at Carl Zeiss (Summary Only)," *Proc. SPIE*, **1720**, 452 (1992b).
- Kuwamura S. and I. Yamaguchi, "Wavelength Scanning Profilometry for Real-Time Surface Shape Measurement," *Appl. Opt.*, **36**, 4473–4482 (1997).

- Kwon, O. Y., "Multichannel Phase-Shifted Interferometer," *Opt. Lett.*, **9**, 59 (1984).
- Kwon, O. Y., "Advanced Wavefront Testing at Lockheed," *Proc. SPIE*, **816**, 196 (1987).
- Kwon O. Y., D. M. Shough, and R. A. Williams, "Stroboscopic Phase-Shifting Interferometry," *Opt. Lett.*, **12**, 855 (1987).
- Kwon O., J. C. Wyant, and C. R. Hayslett, "Rough Surface Interferometry at 10.6 μm ," *Appl. Opt.*, **19**(11), 1862 (1980).
- Larkin K. G., and B. F. Oreb, "Design, Assessment of Symmetrical Phase-Shifting Algorithms," *J. Opt. Soc. Am. A*, **9**, 1740 (1992).
- Li E., J. Yao, D. Yu, J. Xi, and J. Chicharo, "Optical Phase Shifting With Acousto-Optic Devices," *Opt. Lett.*, **30**, 189–191 (2005).
- Liu Y. M., G. N. Lawrence, and C. L. Koliopoulos, "Subaperture Testing of Aspheres with Annular Zones," *Appl. Opt.*, **27**, 4504 (1988).
- Macy W. W., Jr., "Two-Dimensional Fringe-Pattern Analysis," *Appl. Opt.*, **22**, 3898 (1983).
- Malacara D., M. Servín, and Z. Malacara, *Interferogram Analysis for Optical Testing*, 2nd Ed., CRC Press, Boca Raton, (2005).
- Malacara D., Analysis of the Interferometric Ronchi Test," *Appl. Opt.*, **29**, 3633 (1990).
- Malacara-Doblado D. and B. V. Dorrió, "Family of Detuning-Insensitive Phase-Shifting Algorithms: Erratum," *J. Opt. Soc. Am. A*, **18**, 721–721 (2001a).
- Malacara-Doblado D. and B. V. Dorrió, "Family of Detuning-Insensitive Phase-Shifting Algorithms," *J. Opt. Soc. Am. A* **17**, 1857–1863 (2000a).
- Malacara-Doblado D., B. V. Dorrió, and D. Malacara-Hernandez, "Graphic Tool to Produce Tailored Symmetrical Phase-Shifting Algorithms," *Opt. Lett.*, **25**, 64–66 (2000b).
- Malacara-Doblado, M. F. Kuechel, R. Rojas-Laguna, and E. Vargas-Rodriguez, "Wide-Range Detuning Insensitive Phase-Shifting Algorithms," *Proc. SPIE*, **4419**, 269 (2001).
- Massie, N. A., "Digital Heterodyne Interferometry," *Proc. SPIE*, **816**, 40 (1987).
- Massie N. A., "Real-Time Digital Heterodyne Interferometry: A System," *Appl. Opt.*, **19**, 154 (1980).
- Massie N. A., J. Hartlove, D. Jungwirth, and J. Morris, "High Accuracy Interferometric Measurements of Electron-Beam Pumped Transverse-Flow Laser Media With 10-Picosecond Time Resolution," *Appl. Opt.*, **20**, 2372 (1981).
- Massie N. A., M. Dunn, D. Swain, S. Muenter, and J. Morris, "Measuring Laser Flow Fields With a 64-Channel Heterodyne Interferometer," *Appl. Opt.*, **22**, 2141 (1983).
- Massie N. A. and R. D. Nelson, "Beam Quality of Acousto-Optic Frequency Shifters," *Opt. Lett.*, **3**, 46 (1978).
- Massie N. A., R. D. Nelson, and S. Holly, "High-Performance Real-Time Heterodyne Interferometry," *Appl. Opt.*, **18**, 1797 (1979).
- Matthews H. J., D. K. Hamilton, and C. J. R. Sheppard, "Surface Profiling by Phase-Locked Interferometry," *Appl. Opt.*, **25**, 2372 (1986).
- McLaughlin J. L. and B. A. Horwitz, "Real Time Snapshot Interferometer," *Proc. SPIE*, **680**, 35 (1986).
- Mertz L. N., "Optical Homodyne Phase Metrology," *Appl. Opt.*, **28**, 1011 (1989).
- Mertz L., "Complex Interferometry," *Appl. Opt.*, **22**, 1530 (1983).
- Mertz L., "Real-Time Fringe-Pattern Analysis," *Appl. Opt.*, **22**, 1535 (1983).

- Millerd J. E., N. J. Brock, J. B. Hayes, M. B. North-Morris, M. Novak, and J. C. Wyant, "Pixelated Phase-Mask Dynamic Interferometer," *Proc. SPIE*, **5531**, 304 (2004).
- Millerd J. E., N. J. Brock, J. Hayes, B. Kimbrough, M. Novak, M. North-Morris, and J. C. Wyant, "Modern Approaches in Phase Measuring Metrology," *Proc. SPIE*, **5856**, 14 (2005).
- Moore, D. T., *Gradient Index Optics Design and Tolerancing*, Ph.D. Dissertation, Institute of Optics, University of Rochester, Rochester, NY, (1973).
- Moore D. T. and B. E. Truax, "Phase-Locked Moire Fringe Analysis for Automated Contouring of Diffuse Surfaces," *Appl. Opt.*, **18**, 91 (1979).
- Moore D. T., R. P. Murray, and F. B. Neves, "Large Aperture ac Interferometer for Optical Testing," *Appl. Opt.*, **17**, 3959 (1978).
- Moore R. C. and F. H. Slaymaker, "Direct Measurement of Phase in a Spherical-Wave Fizeau Interferometer," *Appl. Opt.*, **19**, 2196 (1980).
- Morgan C. J., "Least-Squares Estimation in Phase-Measurement Interferometry," *Opt. Lett.*, **7**, 368 (1982).
- Mottier, F. M., "Microprocessor Based Automatic Heterodyne Interferometer," *Opt. Eng.*, **18**, 464 (1979).
- Nakadate S. and H. Saito, "Fringe Scanning Speckle-Pattern Interferometry," *Appl. Opt.*, **24**, 2172 (1985).
- Negro J. E., "Subaperture Optical System Testing," *Appl. Opt.*, **23**, 1921 (1984).
- Nugent K. A., "Interferogram Analysis Using an Accurate Fully Automatic Algorithm," *Appl. Opt.*, **24**, 3101 (1985).
- Ohyama N., S. Kinoshita, A. Cornejo-Rodriguez, T. Honda, and J. Tsujiuchi, "Accuracy of Phase Determination With Unequal Reference Phase Shift," *JOSA A*, **5**, 2019 (1988).
- Okada K., H. Sakuta, T. Ose, and J. Tsujiuchi, "Separate Measurements of Surface Shapes, Refractive Index Inhomogeneity of an Optical Element Using Tunable-Source Phase Shifting Interferometry," *Appl. Opt.*, **29**(22), 3280–3285 (1990).
- Onodera R. and Y. Ishii, "Phase-Shift-Locked Interferometer With a Wavelength-Modulated Laser Diode," *Appl. Opt.*, **42**, 91–96 (2003).
- Oppenheim A.V. and R.W. Schafer, *Digital Signal Processing*, Englewood Cliffs, NJ: Prentice-Hall 1975.
- Pez G. and M. Strojnik, "Fringe Analysis, Phase Reconstruction From Modulated Intensity Patterns," *Opt. Lett.*, **22**, 1669–1671 (1997).
- Pfoertner A. and J. Schwider, "Dispersion Error in White-Light Linnik Interferometers, its Implications for Evaluation Procedures," *Appl. Opt.*, **40**, 6223–6228 (2001).
- Phillion D. W., "General Methods for Generating Phase Shifting Interferometry Algorithms," *Appl. Opt.*, **36**, 8098–8115 (1997).
- Powell R. L. and K. A. Stetson, "Interferometric Vibration Analysis by Wavefront Reconstruction," *JOSA*, **55**(12), 1593 (1965).
- Prettyjohns, K.N., "Change-Coupled Device Image Acquisition for Digital phase Measurement Interferometry," *Opt. Eng.*, **23**, 371 (1984).
- Prettyjohns K. N., S. DeVore, E. Dereniak, and J. C. Wyant, "Direct Phase Measurement Interferometer Working at 3.8 Mum," *Appl. Opt.*, **24**, 2211 (1985).
- Pritt M. D., "Automated Subpixel Image Registration of Remotely Sensed Imagery," *IBM J. Res. Development*, **38**, 157–166 (1994).

- Roddiar C. and F. Roddiar, "Interferogram Analysis Using Fourier Transform Techniques," *Appl. Opt.*, **26**, 1668 (1987).
- Rosenbluth A. E. and N. Bobroff, "Optical Sources of Nonlinearity in Heterodyne Interferometers," *Prec. Eng.* **12**, 7 (1990).
- Schmit J. and K. Creath, "Window Function Influence on Phase Error in Phase-Shifting Algorithms," *Appl. Opt.*, **35**, 5642–5649 (1996).
- Schmit J. and K. Creath, "Extended Averaging Technique for Derivation of Error-Compensating Algorithms in Phase-Shifting Interferometry," *Appl. Opt.* **34**, 3610 (1995).
- Schreiber H. and J. Schwider, "Lateral Shearing Interferometer Based on Two Ronchi Phase Gratings in Series," *Appl. Opt.* **36**, 5321–5324 (1997).
- Schreiber H., P. G. Dewa, K. Hanfrod, R. Ligenza, L. R. Rich, and P. J. Tompkins, "157-nm System Test for High-NA Lithographic Lens Systems," *Proc. SPIE*, **4691**, 1675 (2002).
- Schulz G. and J. Schwider, "Interferometric Testing of Smooth Surfaces," in: *Progress in Optics*, vol. XIII, E. Wolf, ed. North Holland, Amsterdam, 1976, pp. 93–167.
- Schwider J., "Detection of Undersampling From Measured Phase-Shifting Data," *Opt. Lett.*, **19**, 231 (1994).
- Schwider J., "Phase Shifting Interferometry: Reference Phase Error Reduction," *Appl. Opt.*, **28**, 3889 (1989).
- Schwider J., O. R. Falkenstoerfer, H. Schreiber, A. Zoeller, and N. Streibl, "New Compensating Four-Phase Algorithm for Phase-Shift Interferometry," *Opt. Eng.*, **32**, 1883 (1993).
- Schwider J., R. Burow, K.-E. Elssner, J. Grzanna, R. Spolaczyk, and K. Merkel, "Digital Wave-Front Measuring Interferometry: Some Systematic Error Sources," *Appl. Opt.*, **22**, 3421 (1983).
- Seara R., A. A. Goncalves, and P. B. Uliana, "Filtering Algorithm for Noise Reduction in Phase-Map Images with 2 Phase Jumps," *Appl. Opt.* **37**, 2046–2050 (1998).
- Selber, L. A., "Interferometer Accuracy and Precision," *Proc. SPIE*, **749**, 8 (1987).
- Selber, L. A., "Interferometer Accuracy and Precision," *Proc. SPIE*, **1400**, 24 (1990a).
- Seligson, J. L., C. A. Callari, J. E. Greivenkamp, and J. W. Ward, "Stability of a Lateral-Shearing Heterodyne Twyman–Green Interferometer," *Opt. Eng.*, **23**, 353 (1984).
- Shagam, R. N., "AC Phase Measurement Technique for Moire Interferograms," *Proc/ SPIE* **429**, 35 (1983).
- Shough D., Kwon O.Y., and Leary D. F., " High-Speed Interferometric Measurement of Aerodynamic Phenomena," *Proc. SPIE*, **1221**, 394 (1990).
- Smythe R. A. and R. Moore, "Instantaneous Phase Measuring Interferometry," *Opt. Eng.* **23**, 361 (1984).
- Smythe R.A., J. A. Soobitsky and B. E. Truax, "Recent Advances in Interferometry at Zygo," *Proc. SPIE*, **816**, 95 (1987).
- Sommargren G. E., "Up/down Frequency Shifter for Optical Heterodyne Interferometry," *J. Opt. Soc. Am.*, **65**, 960 (1975).
- Sommargren G. E., "Optical Heterodyne Profilometry," *Appl. Opt.*, **20**, 610 (1981).
- Soobitsky J. A., *Piezoelectric Micromotion Actuator*, U.S. Patent No. 4,577,131, 1986.
- Stadelmaier A. and J. H. Massig, "Compensation of Lens Aberrations in Digital Holography," *Opt. Lett.*, **25**, 1630–1632 (2000).
- Stahl H. P., Christ L. K., "Interferometric Phase Measurement Using Pyroelectric Vidicons," *Appl. Opt.*, **26**, 1127 (1987).

- Stahl H. P., "IR Interferometry Using Pyroelectric Vidicons," *Lasers Optronics*, **57** (Sept. 1989).
- Steinmetz C. R., "Sub-Micron Position Measurement and Control on Precision Machine Tools With Laser Interferometry," *Precision Eng.*, **12**, 12 (1990).
- Steinmetz C. R., R. Burgoon, and J. Harris, "Accuracy Analysis and Improvements to the Hewlett-Packard Laser Interferometer System," *Proc. SPIE*, **816**, 79 (1987).
- Stephenson D., "Forget Finding the Problem . . . Just Fix it," *Proc. SPIE*, **1049**, 187 (1989).
- Stern A. and B. Javidi, "Three-Dimensional Image Sensing, Reconstruction With Time-Division Multiplexed Computational Integral Imaging," *Appl. Opt.*, **42**, 7036–7042 (2003).
- Stevenson W. H., "Optical Frequency Shifting by Means of a Rotating Diffraction Grating," *Appl. Opt.*, **9**, 649 (1970).
- Stuhlinger, T. W., "Subaperture Optical Testing: Experimental Verification," *Proc. SPIE*, **656**, 118 (1986).
- Stump K. D., "Real-Time Interferometer," *Opt. Eng.* **18**, 648 (1979).
- Surrel Y., "Phase-Shifting Algorithms for Nonlinear, Spatially Nonuniform Phase Shifts: Comment," *J. Opt. Soc. Am. A*, **15**, 1227–1233 (1998).
- Surrel Y., "Additive Noise Effect in Digital Phase Detection," *Appl. Opt.*, **36**, 271–276 (1997a).
- Surrel Y., "Design of Algorithms for Phase Measurements by the Use of Phase Stepping," *Appl. Opt.*, **35**, 51 (1996).
- Surrel Y., "Design of Phase-Detection Algorithms Insensitive to Bias Modulation," *Appl. Opt.* **36**, 805–807 (1997b).
- Surrel Y., "Phase Stepping: A New Self-Calibrating Algorithm," *Appl. Opt.*, **32**, 3598 (1993).
- Suzuki T., X. Zhao, and O. Sasaki, "Phase-Locked Phase-Shifting Laser Diode Interferometer With Photothermal Modulation," *Appl. Opt.* **40**, 2126–2131 (2001).
- Suzuki T., O. Sasaki, and T. Maruyama, "Phase Locked Laser Diode Interferometry for Surface Profile Measurement," *Appl. Opt.*, **28**, 4407 (1989).
- Suzuki, T. and R. Hioki, "Translation of Light Frequency by a Moving Grating," *J. Opt. Soc. Am.*, **57**, 1551 (1967).
- Synborski C. E., "The Interferometric Analysis of Flatness by Eye and Computer," *Proc. SPIE*, **135**, 104 (1978).
- Takeda M., H. Ina, and S. Kobayashi, "Fourier-Transform Method of Fringe-Pattern Analysis for Computer-Based Topography, Interferometry," *JOSA*, **72**, 156 (1982).
- Takeda M., Q. Gu, M. Kinoshita, H. Takai, and Y. Takahashi, "Frequency-Multiplex Fourier-Transform Profilometry: A Single-Shot 3-Dimensional Shape Measurement of Objects With Large Height Discontinuities," *Appl. Opt.*, **36**, 5347–5354 (1997).
- Tarbeyev Y. V., *Proceedings of 3rd IMEKO TC-8-Symposium, Theor. Metrology*, Berlin GDR, 1986, p. 5.
- Tatsuno K. and Y. Tsunoda, "Diode Laser Direct Modulation Heterodyne Interferometer," *Appl. Opt.*, **26**, 37 (1987).
- Thalman R. and R. Dändliker, "Holographic Contouring Using Electronic Phase Measurement," *Opt. Eng.* **24**, 930 (1985).
- Thunen J. G. and O. Y. Kwon, "Full Aperture Testing With Subaperture Test Optics," *Proc. SPIE*, **351**, 19 (1982).
- Towers D. P., T. R. Judge, and P. J. Bryanston-Cross, "A Quasi-Heterodyne Holograph Technique and Automatic Algorithms for Phase Unwrapping," *Proc. SPIE* 1163, 95 (1989).

- Toyooka S. and M. Tominaga, "Spatial Fringe Scanning for Optical Phase Measurement," *Opt. Comm.* **51**, 68 (1984).
- Tricard M., P. Dumas, G. Forbes, and M. DeMarco, "Recent Advances in Sub-Aperture Approaches to Finishing and Metrology," *Proc. SPIE*, **6149**, 614903 (2006).
- Truax B. E., "Programmable Interferometry," *Proc. SPIE*, **680**, 10 (1986).
- Truax B. E., "Absolute Interferometric Testing of Spherical Surfaces," *Proc. SPIE*, **966**, 130 (1988).
- Watkins L. R., S. M. Tan, and T. H. Barnes, "Determination of Interferometerphase Distributions by Use of Wavelets," *Opt. Lett.*, **24**, 905–907 (1999).
- Wiley G. and R. Patchin, "Optical Design Analysis Incorporating Actual System Interferometric Data," *Proc. SPIE* **1168**, 176 (1989).
- van Wingerden J., H. J. Frankena, and C. Smorenburg, "Linear Approximation for Measurement Errors in Phase Shifting Interferometry," *Appl. Opt.*, **30**, 2718 (1991).
- Wizinowich, P. L., "System for Phase Shifting Interferometry in the Presence of Vibration," *Proc. SPIE* 1164, **25** (1989).
- Wizinowich P. L., "Phase Shifting Interferometry in the Presence of Vibration: A New Algorithm, System," *Appl. Opt.*, **29**, 3271 (1990).
- Womack K. H., "Interferometric Phase Measurement Using Spatial Synchronous Detection," *Opt. Eng.* **23**, 391 (1984a).
- Womack K. H., "Fourier Domain Description of Interferogram Analysis," *Opt. Eng.* **23**, 396 (1984b).
- Wyant J. C., "Interferometric Testing of Aspheric Surfaces," *Proc. SPIE*, **816**, 19 (1987).
- Wyant J. C., "Double Frequency Grating Lateral Shear Interferometer," *Appl. Opt.*, **12**, 2057 (1973).
- Wyant J. C., "Use of an ac Heterodyne Lateral Shear Interferometer with Real-Time Wavefront Correction Systems," *Appl. Opt.*, **14**, 2622 (1975).
- Wyant J. C., B. F. Oreb, and P. Hariharan, "Testing Aspherics Using Two-Wavelength Holography: Use of Digital Electronic Techniques," *Appl. Opt.*, **23**, 4020 (1984a).
- Wyant J. C., C. L. Koliopoulos, B. Bhushan, and O. E. George, "An Optical Profilometer for Surface Characterization of Magnetic Media," *ASLE Trans.* **27**, 101 (1984b).
- Wyant J. C. and K. Creath, "Two-Wavelength Phase-Shifting Interferometer and Method," U.S. Patent No. 4,832,489 (1989).
- Wyant J. C. and K. N. Prettyjohns, *Optical Profiler Using Improved Phase-Shifting Interferometry*, U.S. Patent No. 4,639,139 (1987).
- Wyant J. C. and R. N. Shagam, "Use of Electronic Phase Measurement Techniques in Optical Testing," *Proceedings of ICO-11*, International Congress of Optics Madrid, 1978, p. 659.
- Zanio K., "HgCdTe on Si for Hybrid and Monolithic FPAs," *Proc. SPIE*, **1308**, 180 (1990).
- Zhao B. and Y. Surrel, "Effect of Quantization Error on the Computed Phase of Phase-Shifting Measurements," *Appl. Opt.*, **36**, 2070–2075 (1997).
- Zhu Y. and T. Gemma, "Method for Designing Error-Compensating Phase-Calculation Algorithms for Phase-Shifting Interferometry," *Appl. Opt.* **40**, 4540–4546 (2001).
- Zvyagin A. V. and D. D. Sampson, "Achromatic Optical Phase Shiftermodulator," *Opt. Lett.*, **26**(4), 187–189 (2001).

15

Surface Profilers, Multiple Wavelength, and White Light Interferometry

J. Schmit, K. Creath, and J. C. Wyant

15.1. INTRODUCTION TO SURFACE PROFILERS

Over the last 25 years driven by both the development of new technologies such as fast computers and solid state devices and the necessity to precisely inspect these increasingly tiny engineering surfaces, the field of surface metrology has exploded in both its technological sophistication and its range of application. Advances in illumination sources, such as lasers, and in solid state detectors and optoelectronic devices in general have fueled the development of a wide range of instruments that can not only map surface topography but also determine other features such as displacement or dispersion. Innovative techniques and technologies have greatly increased the range of measurable objects, so now even difficult surfaces with high slopes or steps and narrow, deep trenches can be measured. Many of these surface profiling techniques were developed from distance measuring or focus detection techniques, and they often require scanning to obtain the surface profile. This chapter describes instruments such as the stylus profiler, scanning probe microscope, confocal microscope and the interferometric optical profiler that are most often used to determine surface topographies of not only very small, typically engineering surfaces, but also smooth and large surfaces such as aspheres and glass plates.

The first part of this chapter describes scanning probe microscopes and stylus profilers. Then optical methods are detailed with a specific focus on techniques developed over the last 15 years that have found commercial and industrial application. This section describes both interference microscopes that employ both monochromatic and white light illumination and also confocal microscopes that have recently been fairly used to measure engineering surfaces. The next part of this chapter reviews work done in multiple wavelength interferometry, namely two and multiple wavelength, wavelength scanning, and spectrally resolved white light

interferometry. White light and multiple wavelength techniques are often applied to other methods, such as speckles and holography as well as fringe and structured light projection procedures. Finally, we provide a short overview of optical ranging techniques and polarization interferometers.

15.1.1. Contact Profilometers

Stylus Profilers and Scanning Probe Microscopes. These instruments are often described as contact profilers, or tactile sensors, because they use a probe to scan along the surface of an object. Nowadays the forces applied to the tips of these probes are now so minute that it is rather difficult to call them contact instruments anymore. These instruments trade relatively slow measurement speed for excellent lateral resolution, often below the optical resolution of optical microscopes. The stylus profiler, one of the first to be developed, is both inexpensive to build and easy to use; for these reasons it is a standard in many mechanical and optical shops. The scanning probe microscope was developed later but is now common in both labs and the semiconductor and micro-electromechanical systems (MEMS) industries, because of its subnanometer vertical and lateral resolution and its ability to measure many different material properties. The profile of larger objects can be measured using a coordinate measuring machine (CMM) that uses a larger probe and is able to work in scanning mode.

15.1.2. Optical Profilometers

Optical Focus Sensors. Instead of using a mechanical probe as contact profilers do, a broad group of instruments called optical profilometers use an optical focus sensor to obtain profile measurements. A few optical focus sensing techniques that have been developed into commercial profilometers are detailed in this chapter.

Confocal Microscope. The most commercially successful of the optical focus sensing systems is the confocal microscope, which was initially developed to examine biological samples and more recently has found applications for testing engineering surfaces. The confocal microscope is a modified conventional microscope with a single point source and a pinhole placed in front of the detector to filter out spurious light so as to obtain a more distinct irradiance signal at the focus position.

15.1.3. Interferometric Optical Profilometers

Two and Multiple Wavelength Interferometry. The first interferometric optical profilometers used monochromatic illumination and phase shifting methods. Phase shifting methods generated the highest measurement precision; however, monochromatic illumination limited the measurement range. In order to increase the range of these measurements, two and multiple wavelength techniques were developed. Today, the principles of two and multiple wavelength interferometry can be found in holographic and speckle techniques and are applied to the testing of large objects using a fringe projection and structured light techniques (see Chapter 16).

White Light Interferometry. White light interferometric systems have long established themselves as the leading optical profilometers for measuring engineering surfaces. White light interferometry can be thought of as an optical focus sensor where the position of the interference signal determines best focus. Typically, these setups consist of a conventional microscope outfitted with an interferometric objective. The advantage of interferometric systems over most other optical and stylus profilers is their ability to scan the entire field rather than proceed by point to point scanning. Because the whole area is imaged at the same time, lateral scanning becomes unnecessary, which greatly speeds up the measurement process.

Spectral Interferometry. Another group of interferometers is based on the observation of interference fringes for a very large number of wavelengths called spectral interference fringes. Spectral interference can be obtained by using a source with a tunable wavelength or by placing a dispersive element at the exit of the interferometer. These systems are called wavelength scanning interferometers and spectrally resolved white light interferometers, respectively. Rather than detecting best focus at each point, the fringe frequency which carries information about the object's position is measured. This method bypasses the need for mechanical axial scanning.

Optical Ranging Sensors. The techniques used in interferometric optical profilers are often built on or independently developed from techniques found in interferometric optical ranging sensors, which typically measure absolute distances or lengths at a single point. For single point detection, much faster detectors and optoelectronics as well as task-specific signal processing techniques can be employed.

Polarization Interferometers. Some interferometers utilize the polarization properties of light so as to have two beams traveling almost the same path in the interferometer like in the differential interference microscope. Polarization interferometers that use polarization to shift the phase often can be made to be insensitive to vibrations. The polarization properties of light are effective in evaluating some properties of submicron structures.

15.1.4. Terms and Issues in Determining System Performance

All profilometers need to be carefully calibrated especially when measuring small objects such as MEMS. Various aspects of system performance are checked against artifacts that have a traceability certificate to some primary standard. The most common artifact is a step; for example, a step artifact is used for vertical scale calibration while a sample with an etched binary grating is often used for lateral magnification calibration. Other artifacts like sinusoidal gratings or surfaces of different roughness are also used to verify a system's performance. Ideally all measurements should be traceable to the same units. Standards institutes, such as the National Institute of Standards and Technology (NIST), Physikalisch-Technische Bundesanstalt (PTB) or National Research Council (NRC) in Canada, are continuously developing artifacts and measured parameters so as to provide accurate calibration and verification of systems. Artifacts become standards after they are measured using traceable stylus profilometers.

Manufacturers often give, and customers require, as a parameter of system performance the value the vertical resolution; however, rarely is the information provided as to how the manufacturer measured and calculated the parameter and on which artifact. Without knowing how the value of a particular parameter is determined, comparing these values across different manufacturers or systems is worthless.

Along with determining system performance, the other two really important issues in measurement are repeatability and reproducibility. In general repeatability is defined as one sigma standard deviation of a parameter of an object measured multiple times over a short period of time and without any changes in the system. Reproducibility refers to the distribution of multiple measurements over a longer period of time and under different measurement conditions. Specifications for a system are typically presented in terms of the repeatability and reproducibility of certain parameters of a measured object. It is important that the measurement procedure and reported specification values are well defined and agreed upon by user and manufacturer.

Two terms that are often misunderstood in surface metrology are accuracy and precision. Accuracy determines how close the measured value is to the true value (for example, the value of the certified artifact), and precision refers to the distribution of the measurement and can be expressed in terms of repeatability or reproducibility.

15.2. CONTACT PROFILOMETERS

The two main contact profilometers are the stylus profiler and scanning probe microscope, and they use a tactile probe to measure the surface profile. Their measurements differ in lateral and vertical ranges and their resolution, and thus they find different applications. The scanning probe microscope, in addition, measures sample-tip interaction, which allows for the measurement of materials different properties.

15.2.1. Stylus Profilers

Stylus profilers move a small-tipped probe across the surface and sense height variations of the tip to determine the surface height profile. Stylus profilers can measure surfaces up to about one millimeter in height. These profilers work very much like a phonograph; usually the surface is moved under the stylus tip, but the stylus may also be moved over the surface. The vertical motion of the stylus is typically detected by a linear variable differential transformer (LVDT) and this signal is converted to height data. The styli are made of a hard material such as diamond with a tip radius of curvature between 0.05 and 50 μm , which determines the instruments' lateral resolution. To ensure that the test surface is not damaged during measurement, the load of the stylus tip on the surface is variable from 0.1 mg up to 50 mg. A minimum load that keeps the stylus on the surface is chosen so that the surface is not deformed as the stylus moves across it. A schematic of a stylus profiler

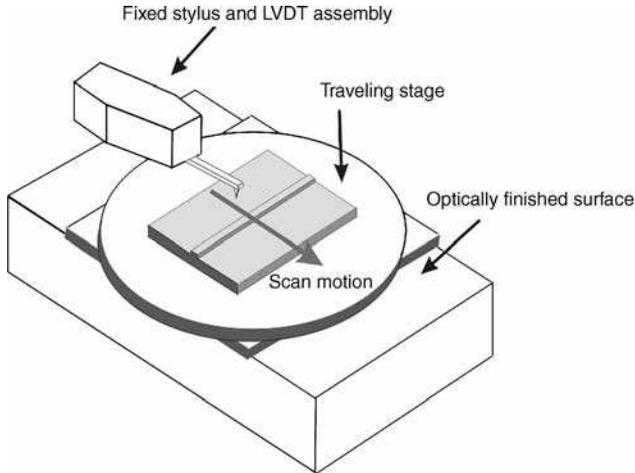


FIGURE 15.1. Schematic of stylus profiler with LVDT as the motion detector. Courtesy Veeco Instruments.

with LVDT as the motion detector is shown in Figure 15.1. Other schemes of tip guidance and its motion detection are possible (Whitehouse, 1997).

The stylus tip shown in Figure 15.2 has a 45° cone angle, but many other shapes and angles are possible. Both the shape and the angle determine the penetration depth of the tip on the test surface. The output of these profilers is the convolution of the size and shape of the stylus tip with the surface profile. Choosing the configuration of the tip is extremely important to ensure penetration to the bottom of steep trenches and prevent rounding off of high surface peaks. Figure 15.3 shows the effect of a stylus tip on the measurement of trenches with various aspect ratios.

However, while a tip radius that is smaller and sharper allows the stylus to follow the shape of the surface more easily, if the tip is too sharp, the local force on the surface over the tip area may be so great that the surface becomes locally deformed. If the surface elastically deforms, the sample will not be damaged but the surface profile may be inaccurate. If the surface plastically deforms, the sample may be permanently damaged and the surface profile will be inaccurate. New, low force technology (less than a milligram tip loading) allows for the measurement of soft materials such as a photoresist. In addition, when a small-radius stylus is used, the scan speed must be greatly reduced, and similarly the stylus load must be reduced to ensure a precise measurement. The most accurate stylus profilers have tip radii of tenths of a micrometer or less and tip loadings of milligrams or less. These profilers also may require enclosures and vibration isolation systems, and completing a scan of a few thousand data points can take many minutes.

The lateral resolution of stylus profilers is determined by the radius of the stylus tip as well as the surface shape and the sampling interval between data points. For a stylus with a spherical tip measuring a sinusoidal surface profile, the shortest

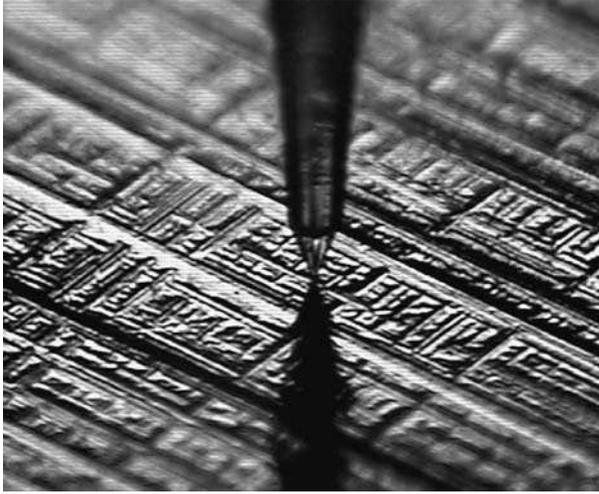


FIGURE 15.2. A stylus tip with a $0.2\ \mu\text{m}$ radius and a 45° cone angle maps the surface of a roughness comparator strip. The camera attached to the stylus profiler observes the position of the tip with respect to the object. Courtesy Veeco Instruments.

measurable wavelength (period) d of the sinusoid depends not only on the stylus radius r but also on the amplitude of the sinusoid a (Bennett and Dancy, 1981). The equation describing the shortest measurable period d is

$$d = 2\pi\sqrt{ar} \tag{15.1}$$

Because two samples per sinusoidal period are required to reconstruct a sinusoid, the lateral resolution will be $d/2$. This means that for a stylus of $10\ \mu\text{m}$ radius measuring $1\ \text{nm}$ surface height variations, the lateral resolution is approximately $0.6\ \mu\text{m}$. To ensure sufficient resolution, it is best to oversample and measure at least four samples

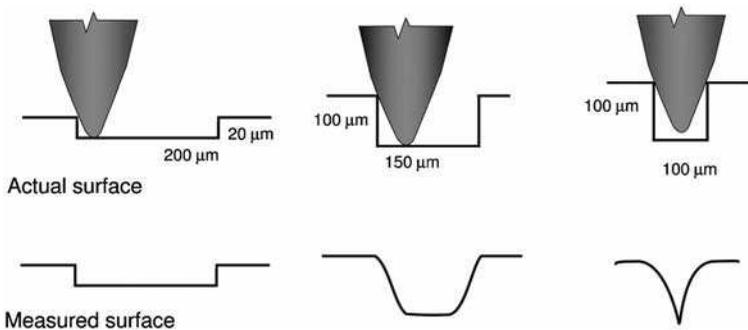


FIGURE 15.3. The convolution of a $25\ \mu\text{m}$ radius stylus tip with a surface profile.

per lateral resolution element (Bennett and Mattson, 1989). Lateral resolution and transfer functions for more complex surface features can also be determined (Al-Jumaily et al., 1987; Bennett and Dancy, 1981). The profile (and radius) of the stylus tip can be determined by viewing the tip with a scanning-electron microscope (SEM) or by scanning it over the edge of a razor blade (Vorburger and Raja, 1990). The smallest stylus tips available on the market have a radius of about 50 nm, they are often etched with a focused ion beam (see Fig. 15.4). These small tips significantly improve the lateral resolution of the profiler.

The cone angle of the stylus tip also determines the measured aspect ratio of the trenches, which typically is 1:1 for a common 60° cone angle. With the necessity of measuring structures with high aspect ratios such as MEMS, sharp styli have been developed to measure trenches with aspect ratios as high as 10:1. Custom tip geometries, that is “chisel-types,” can also be fabricated with today’s technology to measure challenging samples, like solder bumps in integrated circuit packages.

Most stylus profilers have reference datums of some type to ensure measurement accuracy (Vorburger and Raja, 1990). The reference surface can be a skid, that is, moved across the surface with the stylus, or can be a separate reference surface so that another large-radius probe is moved across in a fixed relationship to the measuring stylus. References can also be created using flexures (Vorburger and Raja, 1990). Using a large radius skid near the stylus is the easiest way to generate a reference, but this technique can cause errors and will remove shape and figure information. A separate reference is most accurate but can limit the length of the scan and the measurable height variation. Optical flats with flatness $\lambda/20$ provide a very stable reference.

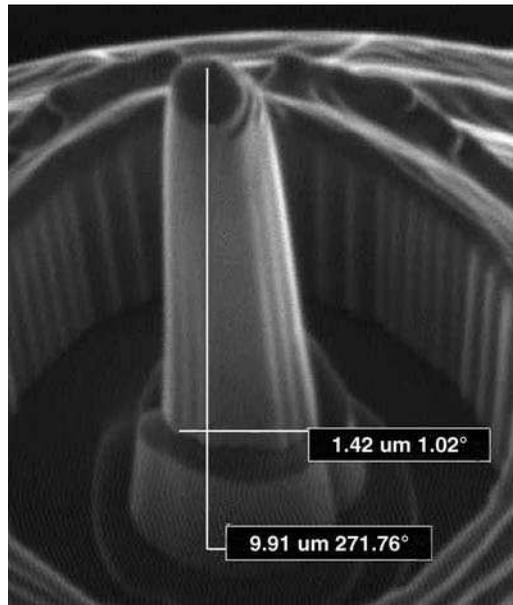


FIGURE 15.4. Scanning electron micrograph of 50 nm radius stylus tip. Courtesy Veeco Instruments.

Stylus profilers are normally calibrated using traceable height and roughness standards; these standards can be purchased from VSLI Standard Incorporated or PTB (Physikalisch-Technische Bundesanstalt, Germany). The most common standards are step of chrome on fused glass or etched in silicon. The step is measured periodically with the profiler to ensure calibration and a scaling factor is calculated to apply to the profile data. Some stylus profilers are not linear over their entire height range; it is important to calibrate the instrument with a step height which is close in height to the test samples being measured. When surface roughness is being determined, it is better to use a roughness standard than a step height standard because both lateral resolution and surface height variation need to be considered. These standards are available in a number of different types. The most common have a sinusoidal height variation with a given amplitude and a number of different spatial wavelengths. Roughness standards are also available as square-wave gratings. Because the stylus may not get down into the valleys and can round off peaks, the sinusoidal standards give a more accurate indication of instrument performance at a single spatial frequency.

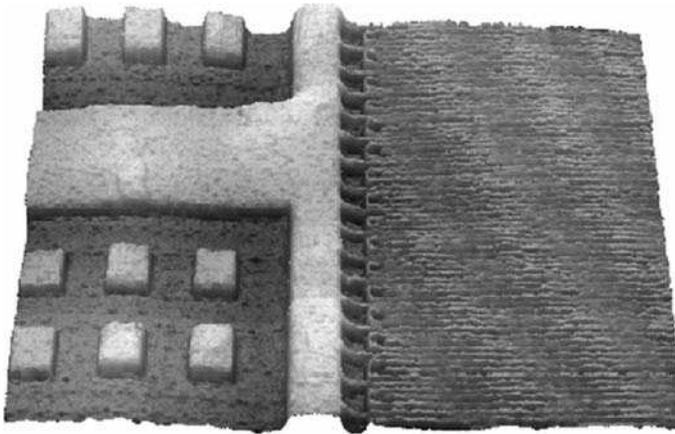
Stylus profilers are capable of measuring surface roughness with a root-mean-square (RMS) as small as 0.5 \AA with lateral resolutions of 0.1 to $0.2 \text{ }\mu\text{m}$. The instrument noise measured at a single point without scanning the surface can be as small as 0.5 \AA RMS. Stylus profilers are capable of measuring 100 nm step height with repeatability of 6 \AA and $60 \text{ }\mu\text{m}$ step height with 7 nm repeatability. Since the stylus profiler is a contact instrument, in order to measure film thickness, the step of the film to substrate is needed.

The stylus profiler has a wide range of applications in general metrology and the semiconductor industry due to its high lateral and vertical range. Stylus profilers are often used when profiles of long surfaces up to 200 mm in length are needed. Profiles of this type are typically obtained in single scans and then stitched together. Multiple scans at low force reduce the possibility of damaging the sample. Other typical applications include a scratch test for measuring the thickness and hardness of a protective coating, wafer planarity, and etch depth rate uniformity across a wafer, testing the stress that thin films induce on a wafer, RGB color filters on flat panel displays, and flip-chip bumps, and monitoring wet etching of MEMS. Figure 15.5 shows the results of a few applications of stylus profiler measurements. Some profilometers are designed specifically to measure aspheric surfaces and their roughness. These systems can measure aspheres up to 12 mm in width, 38 mm in height and over 200 mm in length. A review of recent advances in aspherics measurements using stylus profiles was done by Scott (Scott, 2002).

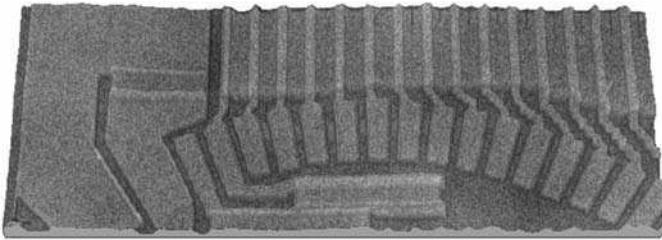
An overview of surface metrology including the stylus profiler, surface characterization and a review of optical methods can be found in a number of sources (Stout and Blunt, 1994; Thomas, 1999; Lehmann, 2003; Whitehouse, 1997, 2003).

15.2.2. Scanning Probe Microscopes

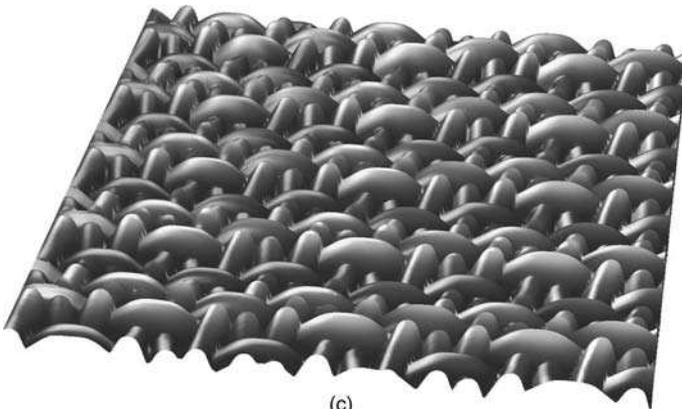
Scanning probe microscopes (SPMs), which are capable of obtaining atomic scale resolution, work by moving a fine tip in close proximity to a test surface. They



(a)



(b)



(c)

FIGURE 15.5. A few examples of stylus profile measurements: (a) Cu line connect $24\ \mu\text{m} \times 45\ \mu\text{m}$ (b) automotive sensor, scan $40\ \text{mm} \times 15\ \text{mm}$, (c) polyester mesh, scan $14\ \text{mm} \times 14\ \text{mm}$ (d) binary optics, scan $240\ \mu\text{m} \times 210\ \mu\text{m}$. (a–b) courtesy KLA-Tencor, (c–d) courtesy Veeco Instruments.

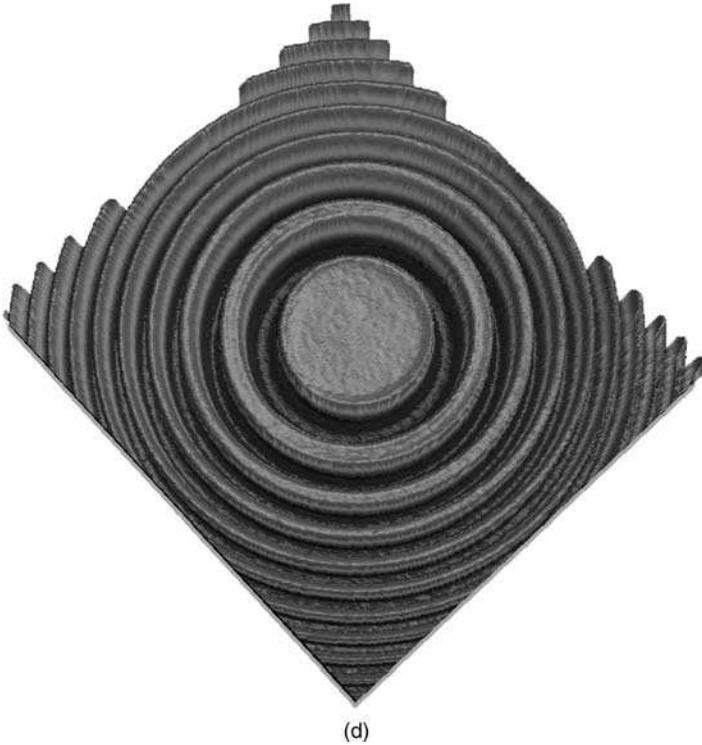


FIGURE 15.5. *(Continued)*

usually scan within a few angstroms of the surface, but the tip can also be in direct contact under forces smaller than a nano-Newton. The first microscope of this type was a scanning tunneling microscope (STM) built by Binnig and Rohrer (Binnig and Rohrer, 1982, 1985) who won the 1986 Nobel Prize in physics for their work in this field. The most popular SPM is the atomic force microscopes (AFM), also known as the scanning force microscopes (SFM).

Numerous modes of operation are possible on the AFM, including magnetic force microscopy (MFM) and electric force microscopy (EFM). The field of scanning probe microscopy is changing rapidly and new probe types are constantly being introduced for a wide variety of applications. With so many different probe types, it is possible to find one that is appropriate for almost any surface to be tested, from integrated circuits to biological objects. Reviews of SPMs have been written by Hansma and Tersoff (1987); Ruger and Hansma (1990), Sarid (1991), Wickramasinghe (1989) and more recently by Bhushan et al. (2004), Bonnell (2001), Cohen and Lightbody (1999), Magnov (1996). This section will concentrate on two types of SPMs: STM and AFM.

Scanning Tunneling Microscopes. In the STM, a metal tip is moved toward the electrically conducting or semi-conducting test surface until a tunneling current

between the tip and test surface is detected. In order to sense tunneling current, voltage must be applied between the probe tip and the test surface. For atomic resolution, the end of the tip has only one atom interacting with the test surface. This tunneling current can only be sensed when the probe is less than 1 nm away from the surface. As the probe is moved closer to the surface, the tunneling current increases exponentially. A change of 0.1 nm causes the current to change by factor of 10 giving STM atomic scale measurement sensitivity. The probe is usually scanned a few angstroms above the surface in a raster fashion using piezo-electric transducers (PZTs). The necessity for a conducting test surface limits the application of the STM, and image resolution is highly dependent on the tip geometry (van Loenen et al., 1990).

Figure 15.6 is a schematic of an STM, consisting of a fine probe tip mounted on x, y, z PZT translators. The STM can operate in either constant-current or constant-height mode. The constant-current mode uses a feedback loop to vary the height of either the probe or the sample during the scan and keep the tunneling current at a constant value. Constant-height mode measures the tunneling current as a function of position while the tip (or the sample) is kept at a constant height. Because larger height variations can be measured, constant-current mode is most often used. Constant-height mode is faster, but the tip can easily crash into the surface if the surface is rough. Crashing not only ruins the tip, but can also harm the test surface.

STMs were originally developed for atomic-resolution applications in a vacuum (Binnig and Rohrer 1982); now they are being used in air and can scan areas larger than 100 μm . Scan range is determined by the PZT; the tradeoff is between range and resolution with longer-range PZTs generally having less resolution. Just as optical microscopes employ various magnifications for measuring with different lateral resolutions and over different sized fields of view, STMs use scan heads with a variety of probe sizes (see Fig. 15.7) and scan ranges in order to change lateral and vertical resolution and scan range. The larger STM scan ranges overlap the measurement span of high-magnification optical profilers and high-resolution stylus profilers. Because STMs require conducting surfaces to produce a tunneling current, surfaces

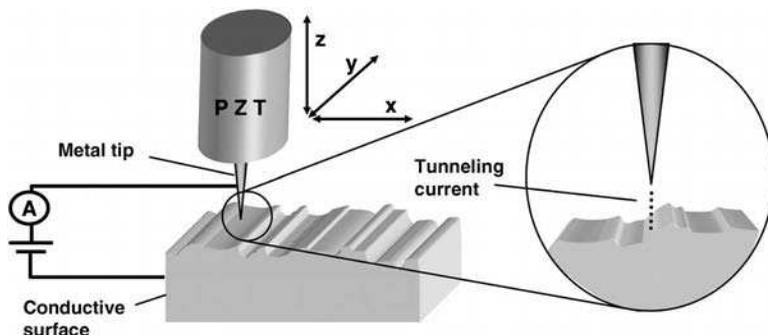


FIGURE 15.6. Schematic of a scanning tunneling microscope (STM), which is able to measure topographies of conductive surfaces with sub-atomic lateral and vertical resolution.

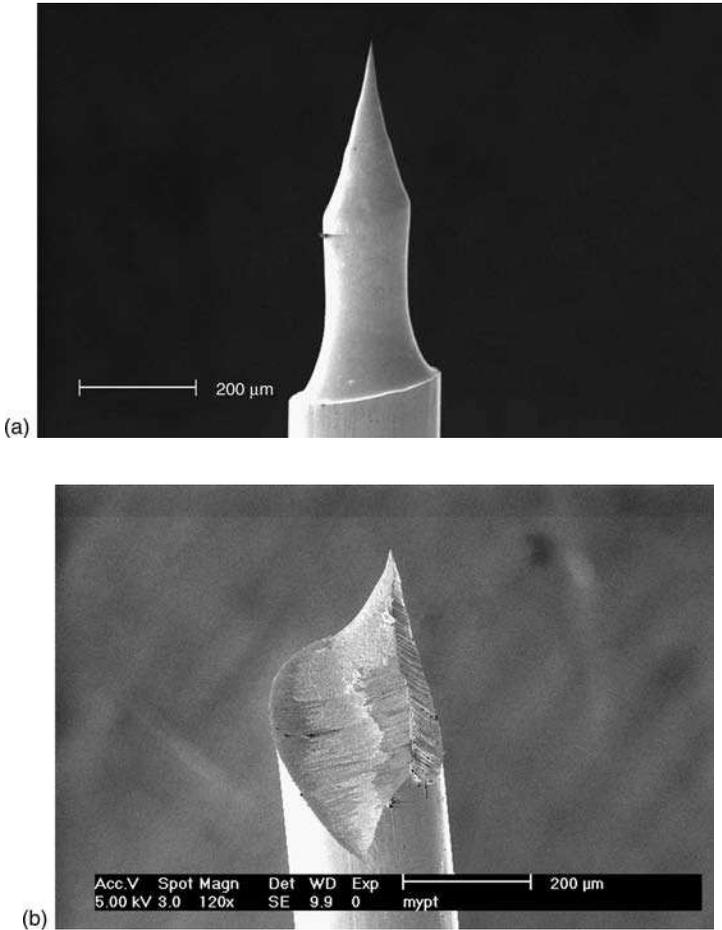


FIGURE 15.7. STM tips can be produced from (a) tungsten wire by computer controlled etching, or (b) tungsten or platinum/iridium wire by mechanical cutting. The latter is recommended for atomic imaging. Courtesy Veeco Instruments.

such as glass cannot be measured. Even with these limitations, STMs have been used to evaluate optical surfaces (Dragoet et al., 1986; Dragoet and Vorburger, 1987; Schneir et al. 1989). Figure 15.8 shows a shaded solid plot of a molecular pattern of a monolayer of liquid crystal film on a conductive graphite-solution interface measured using an STM. The overview of STM's theory, related scanning probe techniques and application were described by Güntherodt and Wiesendanger (1994), Wiesendanger and Güntherodt, (1995, 1996).

Atomic Force Microscopes. The atomic force microscope (AFM) is an extension of the STM (Binnig et al., 1986). In addition to conductive surfaces, the AFM is able to map nonconductive samples by using the force of the tip-surface interaction as the

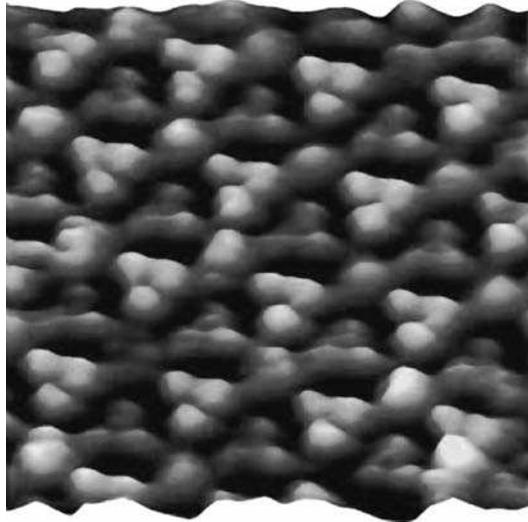


FIGURE 15.8. STM image of a monolayer of liquid crystal with graphite-solution interface. $36 \text{ nm} \times 36 \text{ nm}$ scan. Courtesy Veeco Instruments.

control parameter rather than the tunneling current. Two types of atomic forces can be used for AFMs, a repulsive force and an attractive force; these forces correspond to two basic AFM modes of operation called contact and noncontact modes as shown in Figure 15.9.

In repulsive force mode, the probe acts like a phonograph needle as it is moved essentially in contact with the surface. The tip is typically less than $10 \mu\text{m}$ long with a

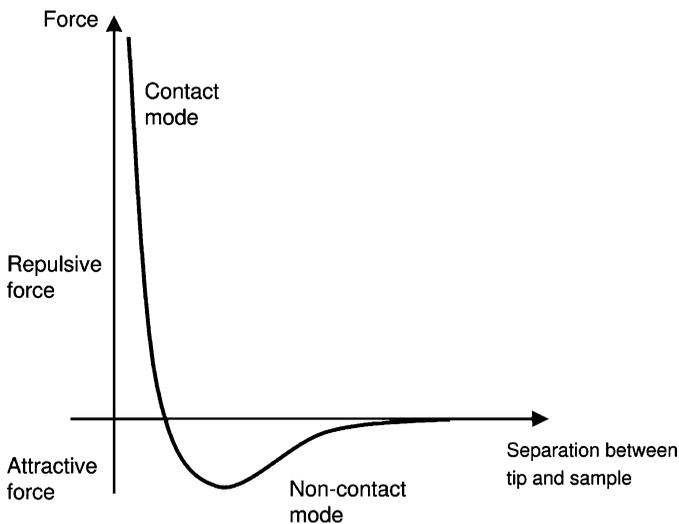


FIGURE 15.9. Force-and-mode type of operation versus tip-to-sample separation.

2–10 nm end radius. The tip is located at the free end of a cantilever (made usually of very flexible material like silicon or silicon nitride) that is typically 100–500 μm long and 2–10 μm thick. A silicon type cantilever and its tip are shown in Figure 15.10. Forces between the tip and the sample surface cause the cantilever to bend (cantilever deflection) as the AFM scans the sample surface. The deflection of the cantilever is measured to determine surface topography. Often, the cantilever is so flexible that the tip-sample contact force is smaller than the force that holds the atoms of many solids together.

Deflection can be measured in a number of ways. Alexander (Alexander et al., 1989) developed a readout system that looks at the deflection of a laser beam

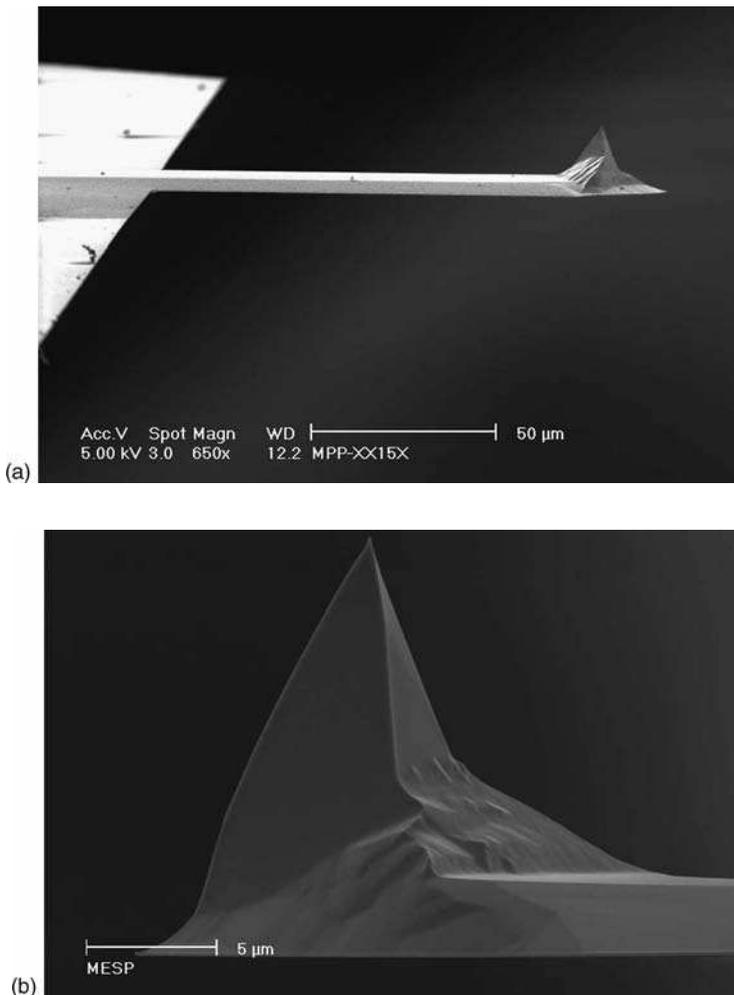


FIGURE 15.10. Scanning electron microscope images of (a) a silicon cantilever and tip and (b) tip only used in AFMs. Courtesy Veeco Instruments.

reflected off a mirror mounted on the cantilever. Deflection measurement is based on the optical triangulation principle. Another readout technique, developed by Sarid (Sarid et al., 1988), uses feedback into a diode laser from the reflection off the back of the cantilever; this cantilever, however, can also be made from piezoresistive material so that its deflection can be measured electrically. Figure 15.11 shows both optical types of readout and Figure 15.12 shows an AFM measurement of the topography of a patterned Si/SiN and metallic glass using laser deflection readout.

Attractive force AFM, also known as non-contact mode, will not damage a sample because it never touches the surface. Because the attractive force is very small, the tip is oscillated at a high frequency, and what is detected is the change in the amplitude and phase of the vibration due to losses or gains in kinetic energy during tip-sample interaction.

A very effective AFM intermittent contact mode called TappingMode (Zhong et al., 1993; Cleveland et al., 1998) has become the most common AFM approach. This patented technique (Virgil and Gurley, 2000) maps topography by lightly tapping the surface with a probe, which is oscillating at a frequency close to the cantilever's resonance frequency. Figure 15.13 shows a schematic of TappingMode. The amplitude of the tip oscillation reaches up to a few tens of nm. This large amplitude of oscillations ensures that the tip does not get stuck in the liquid layer at the surface of the test object, which can happen in noncontact AFM. TappingMode overcomes some of the limitations of both contact and noncontact AFM by eliminating lateral shear forces that can damage soft samples and reduce image resolution. Figure 15.14 shows the measurement of a strand of DNA using TappingMode. TappingMode also allows for the measurement of surfaces with height variations up to several micrometers.

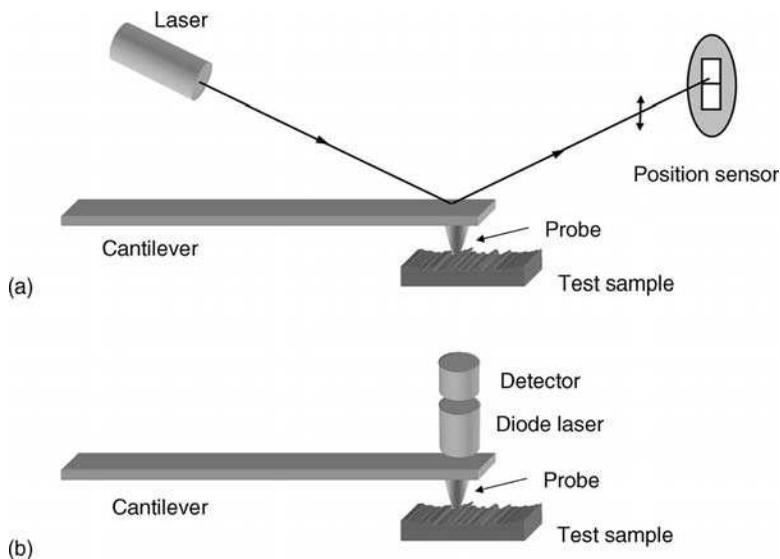


FIGURE 15.11. Schematic of AFM cantilever's tip position readout. (a) Optical lever readout, (b) diode laser readout.

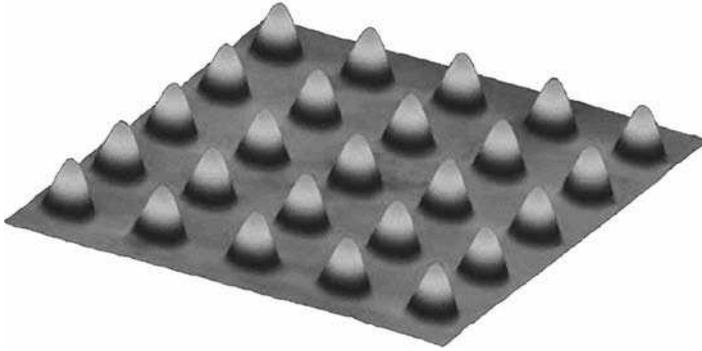


FIGURE 15.12. AFM image obtained in contact mode of a pattern of 80 nm tall features in a Si/SiN substrate produced by ion irradiation through a stencil mask. The features are 1 micrometer in diameter and 1 micrometer apart, $10\ \mu\text{m} \times 10\ \mu\text{m}$ scan. Courtesy Veeco Instruments.

AFM can characterize not only the topography of a sample but also many other sample properties by applying varying motions and signals that drive the probe. For example, variations in material composition, adhesion, friction, viscoelasticity, and electric and magnetic material properties can be determined. Figures 15.15 and 15.16 show the measurement of the surface topographies and the electrical and magnetic properties of samples.

The AFM is also capable of determining changes in the properties of a sample surface by mapping the phase lag between the periodic signal that drives the tip and the oscillations of the tip. A recently developed technique called Torsion Resonance Mode (TRmodeTM) (Su et al. 2003) measures and controls dynamic lateral forces between the probe tip and sample surface. Characterization of torsion oscillations of the cantilever TRmode allows for nanoscale examination of in-plane anisotropy of a sample. A review of different AFM modes and applications of AFM to measurement

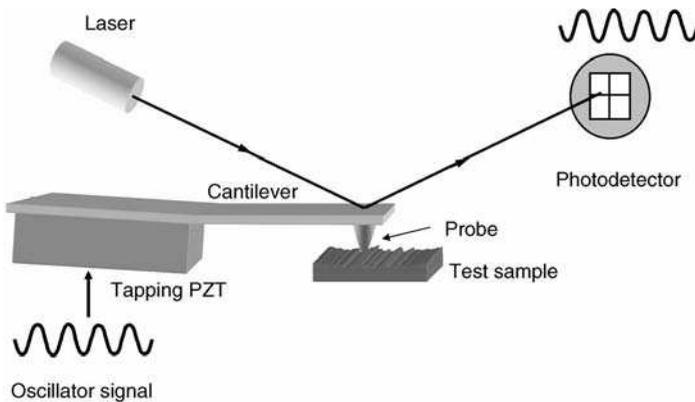


FIGURE 15.13. Schematic of signal detection in AFM Tapping Mode. The probe is kept at a constant level above the sample, which results in a constant amplitude signal. Changes in the amplitude of the signal indicate that the distance between the cantilever and object needs to be changed.

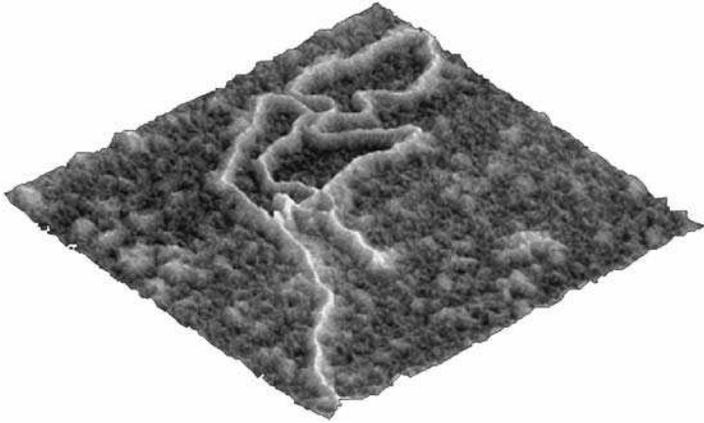


FIGURE 15.14. AFM TappingMode provides a clear and reproducible resolution of lambda phage DNA on mica. This is a $1\ \mu\text{m}$ scan. Courtesy Veeco Instruments.

of microsystems like MEMS was done by Serry and Schmit (Serry and Schmit, 2007).

A number of authors have discussed the uses of the AFM in biological and medical science (Morris et al., 1999; Jena et al., 2002; Braga and Ricci, 2004). The AFM provides important ways for looking at biological samples such as allowing for structural analysis of cells and their functionality. Example of cell measurement is shown in Figure 15.17.

15.2.3. Comparison of AFM and Stylus Profiler

Both the stylus profiler and the later-developed AFM scan across the surface in direct contact with the object at low force. Recent developments in nanotechnology and electronics have pushed the development of the AFM; it is no longer simply a surface

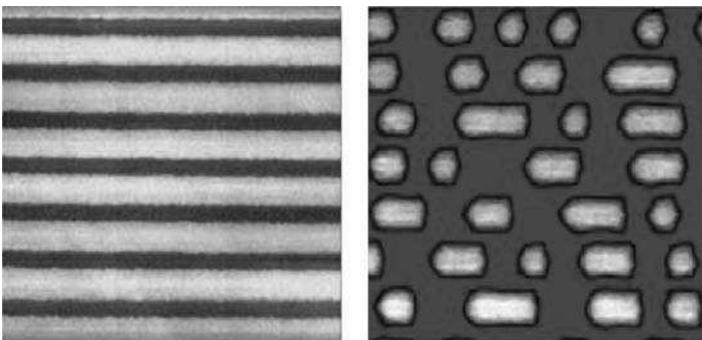


FIGURE 15.15. Surface topography (left) and electrical (right) properties of a DVD-RW. The electrical properties produced using the electron force microscope capabilities of the AFM show amorphous bits formed with the phase change on the crystalline area. Scan area $5\ \mu\text{m} \times 5\ \mu\text{m}$. Courtesy Veeco Instruments.

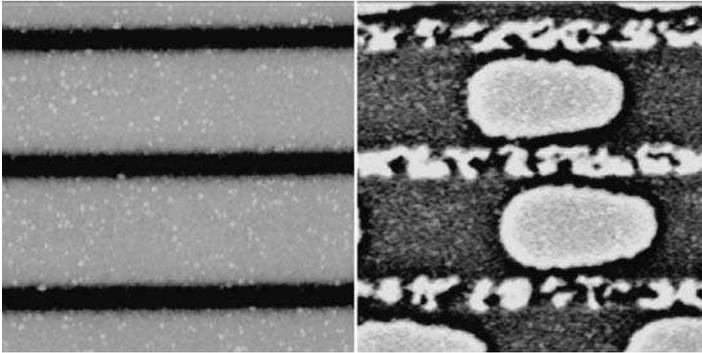


FIGURE 15.16. Bits on a magneto-optical disk. The left image shows surface topography with tracks delineated by grooves. The magnetic force gradient map (right) shows bit edge roughness as well as virgin domain structure in the grooves with features as small as 50 nm. Scan area $5\ \mu\text{m} \times 5\ \mu\text{m}$. Courtesy Veeco Instruments.

topography instrument, for now it can measure a wide range of surface characteristics. The stylus profiler and AFM have complementary capabilities for three-dimensional surface metrology and in fact are sometimes merged together on one platform to fulfill specific industrial needs such as the measurement of semiconductor wafers. Topographies obtained with these instruments are not sensitive to the optical properties of the measured surfaces, and both instruments are well suited for measuring samples made of different materials and films. Their basic characteristics, as compared to the optical profilers described later in this chapter, are collected in Table 15.1. The table

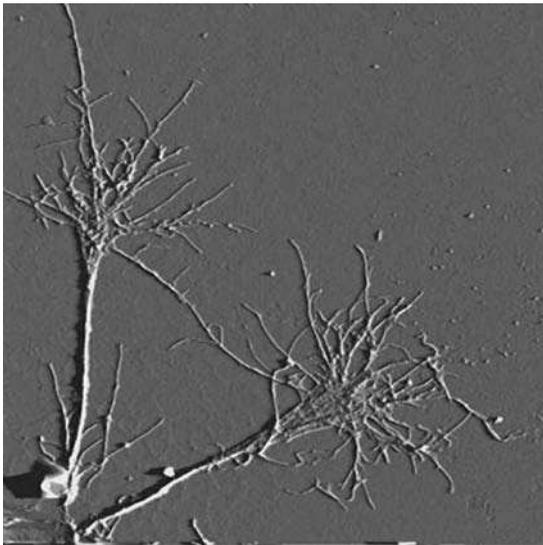


FIGURE 15.17. AFM image of a retinoic acid-induced differentiation of human SH-SY5Y neuroblastoma cells (Dendrites). $100\ \mu\text{m}$ scan. Courtesy Veeco Instruments.

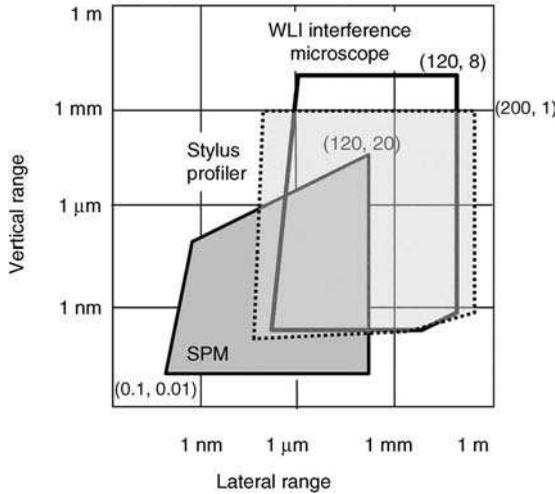


FIGURE 15.18. Plot of measurement ranges for SPM, stylus, and white light interference profilers.

represents typical values at the time of publication. Figure 15.18 shows typical measurement ranges for three types of profilometers.

The broad field of microscopy, including optical, electron and scanning probe microscopy, is reviewed by Hellmuth (2003). A comparison of stylus profiler and AFM measurements on optical surfaces can be found in Bennett et al. (1991).

15.3. OPTICAL PROFILERS

Rather than using a mechanical probe to map surface topography, an optical probe that does not contact the test surface may be used. Optical probes determine shape by sensing the best focus position on a test object. Optical profilers generate measurements by sensing focus at a single point on the surface and adjusting the height of the focusing lens until focus is achieved. Alternately, the signal may be collected during a scan through focus and then analyzed. The amount the lens is moved indicates the surface height at that data point. Either the optical head or the surface is scanned to generate a two- or three-dimensional height-profile map of the test surface. Special hardware is needed to sense the focus. A different group of methods for finding a surface profile that use a conventional microscope or a stereomicroscope is based purely on image processing of the collected images as the object is scanned through focus so as to find the best focus at each point.

The signal collected in optical profilers is often incorrectly called intensity; Palmer (Palmer, 1993) provides a clear discussion of the terms intensity versus irradiance. Intensity describes radiation emanating from source (watts/steradian) while irradiance describes signal collected by detector and is expressed in watts/area. We will use the term irradiance to describe the radiometric quantity detected by the detector.

TABLE 15.1. The main characteristic parameters of a scanning probe microscope, stylus profiler, white light interferometric optical profilometers, and confocal microscope.

	Atomic force/ scanning tunneling microscope	Stylus profiler	White light interferometric optical profiler	Confocal microscope
X,Y resolution	2–10 nm AFM 0.1 nm STM	50 nm (stylus radius dependent)	0.5 μm (NA objective dependent)	0.5 μm (NA objective and lateral sampling dependent)
Z resolution	0.1 nm AFM 0.01 nm STM	0.25 nm	0.3 nm	1–20 nm (Dependent on objective magnification)
Field of view	Typically up to 120 \times 120 mm	Up to 200 mm	100 \times 100 μm to 10 \times 10 mm but can be extended by stitching	100 \times 100 μm to 10 \times 10 mm but can be extended by increased lateral sampling lens array objective
Measurable height range	Up to 20 μm	1 mm	8 mm (or limited by working distance of objective)	Limited by working distance of objective
Sample preparation	Little or none	None	None	None
Contact technique	Optional	Yes	No	No
Special surface requirements	STM - only conductive surfaces	Surface damage possible due to high forces	Needs correction for dissimilar materials and film coatings	Needs correction for dissimilar materials and film coatings
Scanning	Point by point	Point by point	Full field of view	Point by point
Full field dynamic motion of sample measurement	No	No	Yes	No

(Continued)

TABLE 15.1. (Continued)

	Atomic force/ scanning tunneling microscope	Stylus profiler	White light interferometric optical profiler	Confocal microscope
Through the glass measurement	Not possible	Not possible	Possible	Possible
Film thickness	Only if film has a step	Only if film has a step	Minimum 0.1 μm	Minimum 1 μm
Measurable optical properties of surface or film	Indirectly, through correlation with topography for example	no	yes	yes
Other measurable material properties	Numerous: adhesive, electric, magnetic, visco-elastic, elastic, . . .	no	no	no

15.3.1. Optical Focus Sensors

A simple method of determining focus has been implemented in profiling instruments developed by Brodman and Smilga (1987) and Breitmeier and Ahlers (1987). Illumination from a laser source is focused on the test surface, and the return is split at the optical axis into two parts using a prism. Each half of the beam is incident upon a split detector and the difference signal from each split detector is monitored. When the focusing lens is too high, the return beam focuses in front of the split detectors and causes a larger signal on the inner detectors; when the lens is too low, the larger signal is on the outer detectors. The sign of the difference in signal will determine which side of focus the test surface is on, and is used to generate a focus error signal which moves the focusing lens to the correct position. When the focusing lens is in the correct position, both the inner and outer detectors have equal signal and the difference signal is zero. Two sets of split detectors are used to account for variations in tilt of the test surface. A sensor of this type is shown schematically in Figure 15.19 (Brodman and Smilga 1987). Because the focus must be adjusted to null the signal at every sampled surface point, this type of profiler can take a few minutes to generate a three-dimensional surface profiler.

The lateral resolution of optical focus sensors is limited by the size of the focus spot at the test surface, usually 1.0 to 1.5 μm in diameter. The measured surface height at a given sample point will be the average height of the surface over the spot size. This means that the smallest measurable features are about 2 μm . The measurement area will depend upon the sampling interval and number of data points. Another

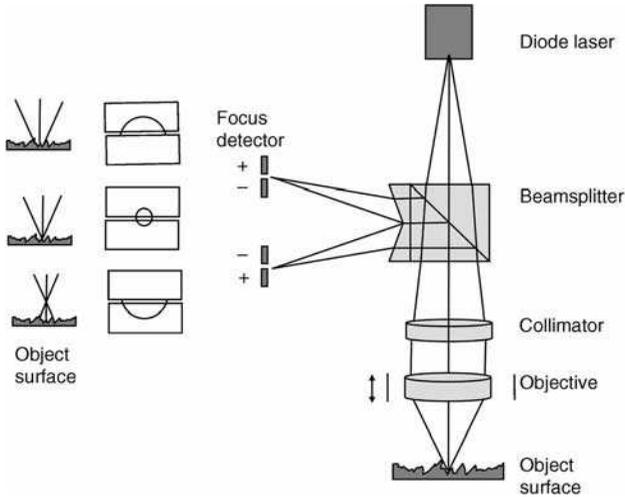


FIGURE 15.19. Schematics of a profilometer with an optical focus sensor.

limitation of this type of profiler is that the light reflected from the test surface must get back into the sensor. If there are steep slopes on the surface, the light may get scattered out of the instrument and the signal will be lost, causing inaccurate results when only the difference signal is monitored. The height resolution of this type of profiler is related to the focusing range and the time to obtain each point. If a large height range is being measured, the movement of the focusing lens needs to be coarser to keep the time per data point the same. Otherwise, finer focusing over larger height ranges will slow the measurement time considerably. Calibration of optical focus sensors is similar to that of stylus and SPM profilers. A traceable standard of approximately the same height or roughness as the test surface is measured and a scale factor is determined to apply to the surface profile data.

Optical focusing techniques have been implemented in many fields, from high density storage applications to surveillance camera systems. Different optical setups can also be used to measure the focus position. Astigmatic lenses (Cohen et al., 1984) and twin micro-Fresnel lenses (Shiono and Setsune, 1989) are two examples. Some profilometric techniques are based on focusing mechanisms developed for other applications like a CD player (Zhang and Cai, 1997; Bartoli et al., 2001). Two most common optical profilers based on a focusing principle are the confocal microscope and the white light interferometer. These will be described in Sections 15.3.2 and 15.5.

Conventional and stereoscopic microscopes without any special hardware can also map surface topographies after doing post-processing of the collected images, not single points, as the microscope scans through focus. Two of these methods are digital deconvolution and stereoscopic imaging.

Digital Deconvolution of Conventional Images. Images registered using a conventional microscope as it scans through focus are deconvolved using a theoretical or

measured point spread function or through blind deconvolution. This method is used in medicine for reconstruction of 3D images from CT and MRI scans. Other methods are based on the measurement of local sharpness of the image (Nahm et al., 1998). These methods work best for objects with distinct lateral features, but these methods have difficulty resolving objects with smooth surfaces that have no lateral features.

Stereoscopic Imaging. Images registered using a stereoscopic microscope are analyzed by identifying the same features in both images and measuring the distance between features. Knowing the angles through which the two stereoscopic images are observed allows for the determination of the relative axial distances between different features on the object. Like deconvolution, this method works best for objects with distinct lateral features, but does not work well for objects with smooth surfaces or with periodic structures. Stereoscopic imaging can be used to image both large and small objects as long as the images are obtained from two different perspectives.

15.3.2. Confocal Microscopy

Like many interferometric methods confocal microscopy really took off with the development of image processing software and affordable, powerful lasers and computers. Back in 1961, Minsky patented modifications to a biological microscope that reduced the stray light in the system in order to improve the obtained image; however, only in the mid 1980s did the development of confocal microscopes to obtain non-invasive, three-dimensional data from biological specimens occur. Since this time, confocal systems have been an important measurement tool in the fields of cell biology, physiology, cytogenetics and developmental biology (Pawley, 1995; Gu, 1996) and ophthalmology. The introduction of two-photon microscopy (Diaspro, 2002) in the late 1990s offered reduced photodamage and increased tissue penetration for better imaging. In recent years confocal microscopes have been adapted for the measurement of microsystems and material applications (Aguilar and Mendez, 1995; Schneider et al., 1997; Smith et al., 2000; Tiziani et al., 2000).

Confocal microscopy sets itself apart from standard light microscopy through the use of confocal (pinhole) apertures that ensure only light at the point of focus on the test surface enters the detector. This elimination of out-of-focus and stray light and the resulting high resolution and high signal-to-noise images are the main advantages that recommend confocal microscopy. The classic setup of a confocal microscope places two small apertures in planes conjugate to the focal plane of the objective, one in front of the illuminating source and the other in front of the detector as shown in Figure 15.20.

In some systems, a spatially coherent light source, namely a laser, eliminates the need for a pinhole at the illumination source; however, in these systems speckle may become a problem. Because confocal microscopy is in principle a single point method, synchronized lateral scanning of the illumination and detection points is required.

Image Build Up. Point-by-point scanning of the illuminating and detecting pinholes typically in *X-Y* raster fashion results in a 2D irradiance image (single optical section) of a sample at a given focus plane. To build up a 3D irradiance image

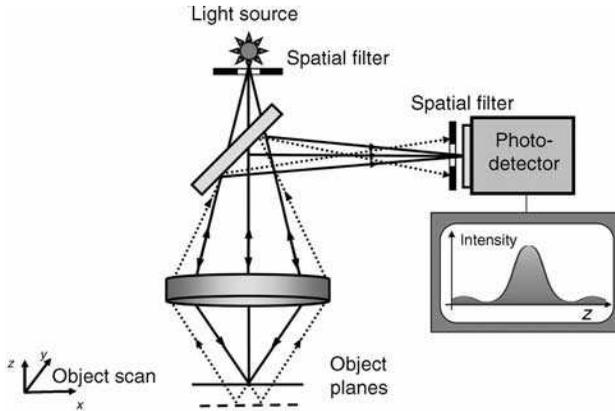


FIGURE 15.20. Depth discrimination in reflective confocal microscopy by eliminating out of focus light via confocal (pinholes) apertures.

(multiple optical sections along the Z direction) either the sample or the objective lens or detector is scanned vertically such that each point on the sample surface passes through the focal plane of the microscope.

For each point on the sample passing through the focal plane of the microscope the collected irradiance (confocal) signal, which is also called the axial point spread function (PSF), falls off as the distance from focal plane increases. The width of this axial point spread function for a point object for aberration-free objectives with $NA < 0.5$ can be described by the simplified expression (Corle et al., 1986; Kino et al., 1988; Ho and Shao, 1991; Corle and Kino, 1996)

$$FWHM = \frac{0.9\lambda}{NA^2} \quad (15.2)$$

The width of the axial PSF depends on the numerical aperture (NA) of the objective and the wavelength. In reality the width of the axial response also depends on the tilt of the object and its shape, namely whether the object is a point or plane-like. The width of the spatial filter and the size of the detector also influence the axial response. Sheppard (Sheppard, 2003) provides a review of many issues in imaging confocal systems.

The signal collected at a single point on the object during a vertical scan is evaluated for maximum irradiance; maximum irradiance corresponds to the imaged point being in focus. In this way a confocal system for surface profiling works like a focal point detection system scanned over the test surface. The position of the axial point spread function can be more consistently determined if, for example, the center of gravity of the signal or the position of the polynomial fitted to the signal is determined instead of simply the position of maximum irradiance. The width of axial point spread function and the sampling rate also influence the consistency in determining the signal position and thus the vertical resolution. Better vertical

resolution is achieved when the axial signal is narrower; because of this the best vertical resolution depends on NA and thus magnification of the objective. Vertical resolution for surface profiling is often defined as RMS (root mean square) of the difference measurement. The vertical resolution also determines the minimum step height that can be measured on a sample. For smooth surfaces the vertical resolution can be on the order of a few nm, but only for the highest magnification objectives. For lower magnification objectives the vertical resolution can be around 10–15 nm to microns. This vertical resolution often is called method's sensitivity.

Confocal microscopy is often used to measure the thickness of transparent layers in the process called optical slicing. Light penetrates the object, is reflected from the interfaces of the layers, and creates additional axial responses. Optical sectioning of confocal systems is commonly used to measure, for example, cells in biological applications and transparent coatings in the semiconductor industry. In cases like these, a different definition of vertical resolution is used, a definition that applies to the measurable optical thickness (Sheppard and Gu, 1992; Sheppard et al., 1994). This definition of axial vertical resolution is based on the vertical two-point resolution of the signal where the width of the axial signal determines the resolution, in this case the minimal measurable optical thickness. The vertical resolution of optical slicing can reach down to 1.5 μm for high NA objectives and is much larger than the vertical resolution related to the measurable step height. Figure 15.21 shows example of axial responses for transparent layer.

Much research has gone into improving vertical resolution mainly through the application of annular filters in the pupil (Sheppard and Gu, 1991; Martinez-Corral et al., 1995) to narrow the width of PSF in both directions as registered by a finite size detector. In order to enhance optical-sectioning capacity, some methods apply symmetrical defocusing of the point source and point (Sheppard and Hamilton, 1984; Ho and Shao, 1991; Kimura and Wilson, 1993). Axial apodization was

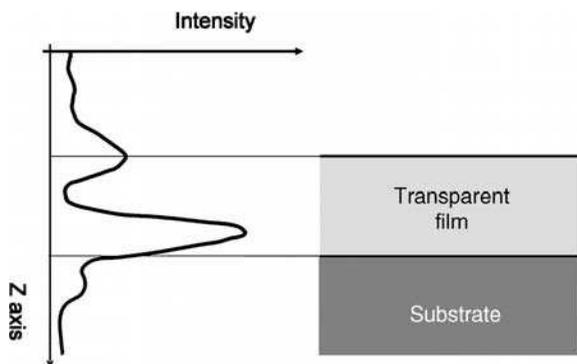


FIGURE 15.21. Schematic of axial response for transparent films in a confocal microscope. The position of the peak corresponding to the interface depends on the index of refraction of the material and appears to be at a shallower depth than the position of the interface, which reduces the capabilities of measuring thinner films. The width of the axial response depends on spherical and axial aberrations introduced by the film. The amplitude of the peaks depends on the ratio of the indices of refraction at the interfaces and thickness and dispersion of film.

proposed so as to improve vertical resolution by the application of destructive-interference apodizers, which provide an axial response with zero irradiance at the focal point (Martinez-Corral et al., 1998). Some have suggested using a phase conjugate mirror to improve vertical resolution (Uhlendorf et al., 1999). However, the one, practical solution for increased resolution still has yet to be developed.

Confocal System Modifications. Over the years a number of methods have been worked out to improve the speed of data acquisition; most employ variations on the confocal aperture. In 1967, Egger and Petran (and later, Petran et al., 1968; Xiao et al., 1988) introduced simultaneous illumination of the sample by many spots of light using an array of pinholes on a rotating disk. This Nipkow disk, invented by Paul Nipkow in 1884, produces images by rotating a disk with multiple pinhole apertures in front of the extended source. A second disk with a matching array of pinholes is placed in front of the detector; however, this method has a significant downside in that it reduces the amount of light at each sample point.

In 1996, Ichihara et al. showed that the amount of light in the system can be increased by up to 10 times placing by a disk with microlenses in front of the pinholes at the disk that is used at the illuminating beam. Pinholes may have different patterns and there may be even up to 20,000 pinholes on the disk with about 1000 of them illuminating the test object at a time. Speed measurements up to one frame/ms may be achieved using this method (Tanaami et al., 2002). Tiziani et al. (2000) suggested a simpler system where pinholes in the Nipkow disk are replaced by the microlenses itself, although this system has a lower lateral resolution. Such system is shown in Figure 15.22.

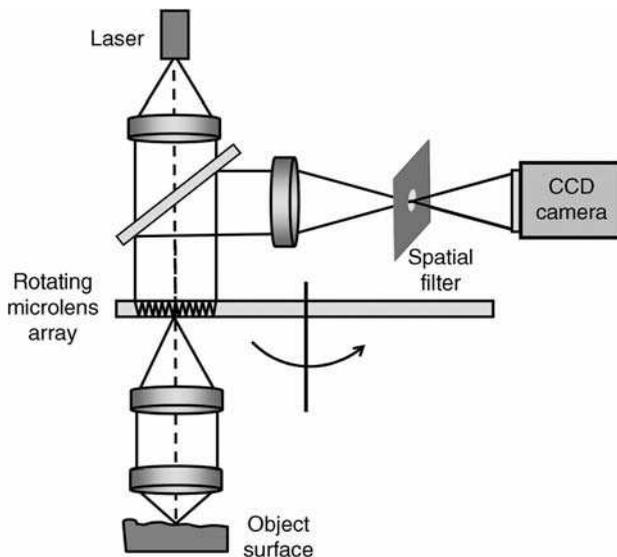


FIGURE 15.22. Confocal microscope with microlens array on Nipkow disk.

Some systems use a scanning slit (Wilson, 1990) instead of a pinhole. A slit aperture represents an alternate geometry to the array of pinholes. In the slit-scan system, multiple detectors are required along the length of the slit, and lateral scanning is necessary only in the direction perpendicular to the slit, which increases the speed of image acquisition (Neil et al., 1997). However, with slit scanning, the width of axial responses is wider than for single point scanning systems, and the vertical resolution is reduced for the given numerical aperture objective.

Yet another method to speed up measurement time is to simultaneously project an equally spaced multiple slit or grid pattern. For these kinds of projected patterns, however, the width of the axial signal with defocus will depend not only on the NA of the objective but also on the spatial frequency of imaged pattern. The fundamental difficulty with this method is the residual unwanted grid pattern in the image. The grid pattern can be removed if the projected pattern is moved in a simple saw-tooth fashion synchronized to the camera frame rate such that any three successive camera images corresponded to a spatial shift of one third of a period in the position of the projected image of the grid. The grid pattern image, because of its periodicity, can be processed also in the Fourier domain.

In the first systems that projected a grid pattern, this pattern was imprinted on a glass plate and laterally shifted by a PZT shifter. Nowadays, the illumination beam can be reflected off a digital micromirror array (DMD made by Texas Instruments) that configures the grid pattern (Hanley et al., 2000), or a polarized illumination beam can be reflected off of a ferroelectric liquid crystal (FLC) (Smith et al., 2000) that configures the pattern (slits or single points) projected onto the sample surface. Each system setup has to match the microdisplay lines and the rows of the CCD pixels. The lateral scanning of the pattern is done by displaying a series of patterns on a microdisplay without the use of a PZT.

If good vertical resolution is a requirement from the confocal profilometer, a high numerical objective needs to be used. The downside of the high NA objective is its small field size. Tiziani and Uhde (1994b) proposed the use of a microlens array in place of the objective as shown in Figure 15.23. This array would allow for the measurement of larger fields; the fields would be determined by the size of the array while each individual microlens would still maintain a high NA. The focal lengths of the individual lenses could be adjusted to the shape of the object in order to decrease the scanning range and further speed up the measurement (Tiziani et al., 2000). The system with microlenses in place of the objectives is a bit different than the typical confocal microscope because the light reflected from each object point while being in the focal plane of that microlens is focused by the lens onto a pinhole, which plays the role of a spatial filter. The pupils of the microlenses are imaged onto the camera rather than the image point as in typical confocal setups. For this system a vertical resolution of 50 nm was achieved for an objective with an $NA = 0.3$.

Chromatic Confocal Microscopy. The chromatic confocal microscope was developed to circumvent the necessity in confocal systems of the vertical scan in order to ascertain relative object height position. Rather than a vertical scan, a chromatic confocal microscope employs an objective with a longitudinal chromatic aberration

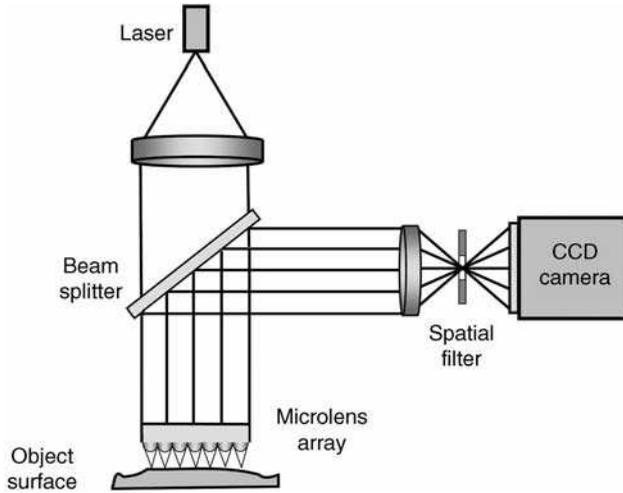


FIGURE 15.23. Wide field confocal microscope with microlens array in place of the objective.

that creates a different focus position for each different wavelength. This idea was first devised by Molesini in 1984 (Molesini et al., 1984) and then adapted to confocal microscopes by Browne et al in 1992; then, others followed (Tiziani and Uhde, 1994a; Dobson et al., 1997; Cha et al., 2000; Lin et al., 1998; Ruprecht et al., 2004). Only the wavelength, whose focus position coincides with object position, is reflected back to the system. For this reason these systems are also called wavelength-to-depth encoding setups. A spectrometer, in place of the CCD camera, detects the wavelength value. Instantaneous measurement of the object focus position by measuring the power spectrum replaces the need for any scanning mechanism and makes the measurement process much faster. Schematic of chromatic confocal microscope is shown in Figure 15.24.

These systems usually use broadband (i.e., white light) sources and a spectrometer, but some systems employ a wavelength tunable source and a CCD camera (Mehta et al., 2002) or white light source and color CCD camera (Tiziani et al., 2000). The wavelength for which the maximum power is detected is encoded to the depth of the object. A drawback to this technique is that the measurable maximum depth depends on the chromatic aberration of the lens and the source spectrum. The diffractive lenses that are used in confocal systems (Dobson et al., 1997) can provide stronger longitudinal dispersion, which is material independent and can be characterized analytically in contrast to refractive lenses. The focal length of the diffractive lens for a given wavelength λ follows the linear dispersion

$$f(\lambda) \approx 2f(\lambda_d) - f(\lambda_d)/\lambda_d \quad (15.3)$$

if $\lambda - \lambda_d$ is much less than the design wavelength λ_d . For lower NA this range can be even 2–3 mm, and for higher NA (i.e., 0.75) this range can be of the order of 10 μm .

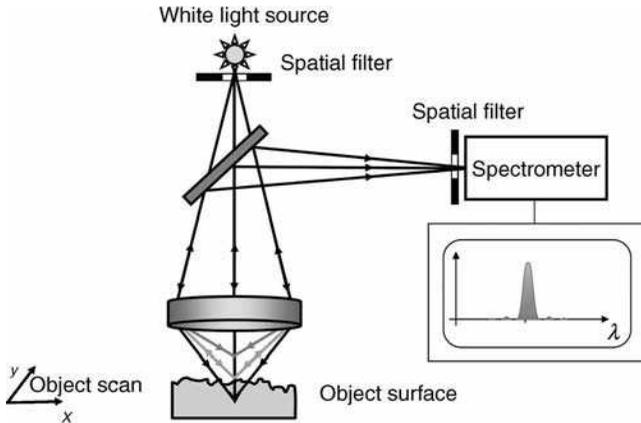


FIGURE 15.24. Schematic of chromatic confocal microscope.

Shi (2004) proposed the use of supercontinuum light which has a high spatial coherence and very wide spectrum, which results in wider measurable object depth and no speckle noise. Supercontinuum light may be a solution to the light efficiency problem in confocal systems. Light efficiency may be also improved by using an extended detection pinhole (Ruprecht, 2004).

Similarly as in a conventional confocal system in order to avoid mechanical lateral scanning of the sample or objective, slit (Lin et al., 1998), microlens (also diffractive) array (Tiziani et al., 2000) can be applied, or a dynamically configurable micromirror like a digital micromirror device (DMD) (Cha et al., 2000), and a liquid crystal display can be used, which serve as a scanning point source (pattern) and detection pinhole.

15.4. INTERFEROMETRIC OPTICAL PROFILERS

Interferometric optical profilers are based on standard microscopes where an interferometer built into the objective replaces the standard objective. The interference signal obtained with these objectives is analyzed to provide quantitative data about an object being measured. Interferometric optical profilers generally record a number of frames of data in order to calculate surface heights at each detector point. During measurement the interference signal is varied by changing the optical path between the object and reference beams; this varying is usually accomplished using phase-shifting or vertical scanning techniques that move the objective relative to the test surface using a piezoelectric transducer or motor scanner. Some methods avoid mechanical scanning by using the spectral properties of the light source; wavelength scanning and wavelength dispersing are two such methods. Which method used for a particular analysis depends on the type of object to be measured.

15.4.1. Common Features

The four interferometric objective setups that are typically used are based on the Michelson, Mirau, Linnik, and Fizeau interferometers. Schematics of these interferometric objectives are shown in Figure 15.25. A number of factors help determine the choice of best objective for a particular measurement. These factors include the level of magnification or more precisely the numerical aperture required to both resolve the features and measure slopes and heights on the sample. At the same time a magnification level that measures the entire area of interest on a sample needs to be chosen. This section first describes various types of interferometric objectives and then examines the characteristics of different objective setups that determine which is best suited to a particular application.

Types of Interferometric Objectives. The design of the interferometric objective is constrained by the mechanical limitations of the system. The four typical objectives differ in the ways that the beam is split into reference and object beams. In all but the Fizeau objective, the reference mirror is placed at the best focus of the objective to obtain best contrast fringes when the sample is in focus. In order to obtain best focus position, the object first needs to be placed in focus, and then the reference mirror is moved to a position for which the best contrast fringes are obtained; this position also corresponds to best focus position. The best contrast fringes correspond to the zero optical path difference (OPD) between the reference mirror and the object position in the arms of the interferometer.

Michelson interferometers (Fig.15.25(a)) are comprised of an objective, a beamsplitter and a separate reference surface. The microscope objective must have a long

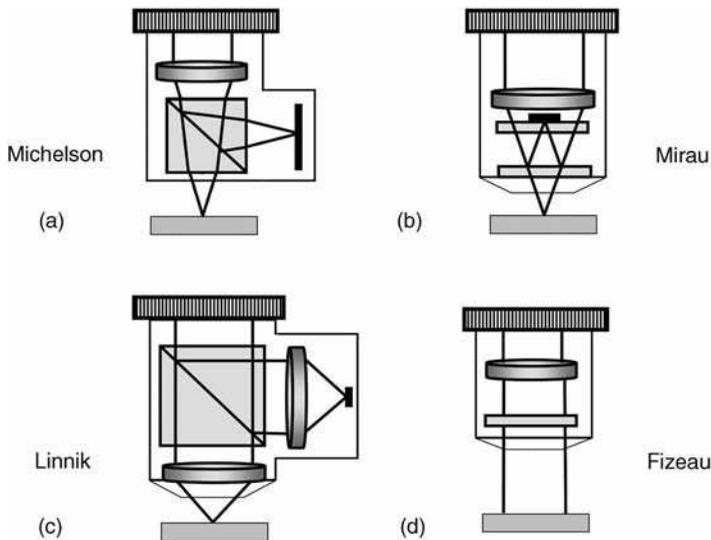


FIGURE 15.25. Schematics of interferometric objectives: (a) Michelson, (b) Mirau, (c) Linnik, and (d) Fizeau.

working distance to fit the beamsplitter in between the objective and the surface. Because of this, Michelson interferometers are only used with low-magnification objectives having low numerical apertures and long working distances.

Mirau interferometers (Fig. 15.25(b)) contain two small glass plates between the objective and the test surface. One plate contains a small reflective spot that acts as the reference surface, and the other plate is coated on one side to act as a beamsplitter. The plate with the reference spot also acts as a compensating plate. These interferometers are used for midrange magnification objectives where not enough space exists to insert a beamsplitter cube so as to create a Michelson interferometer. Mirau setups are not very useful at magnifications of less than about 10X because at these lower magnifications the reference spot obscures too much of the aperture. The reference spot has to be larger than the field of view of the objective since it is a surface conjugate to the best focus plane of the object. For magnifications above 50X, the working distance of the objective is too short to place either a beamsplitter cube or plates underneath it.

The Linnik system (Fig. 15.25(c)) allows an interferometer to be set up for any magnification objective from two identical bright field objectives. However, these interferometers are also very difficult to adjust, and thus their commercial use is rather limited. Linnik systems are often used with a high-magnification objective that has a short working distance. A Linnik setup is comprised of a beamsplitter, two matched microscope objectives and a reference mirror. The entire reference arm provides path-length matching in order to obtain interference fringes. The two objectives need to be matched with a beamsplitter to provide a wavefront with minimum aberration and maximum fringe contrast.

The three objectives discussed above are all variations of the Michelson interferometer and all are equal-path interferometers.

The Fizeau interferometer (Fig. 15.25(d)) is an unequal-path interferometer that requires a source with a long coherence length; this setup is well suited for monochromatic or spectral interferometry, but not for white light interferometry. The objective provides a collimated beam on the test surface while imaging the test surface. Since interference fringes will be visible over a large depth of field when a long coherence length source is used, care must be taken to focus on the test surface. Spurious fringes created by Fizeau cavities have to be watched for while analyzing fringes. Fizeau objectives can be used with a large range of magnifications; however, there must be enough room to place the reference surface between the objective and the test surface.

With all of the interferometers except the Fizeau, there is a cone of light incident upon the test surface, and all of these objectives except the Fizeau can be adjusted to obtain white-light fringes.

Depending on the type of object to be measured different illumination schemes may need to be implemented. For measurements of optically smooth surfaces typically instruments based on phase detection in monochromatic illumination are used. The analysis of monochromatic fringes is described in the chapter on Phase Shifting Interferometry. For objects with larger discontinuities, roughness and heights, alternate illumination schemes and fringe analyses need to be applied. Techniques that extend

the measurement range of monochromatic illumination systems form the basis of this chapter, but first we describe the properties of the objectives and the features of a sample that can be measured with a given objective setup.

Matching Interferometric Objectives with Object Parameters. When using an optical profiler the choice of interferometric objective for a particular measurement first requires determining a number of parameters about the sample and the features to be resolved. Decisions need to be made about the smallest lateral features, the maximum height range and maximum slopes to be measured as well as the area to be measured. While magnification of the objective (plus intermediate optics) and size of CCD camera determine the measurable object area, the NA of the objective and the type of illumination determine which features on the object can be measured. Some of properties of the objectives discussed below can also be applied to the objectives used in confocal profilometry.

Lateral Resolution. In order to determine which lateral feature on a sample can be resolved, the lateral resolution of the system first needs to be known. The lateral resolution of a system that is based on a microscope depends mainly on the optical resolution of objective used. Only systems with low magnification objectives may be limited by detector sampling. Optical resolution depends on the wavelength and the numerical aperture of the microscope objective. Sparrow and Rayleigh cite slightly different criteria (see Born and Wolf (1999)).

$$\text{Sparrow optical resolution criteria} = \frac{0.5\lambda}{\text{NA}} \quad (15.4)$$

$$\text{Rayleigh optical resolution criteria} = \frac{0.6\lambda}{\text{NA}} \quad (15.5)$$

These criteria apply for an incoherent system imaging two radiating points as an object, and they are a good general rule of thumb. However, a microscope is (at best) a partially coherent system and objects may take different forms. Modified criteria do exist that consider the coherence state of the system and the type of test sample, but in practice the Sparrow criterion is a good approximation for an optical microscope. It is worth remembering that registering an image with a CCD camera requires at least three pixels in order to resolve an image of two points. Optical resolution and lateral sampling for a typical interference microscope are shown in Table 15.2.

Measurable Heights Range. A test surface can be measured only when fringes are visible on its surface. When monochromatic, multiple wavelength or spectral techniques are used where the object practically stays at one focus position, fringes have to be visible over the whole height range of the object. This height range of the test surface must be within the depth of field of the interferometric objective. In addition, when the illuminating source has some wavelength bandwidth $\Delta\lambda$, then the test surface must also be within the coherence length of the source.

TABLE 15.2. Characteristics of objectives for following assumptions: wavelength 600 nm, Sparrow criterion for optical resolution calculations, sampling interval of camera 6.8 μm , and number of camera pixels in one direction, 1024.

Interferometer type	Michelson		Mirau		Linnik	
	2.5	5	10	20	50	100
Magnification (X)						
Numerical aperture	0.075	0.13	0.3	0.4	0.55	0.9
Optical resolution (μm)	4.00	2.31	1.00	0.75	0.55	0.33
Depth of field (μm)	106	35	6	3.5	1.6	0.35
Lateral sampling interval (μm)	2.72	1.36	0.68	0.34	0.17	0.068
Field of view area (μm)	2785	1393	696	348	174	70

The depth of field of the objective depends upon the NA and the wavelength λ of illumination; it is defined by (see Benford (1966))

$$\text{Depth of field} = \lambda \frac{\sqrt{1 - (\text{NA})^2}}{(\text{NA})^2} \quad (15.6)$$

The coherence length (this is an approximation) of the source can be determined from central wavelength λ of the source and its bandwidth $\Delta\lambda$:

$$\text{Coherence length} = \frac{\lambda^2}{\Delta\lambda} \quad (15.7)$$

The measurable height of the object is determined by the smaller of these two numbers, namely the coherence length of the source and the depth of field of the interferometric objective. Other criteria, which are dependent on the measurement method used, may further limit the measurable height range.

The measurable height range for methods in which the object is scanned through focus such as with confocal or white light interference profilers is limited by the scanning range and the working distance of the objective.

Measurable Slopes. The maximum of the measurable slope of the sample also depends on the NA of the objective. If light reflected from a sample with a high slope is not gathered by the objective, the slope can not be measured. A common assumption holds that slopes with a maximum 0.75NA return sufficient light to be measurable as long as fringes with sufficient contrast are created. For objects with rough surfaces higher slopes can be measured since diffuse surfaces allow some light to travel back to the objective.

Table 15.2 gives the optical resolution predicted by the Sparrow criterion for 600 nm wavelength of the source along with the depth of a single image at the object plane (the coherence length or depth of field, whichever is smaller), lateral sampling interval on the test surface, and field of view across the test surface. In summary,

objectives of lower numerical aperture can measure samples of larger height ranges and larger field of views, but they cannot resolve high slopes and small lateral features. Good knowledge of object features to be measured, on the contrary of objectives and measurement techniques capabilities, is needed in order to obtain required measurement result.

Interference Fringes. All of these interference objectives create an interference pattern that can be observed by the CCD camera. We will now consider the interference signal as registered by a single point x, y ; however, for simplicity these coordinates will be omitted in all equations. For a single wavelength and a single point source on axis (a spatially and temporally coherent source) and thus a single angle of beam incidence, the interference signal can in its most general form be described as

$$I(k, z, \theta) = I_R(k) + I_O(k) + 2\sqrt{I_R(k)I_O(k)} \cos(2k(h - z) \cos(\theta) + \phi(k)) \quad (15.8)$$

where $k = 2\pi/\lambda$ is the wave number for a source wavelength λ , and I_R and I_O are the detected irradiances reflected from the reference mirror and the object, respectively. I_R and I_O depend on the reflectivity of the object and the reference mirror, the transmissivity of the optical system, camera sensitivity and the spectrum of the source.

Phase Difference. The phase difference between interfering beams is described as the optical phase under the cosine from Eq. (15.8) in the following form:

$$\varphi = (2k(h - z) \cos(\theta) + \phi(k)) \quad (15.9)$$

where $h - z$ is the geometrical path difference between a point on the object and a corresponding point on the reference mirror, h represents object height, and $\cos \theta$ is the direction cosine of the beam's incident angle onto the object. The remaining phase term $\phi(k)$ represents the phase change on reflection introduced by the material of the object; the phase term may also contain both the statistical phase term introduced by speckles and the phase offset due to the dispersion of the instrument, which typically is assumed to be zero.

When $h - z$ equals zero, the object and reference beam are traveling the same total optical path length. Thus, this point corresponds to the zero optical path difference (zero OPD) between the beams. The objective is typically set so that the zero OPD position corresponds to the object and reference mirror position being in focus.

In general, the OPD encoded in the fringes at each point varies with two parameters; the geometrical path difference $h - z$, and the wave number k . These parameters can be used to vary the OPD in a controlled way, and they are the critical variables that distinguish the various methods for creating and analyzing fringes, namely phase shifting, white light, wavelength scanning, spectrally resolved white light. Other methods for changing the OPD, which are not described here, are based on modifying the direction cosine $\cos \theta$ (Duan et al., 2006) and the refractive index in the optical path (Zelenka and Varner, 1969).

Fringe Visibility. Interference fringes as observed by each detector pixel (Eq. (15.8)) can be described in a simpler form as

$$I = I'(1 + \gamma \cos(\varphi)) \quad (15.10)$$

where I' is dc irradiance and γ is the fringe visibility (also called modulation, contrast or amplitude). Good fringe visibility is required for good measurement. To obtain good fringe visibility, the irradiances I_R and I_O need to be as equal as possible (see Eq. (15.8)). For this reason objectives may have reference mirrors with different reflectivities to match the test sample's reflectivity. In reality the fringe visibility is a more complicated function that decreases as the OPD increases and is affected by the temporal and spatial coherence of the source (Born and Wolf, 1999; Hariharan, 2005), thus by the wavelength bandwidth and size of the source. The apparent size of the source is in turn determined by the NA of the objective. Thus, the fringe visibility may vary over different height ranges of the object, and it is this fringe visibility that determines the maximum measurable range.

Influence of the Numerical Aperture of the Objective on Fringes. In addition to the effects of the NA of the objective on measurable height range, the observed wavelength of the fringes can differ slightly from the source. For high NA objectives the range of the incident angle of beam onto the object can affect the interference signal. Thus, a correction factor for the central wavelength (central wave number k_c) of the source must be used in order to get accurate height information (Bruce and Thornton, 1956; Gates, 1956; Tolomon, 1956; Ingelstam, 1960; Biegen and Smythe, 1988; Creath, 1989; Schulz and Elssner, 1991; Sheppard and Larkin, 1995; Dubois et al., 2000; Wan, et al., 2004). While for low NA objectives like NA = 0.1 the correction factor for the wavelength is about 1.0025 and can be neglected, for objectives with NA = 0.5 the correction factor rises to about 1.07 and for NA = 0.9 the correction factor is larger than 1.3. Since these interference microscope systems are complicated to model, most corrections are accomplished by using a traceable step-height standard to calibrate to correct scaling factor.

Limitations of Single-Wavelength Interferometric Optical Profilers. At first interferometric optical profilers for full-field measurement of object shape were based on single wavelength phase shifting interferometry (PSI) (for details see chapter on Phase Shifting Interferometry). These profilers delivered results with low noise, and smooth optical surfaces could be measured with very high precision, on the order of angstroms, while collecting only a few frames of data. However, PSI techniques were limited because they could only resolve smooth objects (optical roughness up to approximately $\lambda/30$) with height discontinuities less than $\lambda/4$. During PSI measurement the phase difference is changed in a few steps of a quarter of the fringe (90° or $\pi/2$), typically by having the PZT shift either the reference mirror or the object. The interference signal is analyzed at each point on the object using one of the many algorithms that were developed. Algorithms that compensate well for the

nonlinear motion of the PZT use eight frames of data (Schmit and Creath, 1996) to calculate the phase difference between beams over the measured area:

$$\varphi = \arctan\left(\frac{5I_2 - 15I_4 + 11I_6 - 2I_8}{I_1 - 11I_3 + 15I_5 - 5I_7}\right) \quad (15.11)$$

where irradiances I_1, I_2, \dots, I_8 are collected from single points in consecutive data frames.

Because PSI algorithms use an arctan (precisely atan $2(N,D)$) function, they can only determine the phase within modulo 2π , which means that only the fractional fringe order (fractional interference number) is determined and the relative fringe order has to be assigned during the spatial unwrapping procedure. The spatial unwrapping procedure assumes that the fringe order can not change from point to point by more than half an order (in phase terms by no more than π). The unwrapped phase has to be converted from radian to height units by means of a simple relation. For the object measured in reflection at normal incidence 2π corresponds to $\lambda/2$.

$$h = \frac{\lambda}{2} \times \frac{\varphi}{2\pi} \quad (15.12)$$

Thus, if the object has a height discontinuity larger than $\lambda/4$, then the fringe order will not be properly assigned and the object will not be measured correctly. This is called a 2π ambiguity problem.

Figure 15.26(a) shows interference fringes obtained in monochromatic illumination for a reflective binary grating. We can see that in this figure that it is not possible to determine the relative fringe order on both sides of the grating's discontinuities. Figure 15.26b shows fringes obtained for the same object using white light illumination. From this figure we see that the zero order fringe can be easily determined on both sides of the grating's discontinuities, which solves the problem of 2π ambiguity.

We will next discuss interferometric methods that solve the 2π ambiguity. These methods can be used to measure smooth and rough objects with step discontinuities up to a few mm. Sometimes even smooth objects without discontinuities cannot be measured correctly; this problem occurs when the slope of the object is so large that height difference between consecutive points is larger than $\lambda/4$, which corresponds to sampling the fringe with less than four pixels. Higher slopes may even not be able to be resolved by the detector.

15.5. TWO WAVELENGTH AND MULTIPLE WAVELENGTH TECHNIQUES

As mentioned previously, single wavelength interferometry has difficulty obtaining accurate measurements for objects that have high slopes. This difficulty occurs because the generated fringes are so dense that they are not able to be resolved by

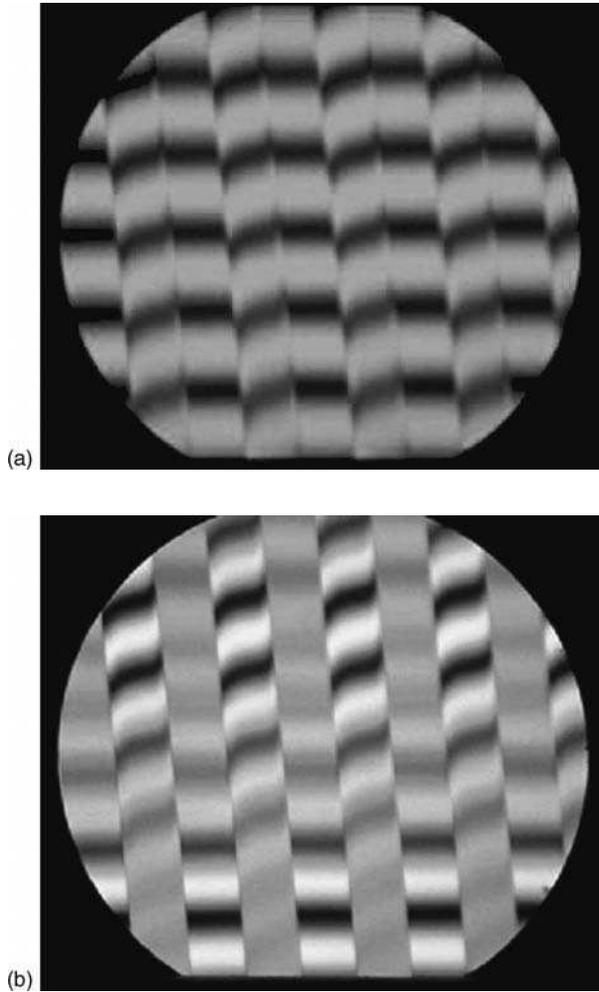


FIGURE 15.26. Fringes for object in form of 3D binary grating in (a) quasi-monochromatic and (b) white light illumination.

the detector. Two-wavelength techniques provide a way to expand the capabilities of single wavelength interferometry by creating fringes at a longer synthesized wavelength that corresponds to fringes that would be created if a long wavelength source were used (i.e., infrared source). Figure 12.57 in Chapter 12 shows fringes that are unresolvable in places using a single, short wavelength (a) and then fringes at different synthetic (effective) wavelengths (b–f). Fringes at the synthesized wavelength itself are analyzed rather than fringes at the two individual wavelengths that comprise the synthesized wavelength. The single synthetic wavelength is generated by using two short visible wavelengths simultaneously, and an interferogram is acquired that is identical to the one that would be obtained if a single longer

wavelength source were used. This technique enables a wider range of surfaces to be unambiguously and accurately measured without use of expensive long wavelength sources and detectors for these wavelengths. These observed fringes of longer effective wavelength are basically moiré fringes, which are described in Chapter 16. Interferometric techniques that employ two or more wavelengths have been described by many authors over the last 110 years (Hildebrand and Haines, 1967; Heflinger and Wuerker, 1969; Wyant, 1971; Polhemus, 1973; Benoit, 1898). Fringes of long effective wavelength were often used in two wavelength holography for aspherics measurement (Wyant, 1971), but they can also be used to measure rough surfaces.

15.5.1. Two-wavelength Phase Measurement

As long as the fringes generated at each wavelength can be resolved by the detection system, two measurement wavelengths can be used with phase-shifting techniques (Cheng and Wyant, 1984; Wyant, et al., 1984; Creath et al., 1985; Fercher et al., 1985; Creath and Wyant, 1986a; Creath and Wyant, 1986b; Wyant and Creath, 1989). These techniques are used to measure objects with height discontinuities larger than $(\lambda/4)$ at either of single wavelengths. By using the information from a second wavelength, the height range of the measurement can be significantly increased.

A two-wavelength phase measurement is performed by first taking data at one wavelength while shifting the phase in appropriate amount for that wavelength. The modulo 2π phase is then calculated for this first wavelength. The illumination wavelength is then changed, and data are taken at the second wavelength with the appropriate phase shifts; the modulo 2π phase is then calculated for this second wavelength. These two modulo 2π phase measurements can then be combined to produce a modulo 2π phase corresponding to a long synthetic wavelength, which is the beat between the two measured wavelengths. The phase corresponding to the new synthetic wavelength can be described as

$$\varphi_e = \varphi_1 - \varphi_2 = 2\pi\text{OPD} \left(\frac{\lambda_2 - \lambda_1}{\lambda_1\lambda_2} \right) = \frac{2\pi\text{OPD}}{\lambda_e} \quad (15.13)$$

where φ_1 and φ_2 are phases at wavelengths λ_1 and λ_2 ($\lambda_2 > \lambda_1$), and the effective wavelength, which then is described as

$$\lambda_e = \frac{\lambda_2\lambda_1}{\lambda_2 - \lambda_1} \quad (15.14)$$

Once the fractional fringe at the effective wavelength is determined, the fractional order is assigned by using the same spatial phase unwrapping procedures used with single wavelength techniques. Now, the new fringe order corresponds to the effective wavelength, and a measurement can be done correctly over larger height disconti-

nities. The effective wavelength in two wavelength interferometry is symbolically presented in Figure 15.27.

An alternate method for calculating the effective wavelength phase is to take all the frames of data for both wavelengths and then calculate the phase difference between individual wavelength phases instead of first calculating each individual wavelength phase and then subtracting it. In this method the effective wavelength phase can be calculated directly from the irradiance data. This calculation can be written as

$$\phi_e = \tan^{-1} \left[\frac{\sin(\phi_1 - \phi_2)}{\cos(\phi_1 - \phi_2)} \right] = \tan^{-1} \left[\frac{\sin \phi_1 \cos \phi_2 - \cos \phi_1 \sin \phi_2}{\cos \phi_1 \cos \phi_2 + \sin \phi_1 \sin \phi_2} \right] \quad (15.15)$$

In general any PSI algorithm can be implemented to two wavelength interferometry. If a PSI algorithm is described by numerator N and denominator D , the phase is expressed as

$$\phi_i = \tan^{-1} \left[\frac{\sin \phi_i}{\cos \phi_i} \right] = \tan^{-1} \left[\frac{N_i}{D_i} \right] \quad (15.16)$$

and then the effective wavelength phase can be obtained from

$$\phi_e = \tan^{-1} \left[\frac{\sin(\phi_1 - \phi_2)}{\cos(\phi_1 - \phi_2)} \right] = \tan^{-1} \left[\frac{N_1 D_2 - D_1 N_2}{D_1 D_2 + N_1 N_2} \right] \quad (15.17)$$

The required phase shift between frames typically equals 90° . In order to realize a 90° phase shift for each wavelength, when using a PZT phase shifter, the PZT needs to shift by different distances that correspond to each wavelength used. However, if a polarization interferometer is used, then an achromatic phase shifter (Hariharan, 1996) introduces the proper 90° phase shift for the selected wavelengths (see also Section 15.9.2). When a system uses a PZT phase shifter that is calibrated only to a

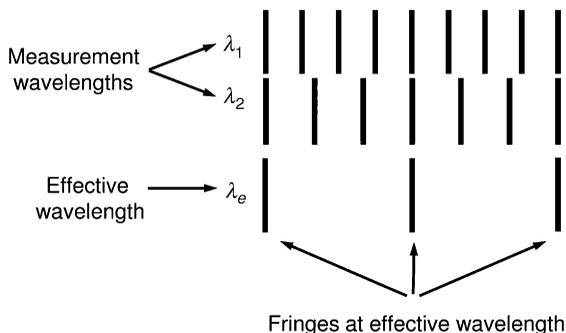


FIGURE 15.27. Beat wavelength, called effective or synthetic wavelength, for two-wavelength interferometry.

single wavelength, for frame sets with phase shifts different than 90° , a phase calculation based on least-square method (Kim et al., 1997) can be used. An even better solution has the phase being calculated using an algorithm that is insensitive to large phase shift miscalibrations (Carré, 1966 Schmit and Havihoran 2006). The phase shift can also be introduced by a frequency change in the laser diodes (Ishii and Onodera, 1991).

Table 12.6 in Chapter 12 lists the values of λ_e that can be obtained using various pairs of wavelengths from an argon ion and a helium–neon laser. By using a dye laser, a large range of equivalent wavelengths can be obtained (Schmidt and Fercher, 1971). Tunable helium–neon lasers with four or five distinct wavelengths ranging from green to red are also available (Wyant, 1971). A range of distinct wavelengths can be obtained with tunable and compact laser diodes for which the wavelength stability needs to be considered (de Groot and Kishner, 1991). In interference microscopes a white light source followed by narrowband spectral filters (Creath 1986) that are typically based on different laser lines are used.

de Groot (1994) has shown that if the fractional phases at single and effective wavelengths are known, the dynamic range does not need to be limited by the effective wavelength in two wavelength interferometry; rather, through analytical manipulations the wavelength can be extended to multiples of the effective wavelength where the multiplier N equals

$$N = \left\lceil \text{int} \left[\frac{1}{\lambda_e/\lambda_1 - \text{int}(\lambda_e/\lambda_1)} \right] \right\rceil \quad (15.18)$$

For example, for green and red spectral emissions and a red wavelength of 644 nm, the effective wavelength equals 2.42 μm , but with the calculated multiplier it can be extended to 10 μm . However, this technique is limited by long calculation times and noise in the measurement. Because of the measurement noise, an approach described in the next section was developed that uses the second wavelength measurement only to correct a single wavelength measurement.

Correction of Single-wavelengths Measurements. The noise in a two-wavelength measurement is proportional to the length of the wavelength used. For example, if there is an RMS measurement noise of 0.01 μm at $\lambda = 0.5 \mu\text{m}$, there will be an RMS noise of 0.1 μm with an effective wavelength of $\lambda_e = 5 \mu\text{m}$. A two-wavelength measurement can be used to correct the phase ambiguities in the modulo 2π single wavelength phase to provide a measurement with visible wavelength precision and extended height range (Creath, 1986). This precision is achieved by comparing a scaled version of the long effective wavelength phase with the single wavelength phase. The number of 2π s to add to the single wavelength data are determined by looking at the height changes in the scaled effective phase. This correction works well for relatively smooth data. If the noise in the scaled effective phase is greater than $\pm\lambda/4$ between adjacent pixels at the single wavelength then unwanted 2π jumps occur in the corrected data.

15.5.2. Multiple-wavelength Phase Measurement

Two-wavelength techniques can be extended to multiple wavelengths in order to correct single wavelength data (Cheng and Wyant, 1985; Dandliker et al., 1992; de Groot, 1991; Decker et al., 2003; Towers et al., 2003). A number of wavelengths are specifically chosen so that a series of effective wavelengths are produced that are proportionally spaced from the single wavelength up to the wavelength necessary to measure the test object. A rule of thumb for good measurements is to keep the ratio between the longer wavelength and the wavelength being corrected between a factor of five and ten, and this ratio is limited depending on the level of phase noise in the measurements. For situations with low phase noise, the ratio can be made larger, and a reduced number of wavelengths are required to span the range between the shortest measurement wavelength and the desired measurement range.

An example of a 13- μm step measured using single, two-wavelength, and multiple-wavelength techniques is shown in Figure 15.28. These measurements were taken using an interferometric optical microscope with phase-shifting capability. Figure 15.28(a) shows the step measured at a wavelength of 657 nm. A two-wavelength measurement is shown in Figure 15.28(b) where the measurement wavelengths are 657 nm and 651 nm producing an effective wavelength of 64 μm . The difference between two consecutive measurements using two wavelengths is shown in Figure 15.28(c) where the RMS is 7.13 nm. This means that the measurement is repeatable to within $\lambda_e/9000$ at the effective wavelength. Using the data from the 651 nm measurement to correct the phase data taken at 657 nm, the result given in Figure 15.28(d) shows uncertainties of 2π , which are caused by noise in the single wavelength measurement. If three measurement wavelengths (657 nm, 651 nm, and 601 nm) are used, the corrected measurement at 657 nm shown in Figure 15.28(e) is much less noisy. The repeatability (difference between two consecutive measurements) of the three-wavelength measurement of Figure 15.28(e) is shown in Figure 15.28(f). The RMS of the difference in measurement is 0.67 nm, which yields a dynamic range for the measurement of almost 20,000. Thus, the use of multiple wavelengths can increase the dynamic range of a measurement by a factor of 10. Towers et al. (Towers, Towers and Jones 2003, 2004a, 2004b, 2005) described how to choose an optimal series of multiple wavelengths to create a geometric series of effective wavelengths that yields the greatest increase in dynamic range for each subsequent effective wavelength. The series of four wavelengths in the example given above is an example of an optimal series.

Correction of Single-wavelength Measurements. In multiple-wavelength interferometry like in two wavelength interferometry (Section 15.5.1), in order to correctly resolve large height discontinuities fringe the order must also be determined. The longer measurement wavelengths enable larger height discontinuities to be measured, but measurement noise increases proportionately. A longer wavelengths' fringe order can be used to determine fringe order at a shorter wavelength; once fringe order at this shorter wavelength is established, the phase with relatively

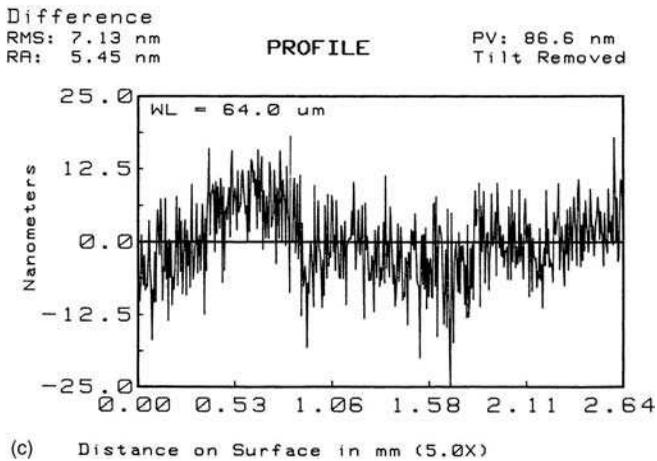
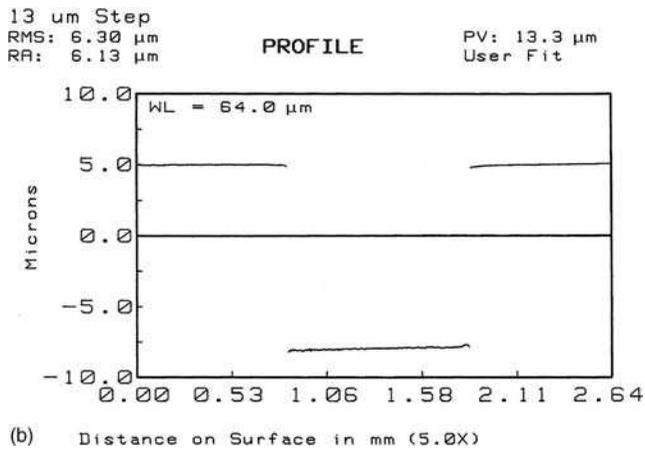
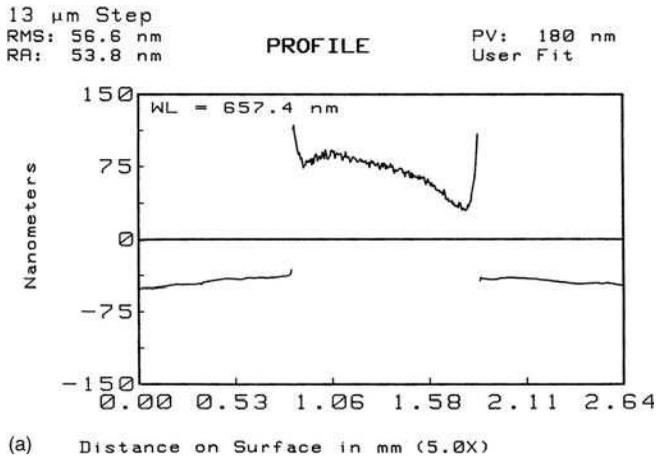
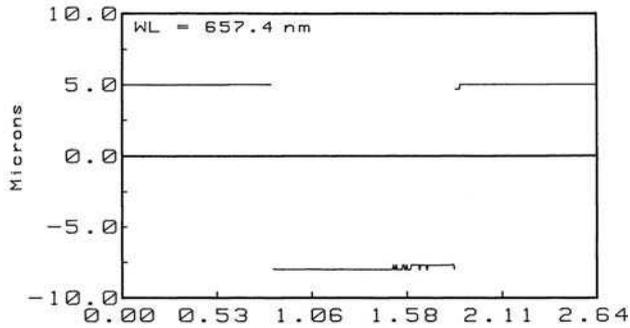


FIGURE 15.28. Measurement of 13 μm step using multiple wavelength interferometry: (a) 657 nm, (b) two-wavelength measurement using 657 and 651 nm with $\lambda_e = 64 \mu\text{m}$, (c) difference of consecutive two-wavelength measurements, (d) 657 nm data corrected using two-wavelength measurement, (e) three-wavelength using 657, 651, and 601 nm, and (f) difference of consecutive three-wavelength measurements.

13 μm Step
RMS: 6.30 μm
RR: 6.14 μm

PROFILE

PV: 13.3 μm
User Fit

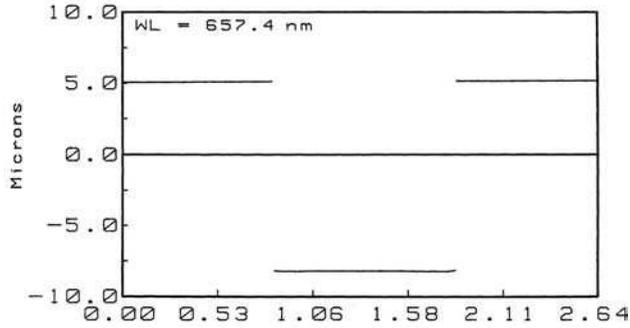


(d) Distance on Surface in mm (5.0X)

13 μm Step
RMS: 6.49 μm
RR: 6.31 μm

PROFILE

PV: 13.4 μm

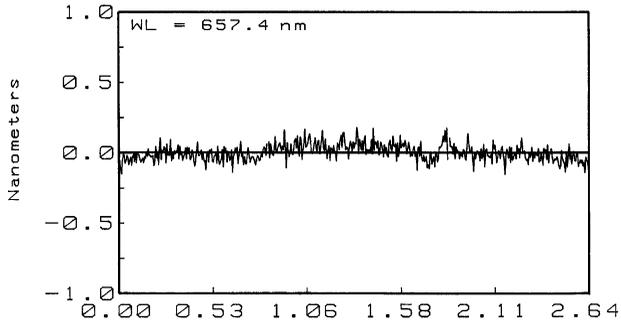


(e) Distance on Surface in mm (5.0X)

Difference
RMS: 0.064 nm
RR: 0.049 nm

PROFILE

PV: 0.671 nm
Tilt Removed



(f) Distance on Surface in mm (5.0X)

FIGURE 15.28. (Continued)

low noise level can be unwrapped. In this way effective wavelengths created in two- and multiple-wavelength interferometry are important for extending the range of resolvable heights. A powerful technique for determining fringe order in multiple wavelength interferometry uses a temporal phase unwrapping process (Huntley and Saldner, 1993, 1997; Saldner and Huntley, 1997a, 1997b) rather than the typical spatial process. This temporal phase unwrapping at each pixel approach works on a series of multiple wavelengths; the values of these multiple wavelengths must form a series of decreasing geometrical numbers.

15.5.3. Reducing Measurement Time

The time required for taking a measurement when using a two- or multiple-wavelength technique is at least twice as long as when a PSI single wavelength technique is used. In an effort to reduce measurement time, two or multiple wavelength superimposed interferograms can be captured in one frame and then analyzed if the fringes have a carrier frequency (Onodera, 1997). In this case moiré-like interference fringes at the effective wavelength are observed, but if the Fourier transform is applied, then the main frequencies corresponding to interferograms for different wavelengths can be separated and the fractional phases for individual wavelengths can be calculated.

Pförtner and Schwider used a color CCD camera to capture in one snapshot three frames of fringes for wavelengths from three laser sources that corresponded to the RGB colors of the camera (Pförtner and Schwider, 2001). This technique, called RGB interferometry, employed a large wavelength separation (633, 532, 473 nm). In addition, so as to reduce error an axial chromatic dispersion was subtracted without doing any additional measurement. If fringes with carrier frequency were used, only one frame would be needed. However, the carrier approach has its limitations.

Different methods capture simultaneously a few phase shifted interferograms that are spatially separated and captured by one or multiple cameras. Two or multiple sets of interferograms corresponding to two or multiple wavelengths can be registered successively over a period of only 100 μs . A recently introduced separation of interferograms can be done with a specially developed pixelated phase mask that introduces a unique phase shift at each pixel (North-Morris et al., 2004). These methods do not need fringes with a carrier frequency.

Multiple wavelength approaches are also found in other areas of interferometric metrology such as speckle or digital holography. The multiple wavelength approach has also been used in fringe projection methods to increase the dynamic range of their measurements by projecting fringes of multiple frequencies that act like fringes of multiple wavelengths (see Chapter 16).

The applicability of multiple-wavelength methods in interference microscopy may be limited by the depth of field of the interference objective that determines measurable heights (see Section 15.4.1) rather than by the effective wavelength. For this reason it is often more practical to use white light interference for which measurable heights are limited by the working distance of the objective.

15.6. WHITE LIGHT INTERFERENCE OPTICAL PROFILERS

White light interference (WLI) optical profilers use broadband illumination and work like an array of optical focus sensors where the position of the interference signal at each sensor determines the best focus position. The use of broadband illumination overcomes some of the limitations that are found in single and even multiple-wavelength methods. WLI methods have long established themselves as the leading optical profilometers for measuring engineering surfaces like MEMS devices, binary optic, and machined surfaces. The vertical resolution of WLI depends on the analysis of the signal and can be as good as single wavelength PSI methods (0.3 nm), but more commonly is around 3 nm. Vertical resolution here is defined as RMS of the difference measurement on smooth sample.

15.6.1. White Light Interference

A white light source used in an interference optical profiler has a broadband visible spectrum with wavelengths from about 380 up to 750 (violet to red) nanometers. The source has low temporal coherence because of the large wavelength bandwidth, and it is not considered a point source, which means that it also has low spatial coherence. The low temporal and spatial coherence of the source creates interference fringes that are localized in space.

In order to obtain fringes at best focus, the position of the reference mirror needs to be set also at the best focus of the objective. This is done in three steps: first, the reference mirror is moved a few or tens of microns away from focus; second, the objective is focused on the object with some features like the edge of a sharp but not too tall step (fringes are not visible at this moment); and third, the reference mirror is brought to focus and stopped when best contrast fringes are obtained. The reference mirror of the interference objective is set at the best focus of the objective in order to obtain the zero OPD.

Because low temporal coherence has a stronger influence on fringe localization than low spatial coherence, temporal effects will be the focus of this discussion. The different wavelengths from the source spectrum are mutually incoherent and the superposition of fringes for individual wavelengths creates white light fringes as shown in Figure 15.29. A monochromatic detector observes the sum of all the fringe intensities. Because the spacing of the fringes for each wavelength of the source is different, the maxima of fringes will align around only one point where the OPD is zero for all wavelengths as shown in Figure 15.29a. Away from this zero OPD position the observed sum of the intensities quickly falls off as shown in Figure 15.29b. It is for this reason that fringes are said to be localized. The fringe with maximum contrast, the fringe that marks the zero OPD, is called the zero order fringe, and each next fringe of smaller amplitude on either side is called +1 and -1, +2 and -2 order fringe and so on. The maximum of the zero order fringe does not need to fall at the maximum of the fringe envelope (see Section 15.6.5).

Looking back at Figure 15.26, we see both white light fringes created for a binary grating and quasi-monochromatic fringes for the same object after a narrow band

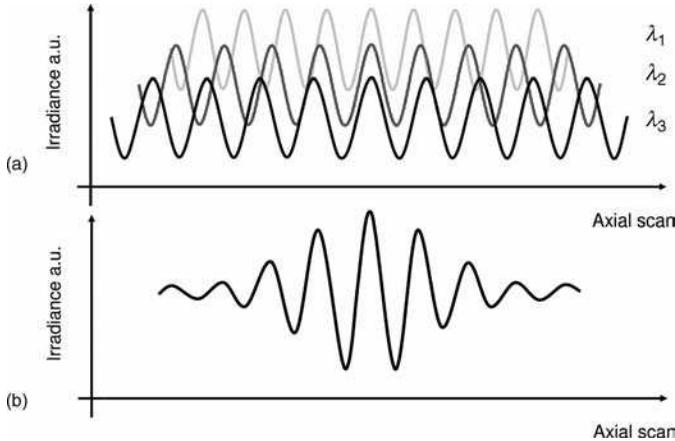


FIGURE 15.29. Formation of white light fringes: (a) fringes for individual wavelengths and (b) sum of fringes of individual wavelengths, which are white light fringes.

filter is placed in front of the white light source. This pair of interferograms illustrates that when looking at fringes created using a white light source, the zero-order fringe can easily be found across the object, and thus surface shape can be determined without ambiguity. This elimination of ambiguity in numbering fringes (2π ambiguity) is a major strength of WLI because it allows for measurement of samples with large discontinuities and rough surfaces.

In mathematical form this white light interference observed by one pixel during an axial scan can be described as the integral of all the fringes for all wave numbers k and for different incident angles (for example, de Groot and de Lega, 2004; Abdulhalim, 2001). The resulting fringes in general can be described as

$$I(z) = I'[1 + \gamma(z) \cos(k_0 z)] \quad (15.19)$$

where I' is the background intensity, $\gamma(z)$ is the fringe visibility function or coherence envelope and $k_0 = 2\pi/\lambda_0$ is the central wave number for fringes under the envelope. $\gamma(z)$ is proportional to the modulus of the Fourier transform of the source spectrum. Generally, if the light source has a Gaussian spectrum, then the envelope of the fringes can be described also as a Gaussian function $\gamma(z)$. The broader the bandwidth of the source spectrum, the narrower the width of the envelope. The width of the fringe envelope is determined by the coherence length of the source (see Eq. 15.7); for a white light source this width is of the order of 1–2 μm . The envelope of the fringes varies with other factors like sensitivity of the camera, the measured object, and dispersion in the system also.

15.6.2. Image Buildup

The important feature of white light fringes for surface topography measurement is the fact that fringes are localized and can only be found within microns or tens of

microns of the zero OPD as shown in Figure 15.30. As the objective (or the sample) is scanned axially through focus, each pixel registers irradiance; the highest point on the fringe envelope determines the best focus position on the sample. Figure 15.31 shows a few interferograms as registered by CCD camera as a sample is progressively scanned through focus. Fringes at individual interferograms show which part of the sample is in focus for a given position of the scan.

15.6.3. Signal Processing of White Light Interferograms

The shape of the object is determined from the localization of the fringes at each spatial point registered during the axial scan. It is assumed that the fringe signal is the same at each point and only its axial position is different due to changes in the topography of the test sample. Since 1980 (Balsubramanian, 1982), a number of methods and algorithms have been developed that describe the use of white light interferometry. Many algorithms first compute the envelope (modulation) of the fringes. The fringe envelope can be calculated in the same way as the modulation of fringes is determined in PSI. During the axial scan, the OPD is changed to introduce typically 90° phase shift between registered frames but over a much longer scan range than with PSI. Then, any PSI algorithm can be implemented to determine the modulation of the fringes at each point along the axial scan.

For a common 5-frame PSI algorithm (Schwider et al., 1983; Hariharan et al., 1987) the modulation would be calculated as

$$\gamma(z) = \frac{\sqrt{(I_2 - I_4)^2 + (I_1 - 2I_3 + I_5)^2}}{\text{normalization}} = \frac{\sqrt{N^2 + D^2}}{\text{normalization}} \quad (15.20)$$

where the N and D represent numerator and denominator of any PSI algorithm. Properties of PSI algorithms that can be used for modulation calculation were nicely reviewed by Larkin (1996b). Kino and Chim (1990) proposed using the Fourier transform technique to calculate the envelope where first the forward Fourier transform of the interference signal is computed. After these calculations three lobes are observed. One sidelobe, which is positioned at the frequency of the fringes, is

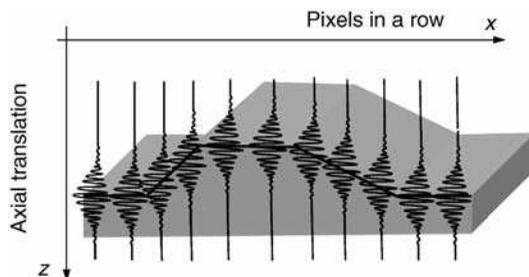


FIGURE 15.30. Irradiance signals as observed by a few pixels in a row for an object placed in white light interferometer.

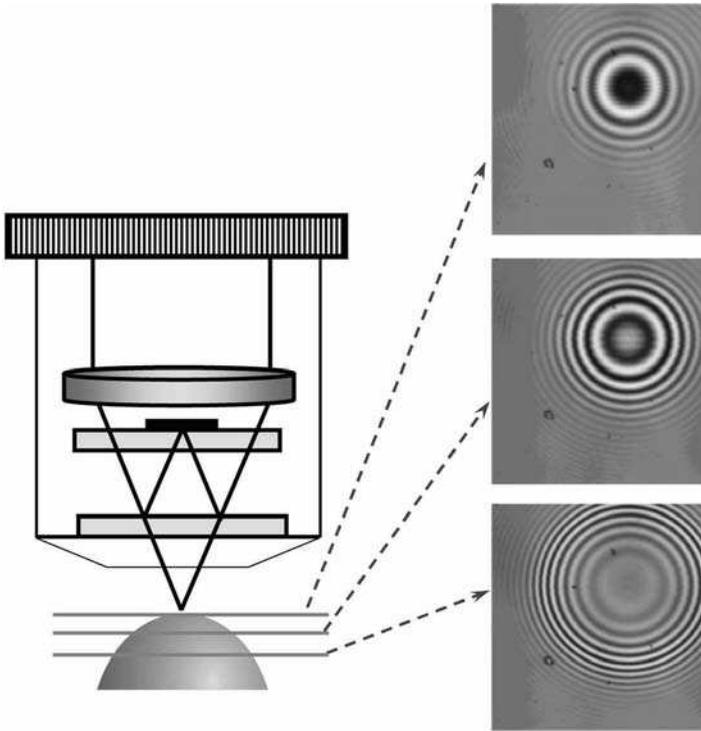


FIGURE 15.31. White light interferograms for a spherical object as obtained for a few positions of the objective during an axial scan.

isolated and shifted to the center. Next the inverse Fourier transform is calculated to obtain the fringe envelope. Caber (1993) proposed using electronic hardware (amplitude demodulation in hardware) to obtain the envelope of fringes.

Once the envelope is determined, its position can be found by fitting the curve to the envelope and finding its position. The position of the envelope can be also found by calculating the envelope's center of mass using the equation:

$$h = \frac{\sum_{i=1}^{N-1} \gamma_{Zi} z_i}{\sum_{i=1}^{N-1} \gamma_{Zi}} \quad (15.21)$$

where γ represents the envelope function, z the axial position, and h the object's height. The center of mass algorithm is very fast and computationally efficient, and often is implemented in confocal systems for finding the maximum of the confocal irradiance signal. Center of mass calculations are equivalent to calculations of the maximum of the envelope position but only for a symmetrical signal. For an asymmetrical signal, a piston is introduced for each point; however, this piston

does not affect the whole measurement. The most precise measurement of the fringe envelope is done using an achromatic phase shifter. At each axial scan position, an achromatic phase shifter, which is described in Section 15.8.9, shifts the fringes underneath the envelope, which results in the registered constant modulation of the fringes. Instead of finding the position of the envelope, it is also possible to find the position of the bright or dark fringe around the envelope's maximum (Park and Kim, 2000).

De Groot and Deck (1995) showed that finding the position of the fringe envelope can be done by processing it in the frequency domain. First, the Fourier transform is calculated (similar to Kino and Chim) and one sidelobe is isolated. The magnitude of the sidelobe represents the strength of the spectrum at a given wavelength, and the phase represents the phase of the interference signal for a given wavelength. Thus, from this sidelobe, each interference component of the white light signal can be recreated $I(k) = I(k) \times \cos(\varphi(k))$, where $\varphi = kz$. If the phase for at least two wavelengths is known, z can be determined without 2π ambiguity from $z = \Delta\varphi/\Delta z$. Other methods like wavelets analysis also can be implemented (Yatagai, 1994; Itoh et al., 1995; Sandoz, 1997; Recknagel and Notni, 1998); the upside of this method is reduction in noise but the downside is longer processing time.

At first, white light interference microscopes were used for testing of smooth surfaces (Davidson et al., 1987; Kino and Chim, 1990; Lee and Strand, 1990) and then extended to measurement of rough surfaces (Häusler and Neumann, 1992) in which the presence of speckles may need to be considered (Häusler and Herrmann, 1992; Pavlíček and Soubusta, 2003). The signal can be analyzed to obtain not only the shape of the object but also to map of an object's different reflectivities or an object image as seen by the objective with infinite depth of focus can be displayed (Sheppard and Roy, 2003). This is not a complete list of the literature describing algorithms for WLI.

Sampling of White Light Interference Signal. Measurement time in WLI varies with the required scan length; thus, the sampling rate of WL fringes has to be carefully considered so as to obtain the best data within the shortest time frame. In order to localize fringes, the value of the envelope's amplitude needs to be known at only a few axial points (Larkin, 1996a, 1996b; Creath, 1997) and not all fringes need to be resolved during the scan. For example, calculating the fringe modulation with PSI algorithms, which was described in Section 15.6.3, requires four samples per fringe; however, if sampling is done not at four samples per fringe but rather at four samples per odd number of fringes, faster but less precise WLI measurements can be obtained. Sampling with four samples per odd number of fringes is equivalent to sampling at every 90° ; not all fringes will be resolved, but calculating the fringe modulation is still possible. Using this method, measurement speeds can be increased 23 times (Schmit, 2003) up to $100 \mu\text{m s}^{-1}$ using a 60 frames per second camera frame rate. The sampling rate would then equal about $1.8\text{--}2 \mu\text{m}$ and is on the order of the coherence length of the white light source. When the sampling rate approaches the coherence length of the source, data become unusable. When this occurs, the

envelope of the fringes needs to be lengthened by reducing the spectral bandwidth of the source. Sampling issues in different WLI algorithms were discussed by a number of authors (Deck and de Groot, 1994; Larkin, 1996a, 1996b; Creath, 1997; Hirabayashi et al., 2002; Schmit, 2003). Fringe projection used in a stereoscopic microscope creates localized fringes that are analogous to fringes in white light interferometry but of longer wavelength and envelope; thus, the sampling rate can be much larger than a few microns (Kröner et al., 2001; Kröner et al., 2006).

Increased Resolution White Light Interferometry. PSI methods achieve about 10 times better vertical resolution (0.3 nm vs. 3 nm) than WLI methods when measuring the position of the fringe envelope; however, if the phase of the fringes under the envelope is found, WLI methods can achieve similar vertical resolution. This high resolution WLI method combines a lower resolution map of the envelope position (relative fringe orders) and a higher resolution map of the phase (position) of the zero-order fringe. The calculation of these two maps and the combination of them can be accomplished using various algorithms (Cohen et al., 1992; Larkin, 1996b; Windecker et al., 1999; Harasaki et al., 2000; de Groot et al., 2002). This high resolution WLI is particularly well suited for determining the shape of smooth surfaces with large height differences such as binary diffractive optics or micro-electromechanical systems (MEMS). The advantage of this method is that the phase is calculated always at the best focus position. Examples of measurements with white light interference optical profilometer are given in Figure 15.32.

15.6.4. Light Sources

Different white light sources, such as a tungsten-halogen, incandescent or arc lamp, LEDs and SLDs can be used for illumination. These sources have different spectra and thus create different fringe envelopes. The width of the fringe envelope is determined by the bandwidth of the source spectra. In Figure 15.33, we see that the two sources, a halogen lamp and a red LED, their spectra having different bandwidths, generate fringes with different envelope widths. The narrower the envelope, the more precisely the localization of fringes can be determined. The spectra of semiconductor light sources, such as light emitting diodes (LEDs) and superluminescing laser diodes (SLDs), are similar in shape to a Gaussian function.

15.6.5. Dispersion in White Light Fringes

In Figure 15.29, it was assumed that a white light interferometer is compensated for all wavelengths, meaning that the position of the maximum of the fringes aligns with the maximum of the envelope, namely where there is a zero phase shift $\varphi_0 = 0$ between fringe and envelope maxima. If there is an odd number of reflections from dielectric surfaces in one arm of the interferometer and an even number in the other, the fringes will be shifted by $\varphi_0 = 180^\circ$ under the coherence envelope, and the minimum of the fringe will align with the maximum of the coherence envelope.

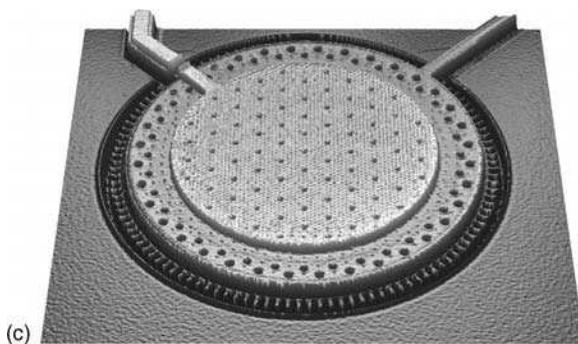


FIGURE 15.32. Examples of object measurement with white light interferometer (a) Forensic bone sample, $460\ \mu\text{m} \times 612\ \mu\text{m}$, (b) salient type micro-motor, $230\ \mu\text{m} \times 304\ \mu\text{m}$ (c) solderless MEMS microphone for cellular phones and other applications, $0.9\ \text{mm} \times 1.2\ \text{mm}$ (d) honed cylinder wall, $1.2\ \text{mm} \times 0.90\ \text{mm}$. Courtesy Veeco Instruments.

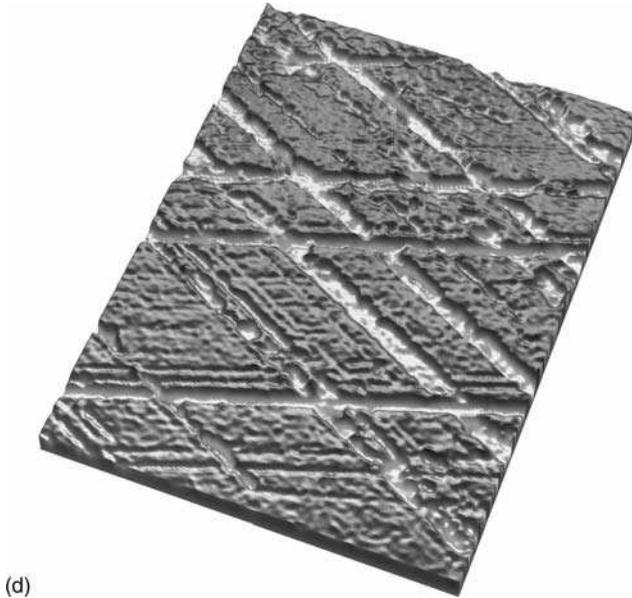


FIGURE 15.32. (Continued)

In a real system the fringes may be shifted with respect to the envelope by any amount of φ_0 , and this shift may be due to any number of factors. These factors include reflection from non-dielectric surfaces and transparent films, dispersion and lateral aberrations of the system or the presence of a transparent plate in one of the arms of the interferometer. These factors cause a phase shift that varies with the wave number, which is also called the spectral phase. The spectral phase introduces changes into the white light interferogram; it changes the envelope and position of the fringes and the fringe frequency. If the introduced spectral phase is linear with the wave number, the fringe location changes but the shape of the envelope of fringes remains the same. The constant phase change on reflection for all wave numbers only shifts the fringes underneath the coherence envelope, like an achromatic phase shifter would (see Section 15.9.2.). Higher order changes in the spectral phase will influence the position, shape, and amplitude of the envelope as well as the fringe frequency.

Dissimilar Materials. As long as the object's surface is comprised of a single material, the spectral phase does not present a problem since a uniform over field change is introduced. However, when two dissimilar materials are side-by-side on the surface, they will introduce different phase shifts upon reflection for different wavelengths (unless both of them are dielectric materials with the imaginary part of the index of refraction $k = 0$) and the measured height difference at the boundary where the two meet will be incorrect. By knowing the optical constants of the different materials for the wavelengths used in the measurement, it is possible to correct for this difference (Bennett, 1964; Church and Lange, 1986; Biegen and

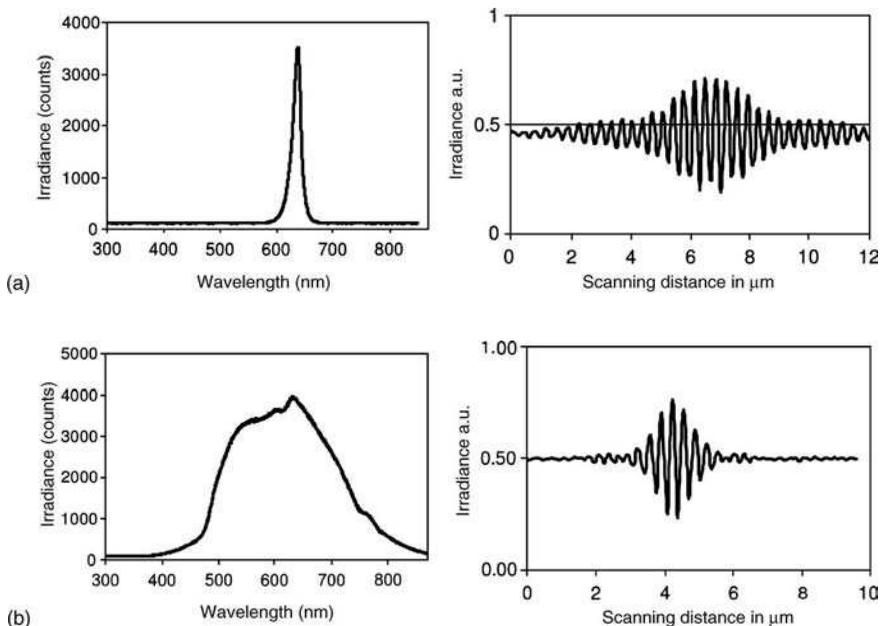


FIGURE 15.33. Spectrum and interferogram for (a) halogen lamp and for (b) red LED source.

Smythe, 1988; Doi et al., 1997; Rogala and Barrett, 1998; Harasaki et al., 2001; Park and Kim, 2001). The phase change on reflection for bulk materials for normal incidence for a given wave number k is described as

$$\varphi_{\text{material}}(k) = \tan^{-1} \left(\frac{2k}{1 - n^2 - k^2} \right) \quad (15.22)$$

where n and k are the real and imaginary index of refractions, and the values of n and k for the indices of refraction for the range of wavelengths of the materials used in the test object can be found in *The Handbook of Optical Constants of Solids* by Palik (1991). Often a better solution when faced with a test object made of a composite material, such as ceramic, is to coat the material with a layer (typically 100 nm is sufficient) of opaque material (i.e., a metal) so as to obtain a good profile of the surface. Alternately, replicas of the test object can be made and then measured.

When a wide spectrum source is used, the phase change on reflection over the entire spectrum needs to be considered. Figure 15.34 shows spectral phases for gold and silicon as examples of metal and semiconductor materials. In white light interferometry the spectral phase introduced by the different materials of the object will shift the peak of the envelope (most of metals) and possibly even change the shape of the envelope (gold and possibly some semiconductors). For typical materials this shift will not be larger than 40 nanometers (Harasaki et al., 2001). Table 15.3 shows the shift in the peak of the envelope for white light fringes and the phase of the fringes for 600 nm wavelength.

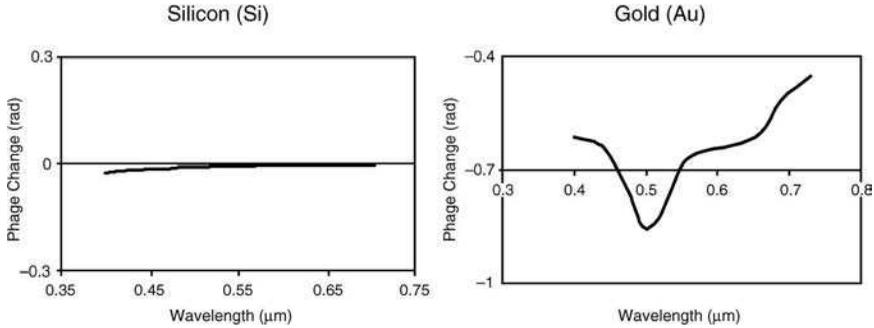


FIGURE 15.34. Phase change on reflection for different wavelengths: (a) silicon and (b) gold.

Thick Film Measurement. If the sample is covered with a transparent film that is more than a few microns thick, two sets of localized fringes separated from each other are generated, one for each interface. A thick film technique is then used to measure the film thickness. For the second interface, the bottom of the film, a phase shift between interference patterns for individual wavelengths proportional to the product of the geometrical path and the index of refraction equal to $d \times n(k)$ is introduced, and the irradiance for the white light fringes for the second interface can be described as

$$I(z) = \int_{k_1}^{k_2} \langle 1 + V(z) \cos\{kz - kd[n(k)]\} \rangle dk \quad (15.23)$$

The dependence of the refractive index on the wave number k can be described to first approximation as a linear expansion:

$$n(k) = n(k_0) + \frac{dn}{dk}(k - k_0) \quad (15.24)$$

The linear dispersion shifts the envelope by the group index of refraction times the thickness of the dispersive element; this dispersion also shifts the fringes under the envelope slightly.

TABLE 15.3. Offset in measured heights due to shift of envelope peak position and monochromatic fringes at 600 nm wavelength due to phase change on reflection for different metals.

Metal	Height offset due to envelope peak shift (nm)	Height offset due to fringes shift (nm)
Silver	36	25.1
Aluminum	13	12.7
Gold	0	33.4
Molybdenum	59	13.4
Nickel	15	20.8
Platinum	13	18.1

A simple technique for finding the relative position of the peaks of the fringe envelopes can be implemented to find the thickness of a film. Figure 15.35 shows two almost clearly separated sets of fringes formed for the air/film and film/substrate interfaces. The typical range of measurable film thicknesses runs from 3 to 150 μm depending on the dispersion of the film and NA of the objective.

Higher order dispersion introduced by a thicker film or an inserted plate may have many effects; the envelope may widen or even become asymmetrical, the position of the fringes may shift under the envelope, fringes may lose contrast, or the period of the fringes may change with the z position (Pavlíček and Soubusta, 2004). Dispersion effects will be stronger for sources with a wider spectrum and objectives with a higher NA. However, the observed changes will be different for different shapes of the spectra. Thus, to measure thicker films, it is better to use a low numerical aperture objectives and a narrower bandpass of the white light source.

Thin Film Measurement. When the optical thickness of the film is shorter than the coherence length of the white light source, typically less than 3 μm , multiple reflections introduce an additional nonlinear term in the spectral phase that causes changes in the fringe envelope and the frequency of fringes (Hariharan and Roy, 1996; Roy et al., 2005). Finding the envelope peak position is not valid any longer since the localized fringes at both interfaces are not separated. A different approach has to be taken. One approach is to apply the Fourier transform of the measured signal and calculate the spectral phase under the first lobe in the frequency domain (Kim and Kim, 1999). Figure 15.36 shows fringes for thin film (a), the magnitude of first spectral lobe (b), and the corresponding spectral phase in the frequency domain (c). In general, when using interferometry to measure thin or thick films, not only their thickness, but the top and bottom profiles are also measured. The spectral phase for the thin film interference has the form of a polynomial; thus, the polynomial for the chosen film model (n and k) is fitted, and regression analysis is used to find the best fit and, therefore, the film thickness. The spectral phase due to the dispersion of the system

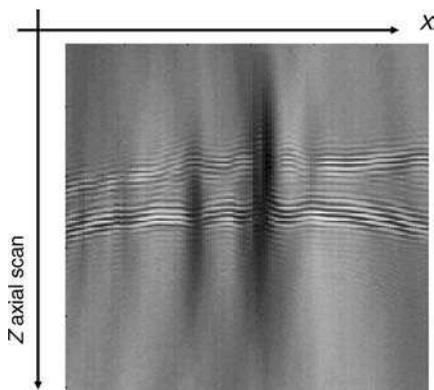


FIGURE 15.35. White light fringes as observed by a row of pixels during axial scan for (a) nonconformal film, where the top fringes are created at the air/film interface while the bottom fringes correspond to film/substrate interface but are located below the interface.

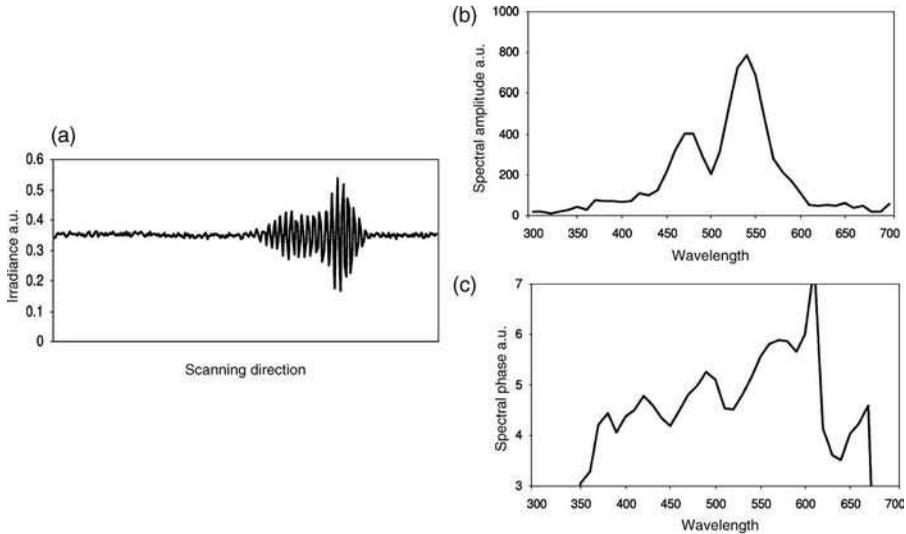


FIGURE 15.36. (a) White light fringes for thin film, (b) the side lobe of the Fourier transform magnitude, and (c) spectral phase under the side lobe.

needs to be known and subtracted for better accuracy. Instead of calculating the spectral phase and finding best fitting simulated spectral phase, the magnitude can be calculated, and the best fitting simulated magnitude can be found to determine film thickness. These methods are used for films of thickness from a few microns down to 100 nanometers. For films of optical thickness less than 100 nm, the sensitivity of the method drastically decreases.

Besides special analysis necessary for measuring objects with dissimilar materials and films, objects with narrow trenches (10 μm wide and less) also require careful examination of the fringes. For a narrow structure, additional unwanted fringes are created that do not correspond directly to the object shape (Schmit et al., 2003; Montgomery et al., 2004; Tavrov et al., 2005)

Measurement through the Glass Plate or Liquid Media. Many engineering objects, like MEMS devices, are often protected by a cover glass, and some devices in environmental chambers need to be tested under different pressures or temperatures. Such objects require testing through a cover glass. Biological samples are often immersed in liquid and require measurement through this liquid. Because of the dispersion of the liquid layer or cover glass white light fringes may be totally washed out. Thus, a compensating plate needs to be introduced in the reference arm of interferometer. This compensation is the most easily done for Michelson-type objectives. In addition, the contrast of fringes diminishes faster for higher numerical aperture objectives, which are used as both illuminating and imaging optics. It is also difficult to introduce compensation for higher NA objectives. In order to increase fringe contrast for systems with high numerical aperture objectives, the numerical aperture of the illumination can be reduced by delivering a nearly collimated

illuminating beam (Han 2006) of very low numerical aperture directly to the interferometer underneath the objective and not through the objective as shown in Figure 15.37(a). Figure 15.37(b) shows a measurement of grating immersed in liquid and a pitch standard as measured through 3 mm cover glass with 20X objective.

15.6.6. Other Names for Interferometric Optical Profilers

White light interferometry with an axial scan for the measurement of engineering surfaces has been variously labeled and can also be found under the following names:

White light interferometry (WLI)

Vertical scanning interferometry (VSI)

Low coherence interferometry (LCI)

Coherence probe

Optical coherence profilometry (OCP)

Optical coherence microscopy

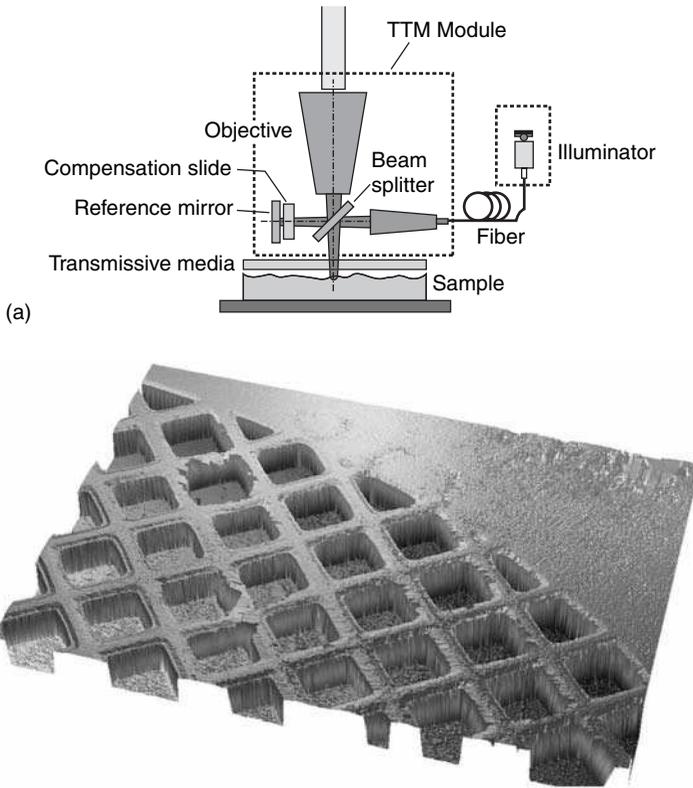


FIGURE 15.37. (a) Michelson type interferometric objective for observation of sample through cover glass with illumination provided not through the objective but from the side of the interferometer with a beam of a very low numerical aperture used to increase the contrast of the fringes, (b) cross hatch grating immersed in water measured with through the media objective with compensating cell/plate in the reference arm. Scan area $620 \times 460 \mu\text{m}$. Grating height: $30 \mu\text{m}$. Courtesy Reed and Gimzewski from UCLA.

Scanning white light interferometry (SWLI)

White light scanning interferometry (WLSI)

Coherence probe microscopy (CPM)

Correlation microscopy

Phase correlation microscopy

Interference microscope

Microscopic interferometry

Wide band interferometry

Full field OCT

Wide field OCT

Coherence radar

Fringe peak scanning interferometry

The equivalent method for biological samples is called mainly optical coherence tomography (OCT), but can also be called time domain OCT (TD-OCT), coherence radar or confocal interference microscope.

15.7. WAVELENGTH SCANNING INTERFEROMETER

An alternative to white light and multiple wavelength interferometers is the spectral interferometer, which takes advantage of spectral interference fringes for a wide range of wavelengths. These spectral fringes can be obtained through the scanning of the source wavelength or dispersing white light fringes with a spectrometer.

Spectral interferometry with wavelength scanning uses a setup that is typically based on a Michelson interferometer with the difference being a light source that is wavelength tunable (see Fig. 15.38). This system does not require point by point axial mechanical scanning as in typical confocal microscopy, or field axial scanning as in white light interferometry in order to find the best focal position for each point. Instead, fringes of different frequencies are observed by sweeping the light source through wavelengths from which the height of the object with respect to the reference mirror is determined. This wavelength scanning profiler delivers topographies of smooth and rough surfaces with no 2π -phase ambiguity problem. One advantage that a wavelength scanning system has over a white light setup is that the contrast of the fringes remains good even for dispersive media.

15.7.1. Wavelength Tunable Light Sources

Because the tuning range of the wavelength determines the resolution of the measurement and the tuning step of the wavelength determines the system's measurable depth (see Section 15.7.2), illuminating systems are continuously being developed to increase the scanning range and decrease the scanning step. At first, large expensive dye lasers, which were inconvenient for industrial purposes, and Ti:Sapphire lasers (Kuwamura and Yamaguchi, 1997; Yamamoto et al., 2001) were used. At the same time, new and much more convenient tunable solid

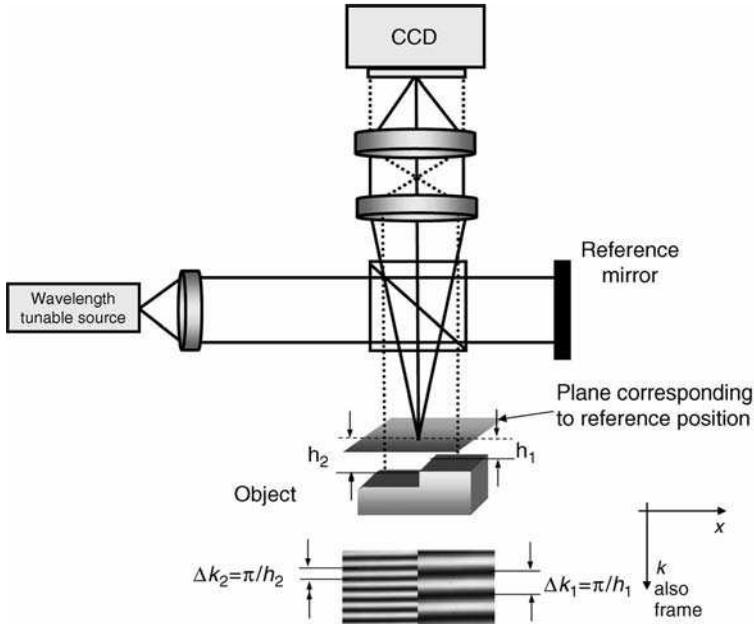


FIGURE 15.38. Schematic of spectral interferometer with wavelength scanning.

state lasers were also being used. A standard method of varying the wavelength in a diode laser is to change the injection current or chip temperature; however, these methods are subject to mode hops and changing of mode shape. To avoid these mode-hops, Tiziani et.al. (1997) used an external resonator like those used in dye lasers. Later, broadband sources like a superluminescent diode in combination with wavelength-tuning devices like an acousto-optical tunable filter or a liquid crystal Fabry–Perot interferometer (Mehta et al., 2002) were proposed as more convenient and stable illuminating systems.

15.7.2. Image Buildup

The interference signal at each x, y point can be described using Eq. (15.8), which is now

$$I(k, h) = I_R(k) + I_0(k) + [2I_R(k)I_0(k)]^{1/2} \cos(2kh) \tag{15.25}$$

where h is the path difference between corresponding points on the reference and the object. Since the object is not moving, it is assumed that $z = 0$. The variation in h across the field determines height variations of the object. As mentioned in Section 15.6.1 the phase difference $2kh$ can be changed in a few different ways. With mechanical axial scanning of the object by δh (previously marked as z) distance at a time (like in scanning white light or phase shifting interferometry), the phase difference for a given wave number k , for each point on the object, would change by

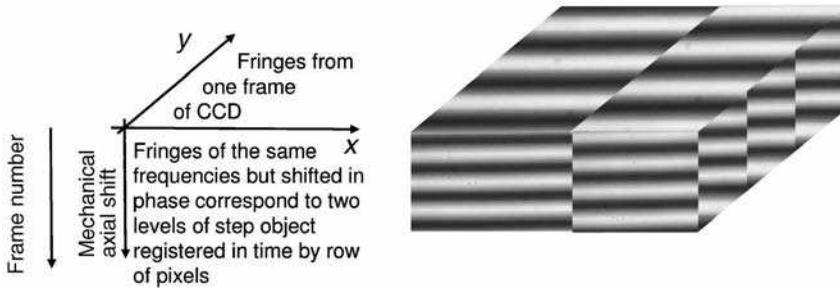


FIGURE 15.39. Fringes in monochromatic illumination for the step object as obtained with PZT scanning.

the same amount $\delta\varphi = 2k\delta h$ creating fringes of the same frequency for each point as shown in Eq. (15.26) and Figure 15.39.

$$I(k, h) = I_R(k) + I_0(k) + 2[I_R(k)I_0(k)]^{1/2} \cos(2k(h + \delta h)) \quad (15.26)$$

In wavelength scanning interferometry the phase change for each point on the object is caused by the wavelength change expressed as δk , and the change in the phase difference is expressed as $\delta\varphi = 2\delta kh$. The irradiance during the wavelength scan for each point can be expressed as in Eq. (15.27)

$$I(k, h) = I_R(k) + I_0(k) + 2[I_R(k)I_0(k)]^{1/2} \cos(2h(k + \delta k)) \quad (15.27)$$

And is called spectral interference. The variation in the optical path difference due to the shape of the object creates fringes of different frequencies for points across the object during the wavelength scan. Figures 15.39 and 15.40 show fringes as obtained for a step object that is tilted around the x axis with the phase change introduced by axial mechanical scanning and wavelength scanning, respectively. For a mechanical scan that uses monochromatic illumination, fringes of the same

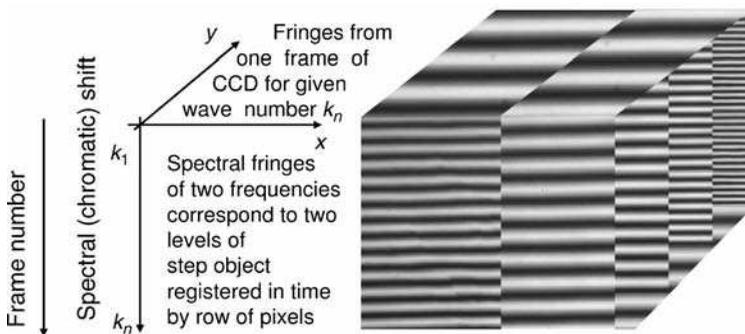


FIGURE 15.40. Fringes for the step object as obtained with wavelength scanning.

frequency but with a different initial phase for different step levels are created; for wavelength scanning two distinct frequencies for two step levels are created. Small variations in object shape result in small changes in the fringe frequency across the object.

The phase change at each point as a function of wavelength change equals

$$\delta\varphi = 2\delta kh = 4\pi h \frac{\delta\lambda}{\lambda^2} \quad (15.28)$$

The period of the fringes at each point can be expressed in terms of wavelength shift $\Delta\lambda$ for which the phase change $\Delta\varphi$ is equal to 2π

$$\Delta\varphi = 2\Delta kh = 2\pi \quad (15.29)$$

From the frequency of the detected fringes, at each point the height of the object h can be determined by Eq. (15.30)

$$h = \pi/\Delta k = \Delta\varphi/2\Delta k. \quad (15.30)$$

Finding the height from the fringe frequency is equivalent to finding the height from the rate of the phase change at each point, called phase slope or spectral phase.

The scanning wavelength range can be from a few nanometers to tens of nanometers and the tuning step of the wavelength is typically on the order of a fraction of nanometers (see Fig. 15.41). With this technique submicron resolution is possible with a height range of a few millimeters. The maximum measurable height is determined by the wavelength's tuning steps as

$$\Delta h_{max} = \pi/\delta k \frac{\lambda^2}{2\delta\lambda} \quad (15.31)$$

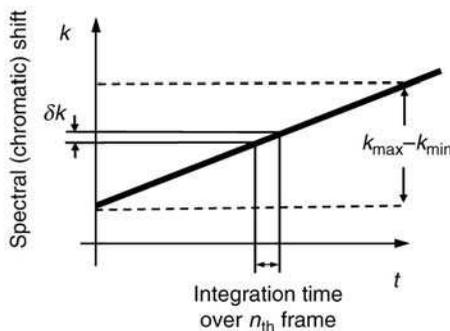


FIGURE 15.41. Wavelength tuning in wavelength scanning interferometer.

and the minimum measurable height difference is determined by the total tuning range as

$$\Delta h_{min} = \pi / (k_{max} - k_{min}) \quad (15.32)$$

However, other factors influence these values; for example, the depth of field of the objective may limit the maximum measurable depth of the object (see Section 15.4.1).

15.7.3. Signal Analysis

As explained in the previous section, information about the relative distance between the object and reference plane can be retrieved by finding the fringe frequency or phase change rate at a given point. At first, the fringe frequency was determined by counting the number of zero-crossings of the signal at each pixel. Later, a method based on the Fourier transform calculation was developed (Takeda and Yamamoto, 1994; Barnes et al., 1996; Yamaguchi et al., 1998). Figure 15.42 shows three interferograms for three different objects at one wavelength during a wavelength scan. Also shown is the signal as recorded by a pixel during a wavelength scan and the Fourier transform of the signal with the peak position determining the height of the object. To further improve the resolution of the Fourier transform technique, the rate of phase change (spectral phase slope) along the spectral interference can be calculated (Kinoshita et al., 1999; Mehta et al., 2003a; Mehta et al., 2003b). Wavelength scanning interferometry requires tens or hundreds of registered interferograms during the wavelength scanning of the source. The measurement time in wavelength scanning interferometry is the same for shallow and deep surfaces because the number of required frames does not depend on sample's depth like in white light interferometry.

In the Michelson interferometer object height can be, in relative terms, on both sides of the reference mirror. Because the methods described here solve only for the magnitude of the object's distance from the reference, not its direction, a height ambiguity problem exists. Figure 15.43 shows the height ambiguity that can be created. In these cases, a special phase unwrapping procedure must be applied (Paulson et al., 2000), or a phase unwrapping procedure known as temporal phase unwrapping (Huntley and Saldner, 1993, 1997; Saldner and Huntley, 1997a, 1997b, Huntley and Coggrave, 1998) must be applied to wavelength scanning interferometry. If the object were shifted above or below the focus plane, no unwrapping procedure would be necessary because the object would be clearly on one side of focus or the other. Shifting the object away from focus increases the frequency of fringes at each point as shown in Figure 15.44. However, this shifting could limit the maximum measurable height range because the object may be shifted out of the focus range or the fringe frequency may become too dense to be resolved (see Eq. (15.31)). The resolution of the measured height, in addition to the wavelength scanning range, also depends on the roughness of the tested object and the distance from the point of best focus (Yamaguchi et al., 1998).

In a Fizeau setup multiple reflections between the reference and the object must be accounted for in the signal analysis (Yamaguchi et al., 2000), and although the

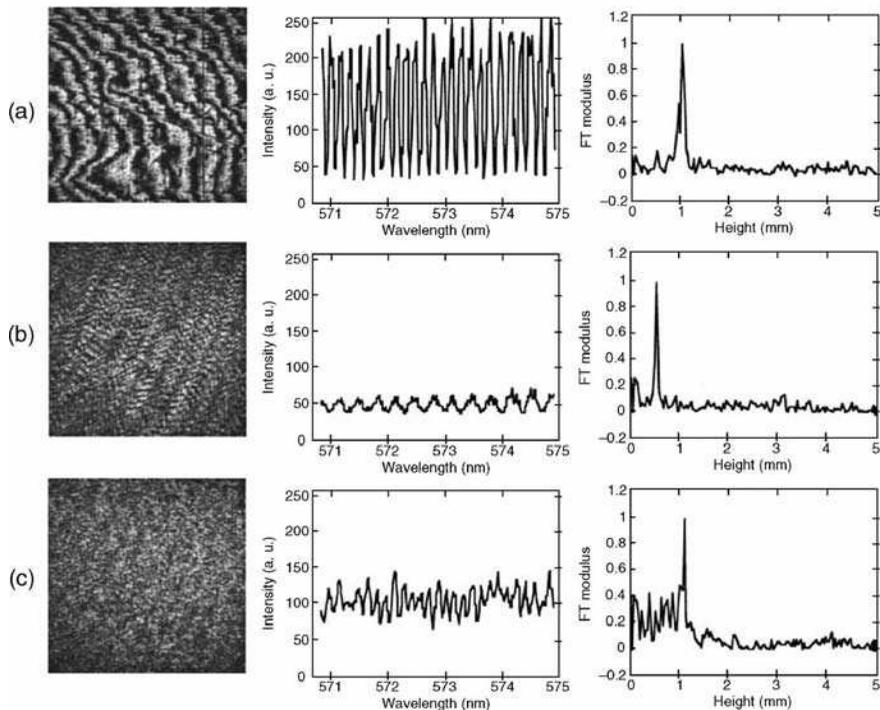


FIGURE 15.42. Interference patterns at a single wavelength, interference signals (spectral fringes) as seen by a single pixel during the wavelength scan in time, and the Fourier transform of the interference signal for (a) polished steel, (b) milled duralumin, and (c) a MgO coated surface. Reprinted with permission from Yamaguchi et al., 2000.

resultant signal is not sinusoidal as in a Michelson setup, the fundamental period of the signal is the same. In the Fizeau setup the object is on one side of the reference mirror and no ambiguity in height derivation exists.

Wavelength scanning is used in a number of optical methods such as optical frequency domain reflectometry, distance measurement (Kikuta et al., 1986), speckle interferometry and, recently, digital holography (Pawlowski et al., 2004). An overview of wavelength scanning interferometry and its applications was done by Tiziani (2000) and Takeda et al. (2005).

15.7.4. Film and Plate Thickness Measurement

A wavelength scanning interferometer can be used not only for shape measurement but also for thickness measurement (or its dispersion) of a transparent object. In this case due to reflections of the wavefront of multiple optical interfaces many interferograms are created and superimposed. Each interfering pair of reflected wavefronts of almost plane and parallel wavefronts creates spectral fringes of the main frequency characteristic for the average optical path difference during the wavelength scan. Some of these

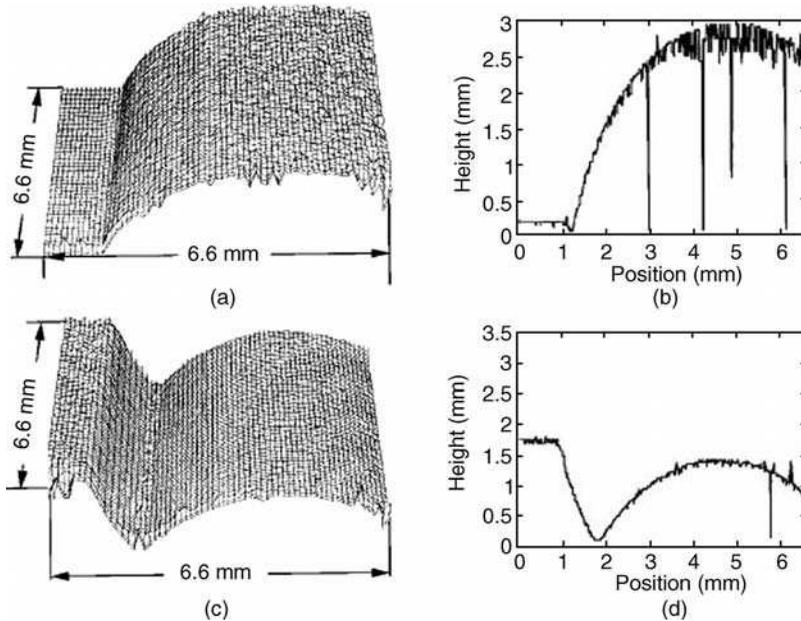


FIGURE 15.43. Wavelength scanning interferometry results for cylindrical surface positioned away from the virtual focus position of the reference mirror (a) surface, (b) plot and results showing possible sign ambiguity problem if the object surface crosses virtual focus position of the reference surface, (c) surface, and (d) plot. Reprinted with permission from Yamaguchi et al., 2000.

reflections are unwanted, and over the last 15 years methods based on separating these different interferograms in the data processing stage have been developed so as to avoid immersing the sample in oil or coating the surfaces with an index-matching lacquer (Okada et al., 1990; Deck, 2003; de Groot, 2000; Burke et al., 2006; Hibino et al., 2003; Hibino et al., 2004; Hibino and Takatsuji, 2002). However, these techniques could not be applied to measure thicknesses down to a few microns. Measurement of transparent layers of a few microns in thickness is often laid on

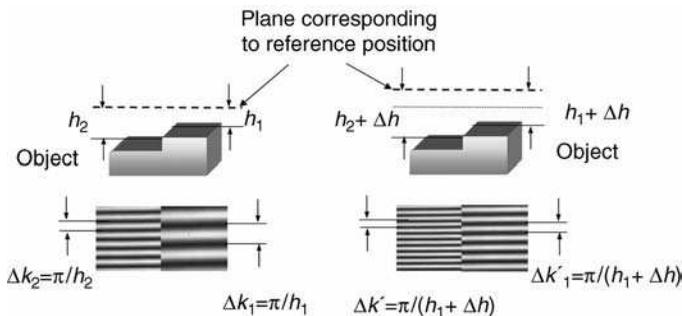


FIGURE 15.44. Fringes in wavelength scanning interferometer for two different positions of the object along the optical axis. Increase in distance between reference mirror and object results in increased fringe frequency.

optoelectronic devices, and at these small thicknesses separating the different frequencies was impossible until Kim (Kim et al., 2002; Kim and Kim, 2004) proposed a technique, similar to the one applied to WLI, to measure these tiny thicknesses.

15.8. SPECTRALLY RESOLVED WHITE LIGHT INTERFEROMETRY (SRWLI)

Spectral interference fringes for a wide range of wavelengths can be also observed by dispersing white light fringes with a spectrometer. This method of observation of spectrally resolved fringes is called spectrally resolved white light interferometry and is an alternative to wavelength scanning that was described in the previous section. The main advantage of SRWLI is that profiles with discontinuities up to tens of microns can be calculated from a single spectral interferogram.

Although the first spectrally resolved white light interferometer for surface profilometry described by Schwider and Zhou (Schwider and Zhou, 1994) was based on the Fizeau interferometer, these systems are typically based on a Michelson system. In order to observe spectrally resolved fringes, these interferometers use a spectrometer in conjunction with the CCD camera. The plane of fringe localization is imaged through the entrance slit of spectrometer onto the CCD camera as shown in Figure 15.45. The spectrometer splits the single line of the object's white light fringes into spectral fringes along the chromatic axis of the CCD camera. Different types of spectrometers can be used, such as a prism or grating spectrometer.

The observed fringes are basically the "channeled spectrum" that scientists observed about 100 years ago. In this respect spectrally resolved interferometry is not a new technique; however, the development of computers and solid state devices and spectrometers has allowed for the utilization of the channel spectrum for many purposes. Spectrally resolved white light fringes are used for many different applications like measuring the differential index of refraction, the index of refraction distribution, wavelength multiplexing, transmission of images, distance and displacement measurements and recently for profile measurement.

15.8.1. Image Buildup

In the spectrally resolved white light interferometer the chromatic axis is along one of the axes of CCD camera which is perpendicular to the slit of spectrometer. The row (or column) of CCD pixels registers a spectral interference signal from which the spectral phase or fringe frequency is calculated for a single object's point delivering information about the object's distance from the corresponding reference point, and one CCD frame delivers information about the profile of the object along one line as shown in Figure 15.46. Thus, the object needs to be scanned laterally in order to obtain a 3D profile.

A spectrally resolved white light interferometer with lateral scanning delivers the same type of fringes as a wavelength scanning interferometer. We can think of the collected data as a cube (compare Figs. 15.40 and 15.47). The difference lies in the

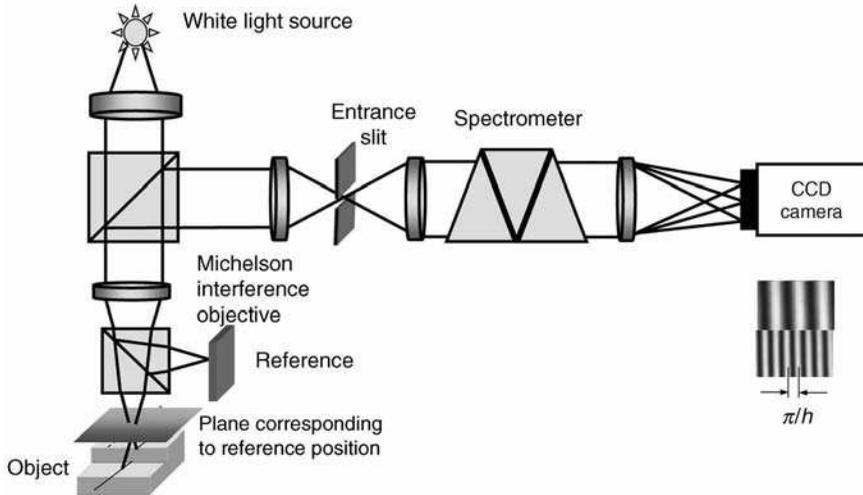


FIGURE 15.45. Schematic of a spectrally resolved white light interferometer.

axis labeling of both cubes of data. In wavelength scanning interferometry (Fig. 15.40) one CCD frame delivers a monochromatic interferogram in spatial coordinates x, y . The third coordinate, the chromatic axis, follows interferograms that are registered in time during the scan of the source's wavelength. In spectrally resolved WLI, one CCD frame delivers spectral fringes in chromatic-spatial coordinates. The third coordinate, the spatial axis, follows interferograms registered in time during the lateral scan.

Like in wavelength scanning interferometry, the dynamic range for the measurement is limited by the spectral resolution of the spectrometer (equivalent to the tuning step in wavelength scanning interferometry), and the resolution is determined by the spectral bandwidth (tuning range) of the light source.

The chromatic axis of the CCD camera is calibrated typically with a cadmium (Cd) spectral lamp, which has four spectral lines, 643.8 nm, 508.6 nm, 480 nm, and 467.8 nm. These wavelengths are assigned to the corresponding pixels on the camera (see Fig. 15.48(a)). The wavelengths for the rest of the pixels are calibrated (see Fig. 15.48(b)) using Hartmann's formula.

15.8.2. Signal Analysis

In order to determine the distance of an object with respect to the reference plane, the spectral interferogram can be analyzed by measuring either the frequency of the fringes (as shown for wavelength scanning interferometry) or the spectral phase slope (Schwider and Zhou, 1994). Both methods require the carrier frequency of the fringes, which can be regulated by the distance of the object from the reference mirror. Spectral phase slope $\Delta\phi/2\Delta k$ can be measured using the Fourier Transform (Takeda and Yamamoto, 1994) or a spatial phase-shifting method (Sandoz et. al. 1996).

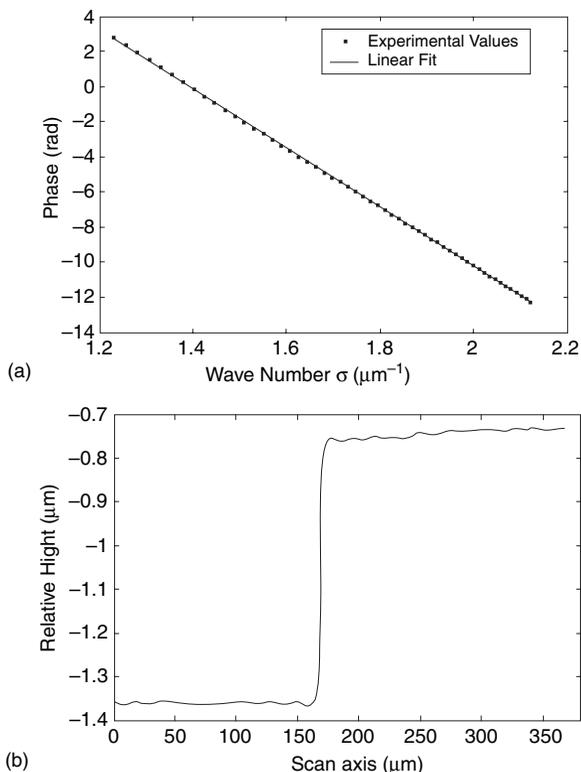


FIGURE 15.46. Obtained from spectrally resolved white light interferometer. (a) spectral phase along chromatic axis for one point on object and (b) profile of the step.

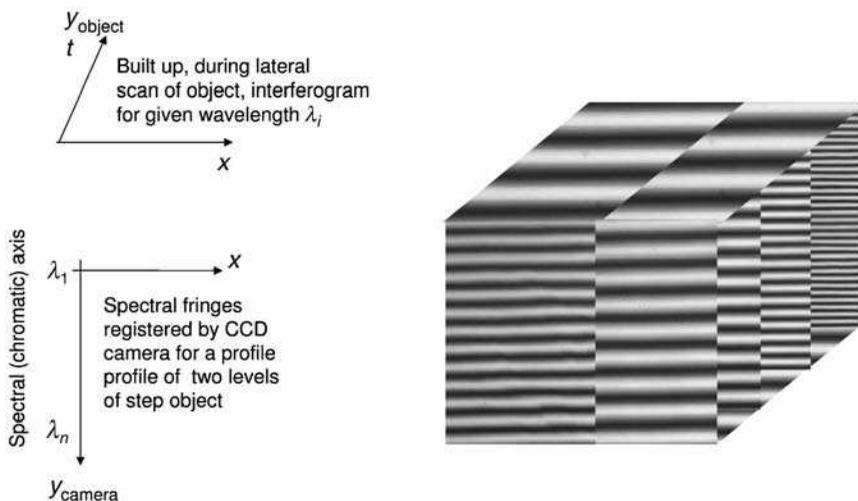


FIGURE 15.47. Cube of data in spectrally resolved interferometer for an object with step tilted with respect to the reference mirror around axis perpendicular to the step.

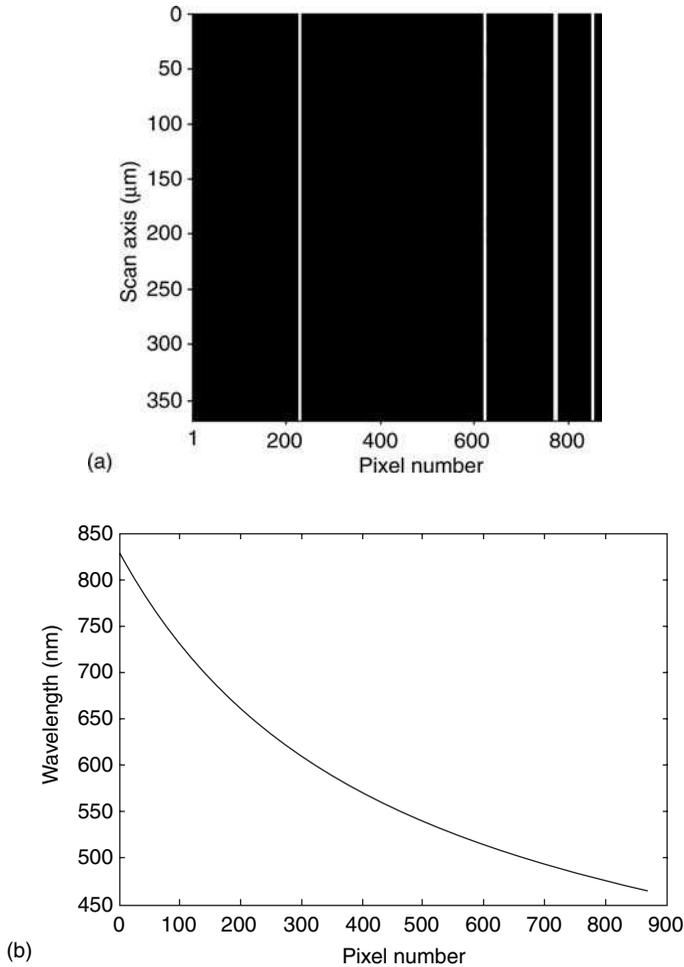


FIGURE 15.48. Calibration of chromatic axis of camera: (a) imaged spectrum of Cd lamp on the CCD and (b) calculated wavelength for each pixel. Reprinted with permission from Debnath and Kothiyal (2005).

Although carrier methods require only a single frame, they do not distinguish the sign of the spectral phase slope. Like in the wavelength scanning interferometer, the sign of the spectral phase slope determines where the object is in relation to the reference mirror, either above it or below it, and this information is lost. In addition, carrier methods may be affected by the background irradiance variations of the object across chromatic axis.

In a spectrally resolved white light interferometer, PSI methods have been implemented to increase the resolution of the measured spectral phase and, thus, also the height, and at the same time to overcome phase sign difficulties. PSI methods deliver information about the spectral phase slope, do not require fringes

of carrier frequency, and are insensitive to spatial variations in background irradiance (Helen et al., 2001; Debnath and Kothiyal, 2005). With this method, points on the object in relative terms can be on either side and very close to the reference mirror.

Thick and Thin Film Measurement. Spectrally resolved interferometry also can be used to measure film thicknesses as thin as 2 microns by mapping the peaks of the Fourier transform (Häusler and Lindner, 1998; Zuluaga and Richards-Kortum, 1999). The problems in this method, like in wavelength scanning interferometry, are ghost frequency peaks coming from multiple reflections of many interfaces and the inability of methods to distinguish the interface of the object as being above or below the reference mirror. Wojtkowski et al. (2002) suggested that using a PSI approach the position of the optical film interfaces can be determined without ambiguity. Recently Debnath et al. (2006) showed that the spectral phase calculated with a PSI method can measure engineering samples that have film thicknesses as thin as 100 nm. A spectrally resolved WLI approach allows for the simultaneous registration of the sample profile and film thickness.

15.8.3. Other Names for Spectral Interferometry

Spectrally resolved white light interferometry and wavelength scanning interferometry are often called spectral interferometry. Each of these methods also has been variously labeled. Spectrally resolved white light interferometry can be found in the literature as dispersive (white light) interferometry, white light channeled spectrum interferometry, or spatially resolved spectral interferometry. The wavelength scanning interferometer can also be called spectrally scanned interferometer, frequency scanning interferometer, wavelength tuning interferometer and optical frequency domain microprofilometer.

Variations of spectral interferometry are employed in the biomedical field; the term used there is frequency domain optical coherence tomography (FD-OCT). Whereas typical OCT is a point method and requires lateral scanning, FD-OCT works in the frequency domain, thus avoiding axial mechanical scanning. In the OCT field, both methods are commonly called frequency (or Fourier) domain optical coherence tomography (FD-OCT) or sometimes spectral radar.

15.9. POLARIZATION INTERFEROMETERS

In commercial interferometers the beam splitting is commonly done by amplitude splitting like in the Twyman–Green, Michelson or Fizeau interferometers; these systems were described in Section 15.4.1. Beam splitting can also be accomplished through the use of a polarizing beam splitter when different polarization states in the reference and object beams are required. These setups are called polarization interferometers and often utilize polarization techniques to introduce phase shifting between interfering wavefronts in order to avoid mechanical axial or wavelength

depended phase shift. Such systems were described by Hettwer et al. (2000). Polarization technique for phase shifting will be described in Section 15.9.2.

Different type of polarization interferometers use polarizing beam splitter to split the incident beam into two sheared beams incident on the object; this technique creates a quasi-common path interferometer. Although the idea of the polarization interferometer was known at the turn of 20th century, these systems became popular when they were inserted into a microscope to create the differential interference contrast microscope (DIC). An excellent review of different types of polarization interferometers can be found in a few books (Françon and Mallick, 1971; Pluta, 1993; Polarvarapu, 1997) and also in an article by Françon (1963). In addition, the previous edition of this book (Creath, 1992) details two other types of polarization interferometers, Sommargren (Sommargren, 1981) and Downs (Downs et al., 1985).

While imaging small features, on the order of illuminating wavelength and well below, in polarization interferometer it is important to understand how the reflected or transmitted wavefront is affected by illuminating beam state of polarization in order to interpret results correctly (Totzeck et al., 2005). To improve optical resolution of polarization interferometers, structure pupil filters are suggested to be used (Totzeck et al., 2002).

15.9.1. Differential Interference Contrast Microscope (Nomarski)

The DIC interferometer was invented, a Polish-born French physicist, Georges (Jerzy) Nomarski in 1955 (Nomarski, 1955) and is also called a polarization interference contrast microscope. In the Nomarski system, a typically broad source (like a halogen lamp) is used for illumination followed by a polarizer (see Fig. 15.49). This polarized light travels through the heart of the DIC microscope, which is a modified Wollaston birefringent prism pair – Nomarski prism. The Wollaston prism splits the polarized light into two orthogonally polarized beams that travel at slightly different angles. Next, both beams are brought onto the object by the objective, and the object is illuminated with orthogonally polarized beams slightly sheared from each other. For this reason this interferometer is also called a shearing interferometer. This shear is on the order of tens to hundreds of nanometers, well below the resolution of the objective. Upon reflection two displaced beams travel back through the objective and are recombined by the Wollaston prism. To allow for observation of interference between both beams, an analyzer is placed before the beams reach the camera.

The interference pattern does not directly represent the shape of the image but rather its gradient in the direction of the shear introduced by the Wollaston prism. Thus, to fully characterize the object, an additional measurement at orthogonal shear needs to be taken, which is achieved by sample or optics rotation. The observed interference colors are phenomenal and give an effect of a pseudo three-dimensional appearance of the object. These colors can be varied by changing the amount of shear between beams via an axial shift of the Wollaston prism.

Initially, all DIC systems were manual and of a qualitative nature only. The user manually chooses a Wollaston prism axial position so as to achieve an interference

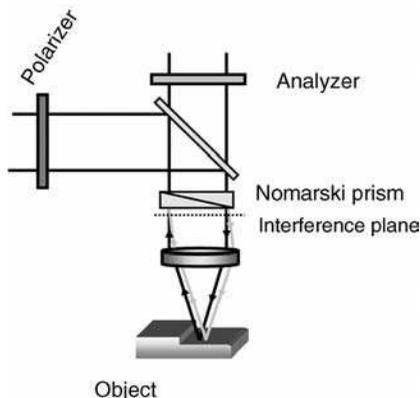


FIGURE 15.49. Differential interference microscope in reflection mode.

pattern that would be the most appealing and emphasize the features of the object that were under test. The most common type of DIC interferometer is used for qualitative observation of semitransparent biological samples; this system works in transmission and not on reflection. Although neither as popular nor as commercially available as white light interference microscopes, a DIC interferometer can also deliver quantitative data about a surface profile while working in reflective mode. Some idea about the optical path difference between sheared wavefronts can be achieved by looking at created interference colors and comparing them to Michel-Lévy and Lacroix who came up in 1889 with Interference Color Chart. A more reliable technique, one that does not require color comparison and can be used with a black and white CCD camera, is based on phase shifting interferometry (see chapter on Phase Shifting Interferometry and Arnison et al., 2004). The phase shift between interfering beams can be obtained by shifting the Wollaston prism laterally or by rotating the polarizer. However, the phase shift will be wavelength dependent, and the fringes will change contrast during the phase shift. To avoid this changing contrast, achromatic phase shifting based on the geometric phase can be used (Hariharan, 1996). Once the optical path difference between sheared replica of the wavefront reflected from the object is obtained in two orthogonal directions, the original surface can be calculated. The advantage of the DIC interferometer is that it is almost insensitive to vibrations since they are almost common path interferometers. The major limitations of these interferometers, which include a short measurable depth and thickness of the object and index gradient, occur because of the limited depth of field of the objective like in any other non-scanning interference microscope. On the contrary, a very small depth of field for objectives with high NA can be utilized for optical sectioning if axial scan between object and objective is introduced. Reflective DIC are often used for quality inspection of integrated circuits in semiconductor industry. A lot of useful information about DIC interferometers working in reflection can be found on websites designed by for example Nikon and Olympus.

15.9.2. Geometric Phase Shifting

Many measurement techniques in interferometry involve shifting the phase of the interfering wavefronts. Mechanical phase shifters, when used in white-light and multiple-wavelength interferometry, introduce the same change in the optical path difference, measured in nanometers, for all wavelengths; however, the resulting phase shift, known as a dynamic phase shift, when measured in degrees or radians, varies with the wavelength. A different technique of phase shifting, involving a cycle of changes in the polarization of the light, can produce the same phase shift, measured in degrees or radians, for all wavelengths. This phase shift, known as the Pancharatnam phase (Pancharatnam, 1956), is a manifestation of the geometric phase (Berry, 1987), and it can generate any required wavelength-independent phase shift without a change in the optical path difference. As a result, geometric phase-shifting has found many applications in interferometry (Hariharan, 2005).

In white-light interference, a change in the geometric phase shifts the fringes under the coherence envelope (Hariharan et al., 1994), as shown in Figure 15.50, while the coherence envelope stays in place, resulting in no change in the fringe contrast at each point. In dynamic phase shifting, the whole white-light interferogram is shifted, resulting in changing fringe contrast at each point. With multiwavelength interferometry, a geometric phase-shifter will produce the same phase shifts for any wavelength used in the interferometer, without the need to make any changes in the phase shifter.

Achromatic phase-shifters operating on the geometric phase use circularly polarized light, as shown in Figure 15.51, and employ polarization elements such as a rotating half-wave plate followed by a quarter-wave plate and polarizer or, in a simpler arrangement, just a rotating polarizer. In order to introduce a phase shift between two interfering beams, the two beams need to be orthogonally polarized.

Figure 15.52 shows an interferometer with two orthogonally linearly polarized beams leaving the reference and object arms of the interferometer. The geometric

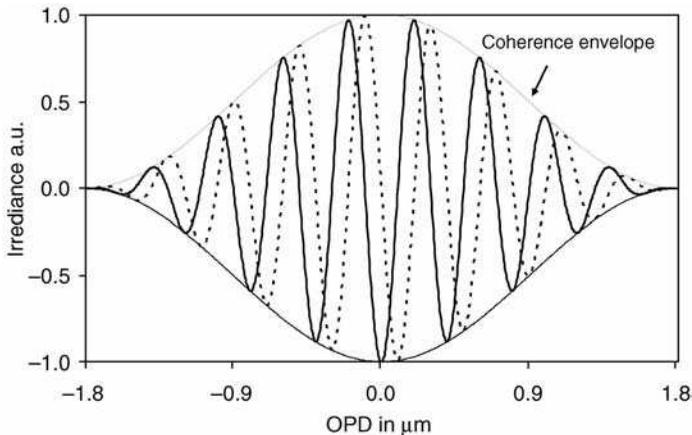


FIGURE 15.50. During achromatic phase shift only fringes shift, while envelope remains stationary.

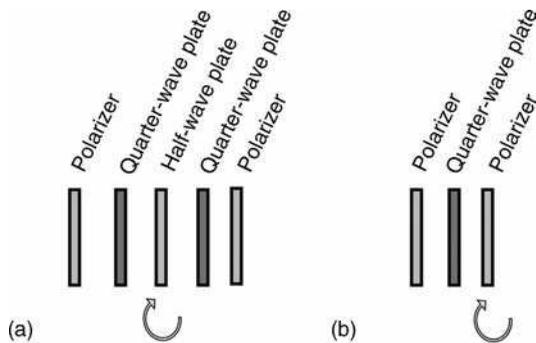


FIGURE 15.51. Geometric phase shifter with rotating (a) wave plate and (b) polarizer.

phase shifter, consisting of a rotating half-wave plate mounted between two quarter-wave plates with their axes set at 45° to the angles of polarization of the two beams (Hariharan and Ciddor, 1994; Hariharan et al., 1994), is placed at the exit of the interferometer. This interferometer employs the first type of geometric phase shifter shown in Figure 15.51a in which the first quarter-wave plate creates left- and right-handed circularly polarized beams. The half-wave plate then changes the right-handed circularly polarized beam to a left-handed one and the left-handed circularly polarized beam to a right-handed one. Finally, the second quarter-wave plate brings the two beams back to their original orthogonal linear polarizations. A rotation α of the half-wave plate shifts the phase of one linearly polarized beam by $+2\alpha$ and the phase of the other, orthogonally polarized beam by -2α , so that a net phase

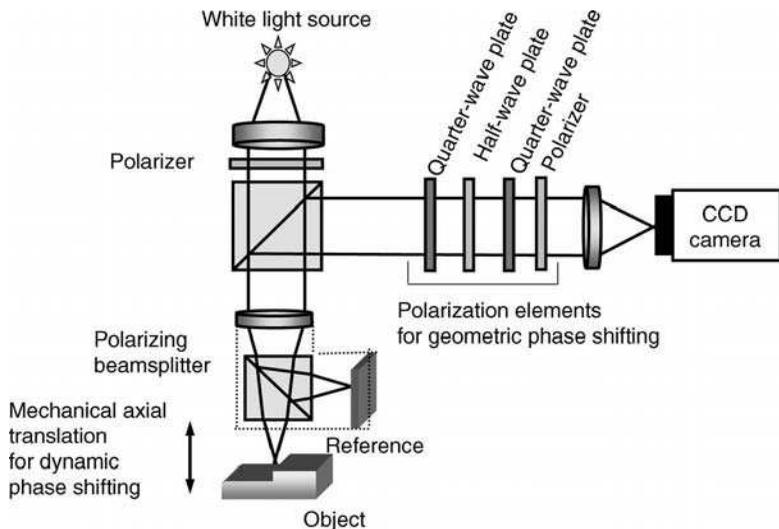


FIGURE 15.52. White light interference microscope with geometric phase shifting. At each position of the axial scan, approximately every 150 nm, geometrical phase shifting is introduced, and a few frames are collected to calculate the contrast of fringes.

difference of 4α is introduced between the two beams. This phase difference is very nearly independent of the wavelength over the whole visible spectrum. The polarizer makes it possible for the two beams to interfere.

The second type of geometric phase shifter shown in Figure 15.51b only requires a rotating polarizer to be placed after the quarter-wave plate that changes the two beams leaving the interferometer to left- and right-circularly polarized beams. In this case, if the test beam is left-circularly polarized and the reference beam is right-circularly polarized and both beams are incident upon the linear polarizer which is set at an angle α with respect to the x axis, both the test and reference beams, upon passing through the polarizer, become linearly polarized at an angle α . However, a phase offset $+\alpha$ is added to the test beam and a phase offset $-\alpha$ is added to the reference beam. A rotation of the linear polarizer by α therefore introduces a phase shift 2α between the two interfering beams. The linear polarizer acts as a phase shifting device and also makes it possible for these beams to interfere. While an achromatic quarter-wave plate could be used to extend the spectral range over which this phase-shifter operates, it turns out that the variations in the phase shift produced by this system due to variations in the retardation of the quarter-wave plate with the wavelength are quite small (Helen et al., 1998; Millerd et al., 2004).

The measurement time, which can be critical in some industrial applications, can be reduced if the interferograms are collected simultaneously. This can be done with the system presented in Figure 15.53 (Millerd et al., 2004). In this setup, a polarizing beamsplitter causes the reference and test beams to have orthogonal polarizations. Quarter-wave plates are placed in the reference and test beams so that the beam initially transmitted through the beamsplitter is reflected when it returns, and vice versa. These two beams pass through a quarter-wave plate, which converts the two orthogonally polarized beams to right- and left-handed circularly polarized beams, and then through a phase mask. The quarter-wave plate can be placed at the exit of the interferometer, or in front of the camera, while the phase mask is placed just in front of the CCD array in the camera.

The phase mask is a micropolarizer array built up of groups of four linear polarizer elements having their transmission axes at 0° , 45° , 90° , and -45° as shown in

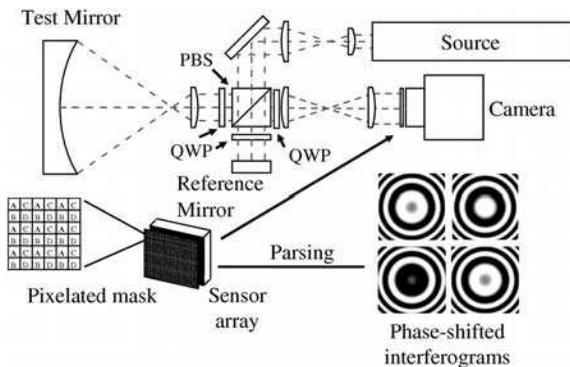


FIGURE 15.53. Simultaneous phase-shifting interferometer. (Courtesy 4D Technology).

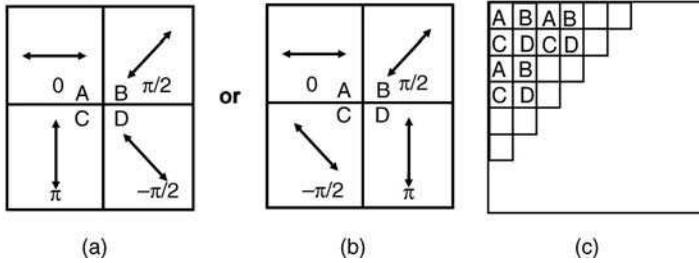


FIGURE 15.54. Phase filter. (a) 4 polarizer elements giving 0° , 90° , 180° , and -90° phase shifts; (b) 4 polarizer elements giving 0° , 90° , -90° , and 180° phase shifts, (c) phase filter made up of array of 4 polarizer elements. Courtesy 4D Technology.

Figure 15.54(a), or at 0° , 45° , -45° and 90° as shown in Figure 15.54(b), and is structured so that each polarizer element is placed over a detector element. These four linear polarizer elements introduce phase shifts between the test and reference beams of 0° , 90° , 180° , and 270° . Thus, four phase-shifted interferograms, obtained from each group of pixels (Fig. 15.54(c)), are recorded simultaneously using a single CCD array.

As can be seen, the phase mask works as a geometric phase shifter. The two essential requirements are that the test and reference beams have orthogonal polarizations, and the micropolarizer array matches the CCD array.

15.10. OPTICAL RANGING METHODS

Optical ranging methods are typically single point methods used to measure for example length, distance, vibrations or index of refraction. They can be based on different principles; for example on geometry of light propagation used in laser triangulations or based on interferometric principles or on time measurement of light pulse flight. If scanning is employed all these methods can provide 3D measurement of the object. Single point methods are usually based on optical fibers and fast single point detectors. Some methods can and were made to be used with array detectors delivering information about ranging parallelly at multiple points to deliver information over the area without necessity of point by point scanning. Review of different methods for optical ranging was presented by Amann et al. (2001), Blais (2004), Chen et al. (2000), de Groot (2003), Friedman (2003), and Wagner and Hausler (2003). Selected papers on laser distance measurements were collected by Bosch and Lescure (1995).

15.10.1. Interferometric Ranging

Interferometric ranging is based on white light, two-wavelength, multiple wavelength, and spectral interferometer (Gerges et al., 1987; Smith and Dobson, 1989; Danielson and Boisrobert, 1991; de Groot and McGarvey, 1992; de Groot, 1992,

Haruna et al., 1998; Bosch et al., 2001; Hariharan, 2003). Multiple wavelength interferometry was already used in 1895 to measure the length of an etalon (Michelson and Benoit, 1895) to overcome limitations of single wavelength interferometry. These methods evolved successfully into methods for very precise 3D profilometry described in this chapter.

15.10.2. Optical Triangulation

Triangulation sensors detect the backscattered light from a narrow laser beam impinging on an object (Dorsch et al., 1994). The reflected light is detected by a position sensitive device. From the change in position of the spot on the sensor the change in distance to the object can be determined. Laser scanning and slit scanning are based on optical triangulation. The principles of optical triangulation can be used in tactile profilometers like stylus profilometers and atomic force microscopes to detect the position of scanning probe.

15.10.3. Time of Flight (TOF)

These systems use light propagation as a measuring tool since the speed of light is one of the fundamental well-determined constants. Distance is measured by counting the time it takes for light to travel to an object and back. Time-of-flight method is a preferred method to measure long distances; this method was used to measure the distance to the moon. With the addition of a scanning system, aerial topographies can be obtained. The resolution of TOF measurements ranges from 0.3 mm to a few centimeters or more. A good review of TOF systems was done by Blais (2004) and Moring et al. (1989)

15.11. SUMMARY

All of the methods described in this chapter allow for the measurement of topographies through various contact (tactile) and noncontact (optical) means. Oftentimes the needs of industry have pushed the development of methods that can extract particular information from a specific kind of sample in a unique environment. Testing MEMS devices, for example, with their irregular surfaces often present unique challenges, such as measuring laterally and vertically moving and deforming surfaces, where these measurements occur through a piece of glass that protects the sample. In addition, the need to reduce measurement times is always a priority.

Tactile methods have striven to develop long, narrow and smaller radius tips, lighter forces and new ways of tip motion so that a wider range of sample can be measured with higher precision and without damaging the sample. The interaction of AFM or SPM tips delivers an increasingly wider range of information about the electrical, magnetic and mechanical properties of different materials. However, since these techniques are point by point scanning techniques, decreasing measurement time is always crucial.

Optical methods continue to improve vertical resolution over a large range of height measurements. Recent trends show the development of techniques to measure more complex samples, those built of different materials or coated with transparent layers. Algorithms have been developed that obtain information about film thickness and profile measurement and correction due to phase change on reflection of different materials. In addition, new optical methods that are insensitive to vibrations are being developed to examine the topographies of very large objects.

We foresee that the current trends in developing new methods will continue to increase vertical and lateral range and resolution. Moreover, methods and systems continue to be modified to increase their application space. For example, foreseeable new applications include ways to measure samples submerged in liquids, in situ monitoring and measurement of difficult to reach areas, such as small millimeter-sized holes, are presently being developed. Finally, we envision a larger overlap with optical coherence tomography and confocal microscopy for measurement of biological specimens.

ACKNOWLEDGEMENT

Thank you Matthew Schmit for working on many of the figures found in this chapter.

REFERENCES

- Abdulhalim I., "Theory for Double Beam Interference Microscopes With Coherence Effects and Verification Using the Linnik Microscope," *J. Mod. Opt.*, **48**, 279–302 (2001).
- Aguilar J. F. and E. R. Méndez, "On the Limitations of the Confocal Scanning Optical Microscope as Profilometer," *J. Mod. Opt.*, **42**, 1785–1794 (1995).
- Al-Jumaily G. A., S. R. Wilson, K. C. Jungling, J. R. Mcneil, and J. M. Bennett, "Frequency Response Characteristics of a Mechanical Stylus Profilometer," *Opt. Eng.*, **26**, 953–958 (1987).
- Alexander S., L. Hellemans, O. M. J. Schneir, V. Elings, P. K. Hansma, M. Longmire, and J. Gurley, "An Atomic-Resolution Atomic-Force Microscope Implemented Using an Optical Lever," *Applied Phys.*, **65**, 164–167 (1989).
- Amann M.-C., T. Bosch, M. Lescure, R. Myllylä, and M. Rioux, "Laser Ranging: A Critical Review of Usual Techniques for Distance Measurement," *Opt. Eng.*, **40**, 10–19 (2001).
- Arnison M. R., K. G. Larkin, C. J. R. Sheppard, N. I. Smith, and C. J. Cosgwell, "Linear Phase Imaging Using Differential Interference Contrast Microscopy," *J. Microscopy*, **214**, 7–12 (2004).
- Balsubramanian N., *Optical System for Surface Topography Measurement*, US Patent **4**, 340,306 (1982).
- Barnes T. H., T. Eiju, and K. Matsuda, "Rough Surface Profile Measurement Using Speckle Optical Frequency Domain Reflectometry With External Cavity Tunable Diode Laser," *Optik*, **103**, 93–100 (1996).
- Bartoli A., P. Poggi, F. Quercioli, B. Tiribilli, and M. Vassalli, "Optical Profilometer with a Standalone Scanning Sensor Head," *Opt. Eng.*, **40**, 2852–2859 (2001).

- Benford J. R., "Microscope Objectives," in: R. Kingslake (Ed.), *Applied Optics and Optical Engineering*, vol. III, Academic Press, New York, 1966, pp. 145–182.
- Bennett J. M., "Precise Method for Measuring the Absolute Phase Change Upon Reflection," *J. Opt. Soc. Amer.*, **54**, 612–624 (1964).
- Bennett J. M. and J. H. Dancy, "Stylus Profiling Instrument for Measuring Statistical Properties of Smooth Optical Surfaces," *Appl. Opt.* **20**, 1785–1802 (1981).
- Bennett J. M. and L. Mattson, *Introduction to Surface Roughness and Scattering*, Optical Society of America, Washington, D.C., 1989.
- Bennett J. M., V. Elings, and K. Kjoller, "Precision Metrology for Studying Optical Surfaces," *Opt. Photonics News*, **2**, 14–18 (1991).
- Benoit M. R., "Application Des Phénomènes D'interférence a Des Déterminations Métrologiques," *J. Phys.(Paris)*, **3**, 57–68 (1898).
- Berry M.V., "The Adiabatic Phase and Pancharatnem's Phase for Polarized Light," *J. Mod. Opt.*, **34**, 1401–1407 (1987).
- B. Jena, J. K. Horber, L. Wilson and P. Matsudaira (Eds.), *Atomic Force Microscopy in Cell Biology*, vol. **68**, Academic Press, San Diego, 2002.
- Bhushan B., H. Fuchs, S. Hosaka, *Applied Scanning Probe Methods*, Springer, Berlin, 2004.
- Biegen B. and R. Smythe, "High Resolution Phase Measuring Laser Interferometric Microscope for Engineering Surface Metrology," *Proc. SPIE*, **1009**, 35–44 (1988).
- Binnig G. and H. Rohrer, "Scanning Tunneling Microscopy," *Helv. Phys. Acta*, **55**, 726–735 (1982).
- Binnig G. and H. Rohrer, "The Scanning Tunneling Microscopy," *Sci. Amer.* (August, 1985).
- Binnig G., C. F. Quate, and C. Gerber, "Atomic Force Microscope," *Phys. Rev. Lett.*, **56**, 930–933 (1986).
- Blais F., "Review of 20 Years of Range Sensors Development," *Journal of Electronic Imaging*, **13**, 231–240 (2004).
- Bonnell D. A. (Ed.), *Scanning Probe Microscopy and Spectroscopy: Theory, Techniques, and Applications*, 2nd ed., Wiley-VCH, New York 2000.
- Born M. and E. Wolf, *Principles of Optics: Electromagnetic Theory of Propagation, Interference and Diffraction of Light*, 7th ed., University Press, Cambridge, 1999.
- Bosch T. and M. Lescure (Eds.), *Laser Distance Measurements*, vol. MS 115, SPIE Optical Engineering Press, Bellingham, 1995.
- Bosch T., N. Servagent, and S. Donati, "Optical Feedback Interferometry for Sensing Application," *Opt. Eng.*, **40**, 0–27 (2001).
- Braga P. C. and D. Ricci, *Atomic Force Microscopy: Biomedical Methods and Applications*, Humana Press, Totowa, NY, 2004.
- Breitmeier U. and R. J. Ahlers, "Dynamically Focusing Electro-Optical Sensor System for Microprofilometry," *Proc. SPIE*, **802**, 170–173 (1987).
- Brodman R. and W. Smilga, "Evaluation of a Commercial Microtopography Sensor," *Proc. SPIE*, **802**, 165–169 (1987).
- Browne M. A., O. Akinyemi, and A. Boyde, "Confocal Surface Profiling Utilizing Chromatic Aberration," *Scanning*, **14**, 145–153 (1992).
- Bruce C. F. and B. S. Thornton, "Obliquity Effects in Interference Microscopes," *J. Sci. Instrum.*, **34**, 203–204 (1956).

- Burke J., K. Hibino, R. Hanayama and B. F. Oreb., "Simultaneous Measurement of Several Near-Parallel Surfaces With Wavelength-Shifting Interferometry and a Tunable Phase-Shifting Method," *Optics Lasers Eng.* **45**, 326-341 (2007).
- Caber P. J., "Interferometric Profiler for Rough Surfaces," *Appl. Opt.*, **32**, 3438-3441 (1993).
- Carré P., "Installation Et Utilisation Du Comparsateur Photoelectrique Et Interferentiel Du Bureau International De Poids Et Mesures," *Metrologia*, **2**, 13-23 (1966).
- Cha S. P., C. Lin, L. Zhu, P. Sun, and Y. Fainman, "Nontranslational Three-Dimensional Profilometry by Chromatic Confocal Microscopy with Dynamically Configurable Micro-mirror Scanning," *Appl. Opt.*, **39**, 2605-2613 (2000).
- Chen F., G. M. Brown, and M. Song, "Overview of Three-Dimensional Shape Measurement Using Optical Methods," *Opt. Eng.*, **39**, 10-22 (2000).
- Cheng Y. Y. and J. C. Wyant, "Two-Wavelength Phase Shifting Interferometry," *Appl. Opt.*, **23**, 4539-4543 (1984).
- Cheng Y. Y. and J. C. Wyant, "Multiple-Wavelength Phase-Shifting Interferometry," *Appl. Opt.*, **24**, 804-807 (1985).
- Church E. L. and S. R. Lange, "Structure Effects in Optical Surface Metrology," *Proc. SPIE*, **680**, 124-140 (1986).
- Cleveland J. P., B. Anczykowski, A. E. Schmid, and V. B. Elings, "Energy Dissipation in Tapping Mode Atomic Force Microscopy," *App. Phys. Lett.*, **72**, 2613-2615 (1998).
- Cohen D. K., W. H. Gee, M. Ludeke, and J. Leikowics, "Automatic Focus Control: The Astigmatic Lens Approach," *App. Opt.* **23**, 565-570 (1984).
- Cohen D., P. J. Caber, and C. Brophy, *Rough Surface Profiler and Method*, U.S. Patent 5,133,601 (filed June, 1991; issued July, 1992).
- Cohen S. H. and M. L. Lightbody (Eds.), *Atomic Force Microscopy/Scanning Tunneling Microscopy 2*, Springer, New York 1999.
- Corle T. R., C. H. Chou, and G. S. Kino, "Depth Response of Confocal Optical Microscopes," *Opt. Lett.*, **11**, 770-772 (1986).
- Corle T. R. and G. K. Kino, *Confocal Scanning Optical Microscopy and Related Imaging Systems*, Academic Press, New York, 1996.
- Creath K., "Measuring Step Heights Using an Optical Profiler," *Proc. SPIE*, **661**, 296-301 (1986).
- Creath K., "Calibration of Numerical Aperture Effects in Interferometric Microscope Objectives," *Appl. Opt.*, **28**, 3333-3338 (1989).
- Creath K., "Sampling Requirements for White Light Interferometry," *Fringe'97 Proceedings of the 3rd International Workshop on Automatic Processing of fringe patterns* 52-59 (1997).
- Creath K. and A. Morales, "Contact and Non contact Profilers," in D. Malacara, Ed., *Optical Shop Testing* Wiley, New York, 1992.
- Creath K. and J. C. Wyant, "Direct Phase Measurement of Aspheric Surface Contours," *Proc. SPIE*, **645**, 101-106 (1986a).
- Creath K. and J. C. Wyant, "From Angstroms to Microns: Extending the Measurement Range of Optical Profilers," *Proc. OSA. Fabrication and Testing Meeting Digest*, Seattle, Washington (1986b).
- Creath K., Y. Y. Cheng, and J. C. Wyant, "Contouring Aspheric Surfaces Using Two-Wavelength Phase-Shifting Interferometry," *Opt. Acta*, **32**, 1455 (1985).

- Dandliker R., E. Zimmermann, and G. Frosio, "Electronically Scanned White-Light Interferometry: A Novel Noise-Resistant Signal Processing," *Opt. Lett.*, **17**, 679–681 (1992).
- Danielson B. L. and C.Y. Boisrobert, "Absolute Optical Ranging Using Low Coherence Interferometry," *Appl. Opt.*, **30**, 2975–2979 (1991).
- Davidson M., K. Kaufmann, I. Mazor, and F. Cohen, "An Application of Interference Microscopy to Integrated Circuit Inspection and Metrology," *Proc. SPIE*, **775**, 233–247 (1987).
- Debnath S. K. and M. P. Kothiyal, "Optical Profiler Based on Spectrally Resolved White Light Interferometry," *Opt. Eng.*, **44**, 013606 (2005).
- Debnath S. K., M. P. Kothiyal, J. Schmit, and P. Hariharan, "Spectrally Resolved White-Light Phase-Shifting Interference Microscopy for Thickness Profile Measurement of Transparent Thin-Film Layers on Patterned Substrates," *Opt. Express*, **14**, 4662–4667 (2006).
- Deck L. L., "Fourier-Transform Phase-Shifting Interferometry," *Appl. Opt.*, **42**, 2354–2365 (2003).
- Deck L. and P. de Groot, "High-Speed Noncontact Profiler Based on Scanning White-Light Interferometry," *Appl. Opt.*, **33**, 7334–8 (1994).
- Decker J. E., J. R. Miles, A. A. Madej R. Siemsen, K. J. Siemsen, S. de Bonth, K. Bustraan, S. Temple, and J. R. Pekelski, "Increasing the Range of Unambiguity in Step-Height Measurement with Multiple-Wavelength Interferometry-Application to Absolute Long Gauge Block Measurement," *Appl. Opt.*, **42**, 5670–5678 (2003).
- de Groot P. J., "Three-Color Laser-Diode Interferometer," *Appl. Opt.*, **30**, 3612–3616 (1991).
- de Groot P. J., "Chromatic Dispersion Effects in Coherent Absolute Ranging," *Opt. Lett.*, **17**, 898–900 (1992).
- de Groot P. J., "Extending the Unambiguous Range of Two-Color Interferometers," *Appl. Opt.*, **33**, 5948–5953 (1994).
- de Groot P. J., "Measurement of Transparent Plates with Wavelength-Tuned Phase-Shifting Interferometry," *Appl. Opt.*, **39**, 2658–2663 (2000).
- de Groot P. J., "Optical Metrology," in: T. G. Brown, K. Creath, H. Kogelnik, et al. (Eds.), *The Optics Encyclopedia* (Vol. 3), Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim, 2003, pp. 2085–2117.
- de Groot P. J. and L. Deck, "Surface Profiling by Analysis of White-Light Interferograms in the Spatial Frequency Domain," *J. Mod. Opt.*, **42**, 389–401 (1995).
- de Groot P. J., X. C. de Lega, J. Kramer, and M. Turzhitsky, "Determination of Fringe Order in White Light Interference Microscopy," *Appl. Opt.*, **41**, 4571–4578 (2002).
- de Groot P. J. and X.C. de Lega, "Signal Modeling for Low-Coherence Height-Scanning Interference Microscopy," *Appl. Opt.*, **43**, 4821–4830 (2004).
- de Groot P. J. and S. Kishner, "Synthetic Wavelength Stabilization for Two-Color Laser-Diode Interferometry," *Appl. Opt.*, **30**, 4026–4033 (1991).
- de Groot P. J. and McGarvey, "Chirped Synthetic-Wavelength Interferometry," *Opt. Lett.* **17**, 1626–1628 (1992).
- Diaspro A., *Confocal and Two-Photon Microscopy*, Wiley- Liss, New York, 2002.
- Dobson S. L., P. C. Sun and Y. Fainman, "Diffractive Lenses for Chromatic Confocal Imaging," *Appl. Opt.*, **36**, 4744–4748 (1997).
- Doi T., K. Toyoda and Y. Tanimura, "Effects of Phase Changes on Reflection and Their Wavelength Dependence in Optical Profilometry," *Appl. Opt.*, **36**, 7157–7161 (1997).

- Dorsch R. G., G. Hausler and J. M. Herrmann, "Laser Triangulation: Fundamental Uncertainty in Distance Measurement," *Appl. Opt.*, **33**, 1306–1314 (1994).
- Downs M. J., W. H. McGiven, and H. J. Ferguson, "Optical System for Measuring the Profiles of Supersmooth Surfaces," *Precision Engineering*, **7**, 211–215 (1985).
- Dragoset R. A., R. D. Young, H. P. Layer, S. R. Mielczarek, E. C. Teague, and R. J. Celotta, "Scanning Tunelling Microscope Applied to Optical Surfaces," *Opt. Lett.*, **11**, 560–562 (1986).
- Dragoset R. A. and T. V. Vorburger, "Scanning Tunneling Microscopy (STM) of a Diamond-Turned Surface and Grating Replica," *Proc. SPIE*, **749**, 54 (1987).
- Duan Z., Y. Miyamoto, and M. Takeda, "Dispersion-Free Absolute Interferometry Based on Angular Spectrum Scanning," *Optics Express*, **14**, 655–663 (2006).
- Dubois A., J. Selb, L. Vabre, and A.-C. Boccaro, "Phase Measurement with Wide-Aperture Interferometers," *Appl. Opt.*, **39**, 2326–2331 (2000).
- Egger M. D. and M. Petráň, "New Reflected-Light Microscope for Viewing Unstained Brain and Ganglion Cells," *Science*, **157**, 305–307 (1967).
- Ei-Ghandoor H., "Tomographic Investigation of the Refractive Index Profiling Using Speckle Photography Techniques," *Opt. Com.*, **133**, 33–38 (1997).
- Fercher A. F., H. Z. Hu, and U. Vry, "Rough Surface Interferometry with Two-Wavelength Heterodyne Speckle Interferometry," *Appl. Opt.*, **24**, 2181–2188 (1985).
- Françon M., "Polarization Interference Microscopes," *Appl. Opt.*, **3**, 1033–1036 (1963).
- Françon M. and S. Mallick, *Polarization Interferometers*, Wiley-Interscience, London, New York, 1971.
- Friedman M., "Metrology," in R. C. Driggers, Ed., *Encyclopedia of Optical Engineering*, Taylor and Francis 2003.
- Gates J. W., "Fringe Spacing in Interference Microscopes," *J. Sci. Instrum.*, **33**, 507 (1956).
- Gerges A. S., F. Farah, T. P. Newson, J. D. C. Jones, and D. A. Jackson, "An Interferometric Fiber-Optic Sensor Using Short Coherence Length Source," *Electron. Lett.*, **23**(21), 1110–1111 (1987).
- Gu M., *Principles of Three-Dimensional Imaging in Confocal Microscopes*, World Scientific, Singapore, 1996.
- Güntherodt H. J. and Wiesendanger, R., *Scanning Tunneling Microscopy I: General Principles and Applications to Clean and Absorbate-Covered Surfaces*, vol. 20, Springer-Verlag, Berlin, 1994.
- Han S., "Packaged MEMS and MOEMS Measurement Using Through Transmissive Media (TTM) Profiler," *Proc. SPIE*, **6032**, 603203–1–5 (2006).
- Hanley Q. S., P. J. Verveer, M. J. Gemkow D. Arndt-Jovin, and T. M. Jovin, "An Optical Sectioning Programmable Array Microscope Implemented with a Digital Micromirror Device," *J. Microsc.*, **196**, 317–331 (2000).
- Hansma P. K. and J. Tersoff, "Scanning Tunneling Microscopy," *J. Appl. Phys.*, **61**(2), R1–R23 (1987).
- Harasaki A., J. Schmit, and J. C. Wyant, "Improved Vertical Scanning Interferometry," *Appl. Opt.*, **39**, 2107–2115 (2000).
- Harasaki A., J. Schmit, and J. C. Wyant, "Offset Envelope Position Due to Phase Change on Reflection," *Appl. Opt.*, **40**, 2102–2106 (2001).

- Hariharan P., "Achromatic Phase Shifting for Polarization Interferometers," *J. Mod. Opt.*, **43**, 1305–1306 (1996).
- Hariharan P. *Optical Interferometry*, Academic Press, San Diego, 2003, pp. 113–118.
- Hariharan P. "The Geometric Phase," in *Progress in Optics*, vol. 48, Elsevier, Amsterdam, 2005, pp. 149–201.
- Hariharan P., B. F. Oreb, and T. Eiju, "Digital Phase-Shifting Interferometry: A Simple Error Compensation Phase Calculation Algorithm," *Appl. Opt.*, **26**, 2504 (1987).
- Hariharan P. and P. E. Ciddor, "An Achromatic Phase Shifter Operating on the Geometric Phase," *Opt. Com.*, **110**, 13–17 (1994).
- Hariharan P., K. G. Larkin, and M. Roy, "The Geometric Phase: Interferometric Observations with White Light," *J. Mod. Opt.*, **41**, 663–667 (1994).
- Hariharan P. and M. Roy, "Interferometric Surface Profiling with White Light: Effects of Surface Films," *J. Mod. Opt.* **43**, 1797–1800 (1996).
- Haruna M., M. Ohmi, T. Mitsuyama, T. Tajiri, M. Maruyama, and M. Hashimoto, "Simultaneous Measurement of the Phase and Group Indices and the Thickness of Transparent Plates by Low-Coherence Interferometry," *Opt. Lett.*, **23**, 966–968 (1998).
- Häusler G. and J. M. Herrmann, "Physical Limits of 3D-Sensing," *Proc. SPIE: Optics, Illumination, and Image Sensing for Machine Vision VII*, **1822**, 150–158 (1992).
- Häusler G. and J. Neumann, "Coherence Radar—an Accurate 3D Sensors for Rough Surfaces," *Proc. SPIE*, **1822**, 200–205 (1992).
- Häusler G. and M. W. Lindner, "Coherence Radar" and "Spectral Radar"—New Tools for Dermatological Diagnosis," *J. Biomed. Optics*, **3**, 21–31 (1998).
- Heflinger L. O. and R. F. Wuerker, "Holographic Contouring Via Multifrequency Lasers," *Appl. Phys. Lett.*, **15**, 28–30 (1969).
- Helen S. S., M. P. Kothiyal, and R. S. Sirohi, "Achromatic Phase-Shifting by a Rotating Polarizer," *Opt. Com.*, **154**, 249–254 (1998).
- Helen S. S., M. P. Kothiyal, and R. S. Sirohi, "Analysis of Spectrally Resolved White Light Interferograms: Use of a Phase Shifting Technique," *Opt. Eng.*, **40**, 1329–1336 (2001).
- Hellmuth T., "Microscopy," in: T. G. Brown, K. Creath, H. Kogelnik, *The Optics Encyclopedia*, vol. 3, Wiley-VCH, Weinheim, 2003, pp. 1469–1531.
- Hettwer A., J. Kranz, and J. Schwider, "Three Channel Phase-Shifting Interferometer Using Polarization-Optics and a Diffraction Grating," *Opt. Eng.*, **39**, 960–966 (2000).
- Hibino K. and T. Takatsuji, "Suppression of Multiple-Beam Interference Noise in Testing an Optical-Parallel Plate by Wavelength-Scanning Interferometry," *Opt. Rev.*, **9**, 60–65 (2002).
- Hibino K., B. F. Oreb, and P. S. Fairman, "Wavelength-Scanning Interferometry of a Transparent Parallel Plate with Refractive Index Dispersion," *Appl. Opt.*, **42**, 3888–3895 (2003).
- Hibino K., B. F. Oreb, P. S. Fairman, and J. Burke, "Simultaneous Measurement of Surface Shape and Variation in Optical Thickness of a Transparent Parallel Plate in Wavelength-Scanning Fizeau Interferometer," *Appl. Opt.*, **43**, 1241–1249 (2004).
- Hildebrand B. P. and K.A. Haines, "Multiple Wavelength and Multiple Source Holography Applied to Contouring Generation," *J. Opt. Soc. Am.*, **57**, 155–156 (1967).
- Hirabayashi A., H. Ogawa, and K. Kitagawa, "Fast Surface Profiler by White-Light Interferometry by Use of a New Algorithm Based on Sampling Theory" *Appl. Opt.*, **41**, 4876–4883 (2002).
- Ho R. and S. Shao, "Axial Resolution of Confocal Microscopes Revisited," *Optik*, **88**, 147–154 (1991).

- Huntley J. M. and C. R. Coggrave, "Progress in Phase Unwrapping," *Proc. SPIE*, **3407**, 86–93 (1998).
- Huntley J. M. and H. O. Saldner, "Temporal Phase-Unwrapping Algorithm for Automated Fringe Analysis," *Appl. Opt.*, **32**, 3047–3052 (1993).
- Huntley J. M. and H.O. Saldner, "Error-Reduction Methods for Shape Measurement by Temporal Phase Unwrapping" *J. Opt. Soc. Am. A*, **14**, 3188–3196 (1997).
- Ichihara A., T. Tanaami, K. Isozaki et al., "High-Speed Confocal Fluorescent Microscopy Using a Nipkow Scanner with Microlenses for 3-D Imaging of Single Fluorescent Molecule in Real Time," *Bioimages*, **42**, 57–62 (1996).
- Ingelstam E., "Problems Related to the Accurate Interpretation of Microinterferograms," Paper presented at the Interferometry, National Physical Laboratory Symposium No. 11 (1960).
- Ishii Y. and R. Onodera, "Two-Wavelength Laser-Diode Interferometry That Uses Phase-Shifting Techniques," *Opt. Lett.*, **16**, 1523–1525 (1991).
- Itoh M., R. Yamada, R. Tian, M. Tsai, and T. Yatagai, "Broad-Band Light-Wave Correlation Topography Using Wavelet Transform," *Opt. Rev.*, 135–138 (1995).
- Kikuta H., K. Iwata, and R. Nagata, "Distance Measurement by the Wavelength Shift of Laser Diode Light," *Appl. Opt.*, **25**, 2976–2980 (1986).
- Kim D., S. W. Kim, H. J. Kim, H. J. Kong, and Y. Lee, "Measurement of the Thickness Profile of a Transparent Thin Film Deposited Upon a Pattern Structure with an Acousto-Optic Tunable Filter," *Opt. Lett.*, **27**, 1893–1895 (2002).
- Kim D. and S. W. Kim, "Direct Spectral Phase Function Calculation for Dispersive Interferometric Thickness Profilometry," *Optics Express*, **12**, 5117–5124 (2004).
- Kim S. W., M. G. Kang, and G. S. Han, "Accelerated Phase-Measurement Algorithm of Least Squares for Phase-Shifting Interferometry," *Opt. Eng.*, **36**, 3101–3106 (1997).
- Kim S. W. and G. H. Kim, "Thickness-Profile Measurement of Transparent Thin-Film Layers by White-Light Scanning Interferometry," *Appl. Opt.*, **38**, 5968–5973 (1999).
- Kimura S. and T. Wilson, "Effect of Axial Pinhole Displacement in Confocal Microscopes," *Appl. Opt.*, **32**, 2257–2261 (1993).
- Kino G. S., T. R. Corle, and G. Q. Xiao, "New Types of Scanning Optical Microscopes," *Proc. SPIE*, **921**, 116–22 (1988).
- Kino G. S. and S. S. C. Chim, "The Mirau Correlation Microscope," *Appl. Opt.*, **29**, 3775–3783 (1990).
- Kinoshita M., M. Takeda, H. Yago, Y. Watanabe, and T. Kurokawa, "Optical Frequency-Domain Microprofilometry with a Frequency-Tunable Liquid-Crystal Fabry-Perot Etalon Device," *Appl. Opt.*, **38**, 7063–7068 (1999).
- Kröner K., R. Windecker, M. Fleisher, and H. J. Tiziani, "One-Grating Projection for Absolute Three-Dimensional Profiling," *Opt. Eng.*, **40**, 1653–1660 (2001).
- Körner K., A. Ruprecht, and T. Wiesendanger, "Optical Profiling Techniques for Mems Measurement," in: W. Osten (Ed.), *Optical Inspection of Microsystems*, Taylor and Francis, Boca Raton, 2006.
- Kuwamura S. and I. Yamaguchi, "Wavelength Scanning Profilometry for Real-Time Surface Shape Measurement," *Appl. Opt.*, **37**, 4473–4482 (1997).
- Larkin K.G., "Efficient Demodulation for Bandpass Sampled Am Signals," *Electronics Letters*, **32**, 101–102 (1996a).

- Larkin K.G., "Efficient Nonlinear Algorithm for Envelope Detection in White Light Interferometry," *J. Opt. Soc. Am. A*, **13**, 832–842 (1996b).
- Lee B. S. and T. C. Strand, "Profilometry with a Coherence Scanning Microscope," *Appl. Opt.*, **29**, 3784–3788 (1990).
- Lehmann P., "Optical Versus Tactile Geometry Measurement – Alternatives or Counterparts" *Proc. SPIE*, **5144**, 183–196 (2003).
- Lin P. C., P. C. Sun, L. Zhu, and Y. Fainman, "Single-Shot Depth-Section Imaging through Chromatic Slit-Scan Confocal Microscopy," *Appl. Opt.*, **37**, 6764–6770 (1998).
- Magnov S. N., *Surface Analysis with Stm and Afm: Experimental and Theoretical Aspects of Image Analysis*, Wiley-VCH, Weinheim, 1996.
- Martínez-Corral M., P. Andres, J. Ojeda-Castaneda, and G. Saavedra, "Tunable Axial Super-resolution by Annular Binary Filters. Application to Confocal Microscopy," *Opt. Com.*, **119**, 491–498 (1995).
- Martínez-Corral M., M. Kowalczyk, C. J. Zapata-Rodríguez, and P. Andrés, "Tunable Optical Sectioning in Confocal Microscopy by Use of Symmetrical Defocusing and Apodization," *Appl. Opt.*, **37**, 6914–6921 (1998).
- Mehta, D. S., H. Hinosugi, S. Saito, M. Takeda, T. Kurokawa, H. Takahashi, M. Ando M., Shishido, and T. Yoshihawa, "Spectral Interferometric Microscope with Tandem Liquid-Crystal Fabry Perot Interferometers for Extension of the Dynamic Range in Three-Dimensional Step-Height Measurement" *Appl. Opt.*, **42**, 682–690 (2003a).
- Mehta D. S., S. Saito, H. Hinosugi, M. Takeda, and K. Kurokawa, "Spectral Interference Mirau Microscope with an Acousto-Optic Tunable Filter for Three-Dimensional Surface Profilometry" *Appl. Opt.*, **42**, 1296–1305 (2003b).
- Mehta D. S., M. Sugai, H. Hinosugi, S. Saito, M. Takeda, T. Kurokawa, H. Takahashi, M. Ando, M. Shishido, and T. Yoshihawa, "Simultaneous Three-Dimensional Step-Height Measurement and High-Resolution Tomographic Imaging with a Spectral Interferometric Microscope" *Appl. Opt.*, **41**, 3874–3885 (2002).
- Michelson A.A. and J. R. Benoit, "Détermination Expérimentale De La Valeur Du Mètre En Longueurs D'ondes Lumineuses," *Trav. Mem. Bur. Int Poids Mes.*, **11**, 1 (1895).
- Millerd J. E., N. J. Brock, J. B. Hayes, M. North-Morris, M. Novak, and J. C. Wyant, "Pixelated Phase-Mask Dynamic Interferometer," *Proc. SPIE*, **5531**, 304–314 (2004).
- Minsky M., *Microscopy Apparatus*, US Patent No. 3,013,467, (1961).
- Molesini G., G. Pedrini, P. Poggi, and F. Quercioli, "Focus Wavelength Encoded Optical Profilometer," *Opt. Com.*, **49**, 229–233 (1984).
- Montgomery, P. D. Montaner, O. Manzardo, and H. P. Herzig, "Lateral Error Reduction in the 3D Characterization of Deep MOEMS Devices Using White Light Interference Microscopy," *Proc. SPIE*, **5458**, 34–42 (2004).
- Moring I., H. Ailisto, V. Koivunen, and R. Myllylä, "Active 3D Vision System for Automatic Model-Based Shape Inspection," *Opt. Lasers Eng.*, **10** (3–4), 149–160 (1989).
- Nahm K. B. E., S. Shin and S.-M. Ryoo, "Three-Dimensional Reconstruction by Focus Sensing from Conventional 2D Image Sliced," *Proc. SPIE*, **3261**, 88–92 (1998).
- Neil M. A. A., R. Jujskaitis, and T. Wilson, "Method of Obtaining Optical Sectioning by Using Structured Light in a Conventional Microscope," *Opt. Lett.*, **22**, 1905–1907 (1997).
- Nomarski G., "Microinterferometre a Ondes Polarisees," *J. Phys. Radium* **16**, 9S–16S (1955).

- North-Morris M. B., J. E. Millerd, N. J. Brock, and J. B. Hayes, "Phase Shifting Dynamic Multiwavelength Interferometer," *Proc. SPIE*, **5531**, 64–75 (2004).
- Okada K., H. Sakuta, T. Ose, and J. Tsujiuchi, "Separate Measurements of Surface Shapes and Refractive Index Inhomogeneity of an Optical Element Using Tunable-Source Phase Shifting Interferometry," *Appl. Opt.*, **29**, 3280–3285 (1990).
- Onodera R. I., "Fourier Description of the Phase-Measuring Process in Two-Wavelength Phase-Shifting Interferometry," *Opt. Com.*, **137**, 27–30 (1997).
- Palik E. D. (Ed.) *Handbook of Optical Constants of Solids*, vol. I–III, Academic Press, San Diego, 1991.
- Palmer J. M., "Getting Intense About Intensity," *Metrologia*, **30**, 371 (1993).
- Pancharatnam S., "Generalized Theory of Interference and Its Applications Part I. Coherent Pencils," *Proc. Indian Acad. Sci.*, **A44**, 247–262 (1956).
- Park M. C. and S.-W. Kim, "Compensation of Phase Change on Reflection in White-Light Interferometry for Step Height Measurement," *Opt. Lett.*, **26**, 420–422 (2001).
- Park M. C. and Kim S.-W., "Direct Quadratic Polynomial Fitting for Fringe Peak Detection of White Light Scanning Interferograms," *Opt. Eng.*, **39**, 952–959 (2000).
- Paulson L., M. Sjö Dahl, J. Kato, and I. Yamaguchi, "Temporal Phase Unwrapping Applied to Wavelength-Scanning Interferometry," *Appl. Opt.*, **39**, 3285–3288 (2000).
- Pavlíček P. and J. Soubusta, "Theoretical Measurement Uncertainty of White-Light Interferometry on Rough Surfaces," *Appl. Opt.*, **42**, 1809–1813 (2003).
- Pavlíček P. and J. Soubusta, "Measurement of the Influence of Dispersion on White Light Interferometry," *Appl. Opt.*, **43**, 766–770 (2004).
- Pawley J. B. (Ed.), *Handbook of Biological Confocal Microscopy*, 2nd ed., Plenum Press, New York, 1995.
- Pawlowski M., Y. Sakano, Y. Miyamoto, and M. Takeda, "Phase-Crossing Algorithm for White-Light Fringes Analysis," *Opt. Commun.*, **260**, 68–72 (2006).
- Petran M., M. Hadravsky, M. D. Egger, and R. Galambos, "Tandem Scanning Reflected-Light Microscope," *J. Opt. Soc. Am. B*, **58**, 661–664 (1968).
- Pförtner A. and J. Schwider, "Dispersion Error in White-Light Linnik Interferometers and Its Implications for Evaluation Procedures," *Appl. Opt.*, **40**, 6223 (2001).
- Pluta M., *Advanced Light Microscopy: Measuring Techniques*, Elsevier Science Publishing Co, Amsterdam, 1993.
- Polarvarapu P. L. (Ed.), *Principles and Applications of Polarization-Division Interferometry*, John Wiley & Sons, New York, 1997.
- Polhemus C., "Two-Wavelength Interferometry," *Appl. Opt.*, **12**, 2071–2074 (1973).
- Recknagel R.-J. and G. Notni, "Analysis of White Light Interferograms," *Opt. Com.*, **148**, 122–128 (1998).
- Rogala E. W. and H. H. Barrett, "Phase-Shifting Interferometric Ellipsometer Capable of Measuring the Complex Index of Refraction and the Surface Profile of a Test Surface," *J. Opt. Soc. Am. A*, **15**, 538–548 (1998).
- Roy M., I. Cooper, P. Moore, C. Sheppard, and P. Hariharan, "White-Light Interference Microscopy: Effects of Multiple Reflections within a Surface Film," *Optics Express*, **13**, 164–170 (2005).
- Ruger D. and P. Hansma, "Atomic Force Microscopy," *Physics Today*, **43**, 23–30 (1990).

- Ruprecht A. K., Wiesendanger, and T. F., Tiziani, H. J., "Chromatic Confocal Microscopy with a Finite Pinhole Size," *Opt. Lett.*, **29**, 2130–2132 (2004).
- Saldner H. O. and J. M. Huntley, "Shape Measurement by Temporal Phase Unwrapping: Comparison of Unwrapping Algorithms," *Measurement Science & Technology*, **8**, 986–992 (1997a).
- Saldner H. O. and J. M. Huntley, "Temporal Phase Unwrapping: Application to Surface Profiling of Discontinuous Objects," *Appl. Opt.*, **36**, 2770–2775 (1997b).
- Sandoz P., "An Algorithm for Profilometry by White Light Phase Shifting Interferometry," *J. Mod. Opt.*, **43**, 1545–1554 (1996).
- Sandoz P., "Wavelet Transform as a Processing Tool in White-Light Interferometry," *Opt. Lett.*, **22**, 1065–1067 (1997).
- Sarid D., *Scanning Force Microscopy*, Oxford University Press, New York, 1991.
- Sarid D., D. Iams, V. Weissenberger, and L. S. Bell, "Compact Scanning-Force Microscope Using a Laser Diode," *Opt. Lett.*, **13**, 1057–1059 (1988).
- Schmidt W. and A. F. Fercher, "Holographic Generation of Depth Contour Using a Flash-Lamp-Pumped Dye Laser," *Opt. Commun.*, **3**, 363–365 (1971).
- Schmit J., "High Speed Measurements Using Optical Profiler," *Proc. SPIE*, **5144**, 46–56 (2003).
- Schmit J. and K. Creath, "Window Function Influence on Phase Error in Phase-Shifting Algorithms," *Appl. Opt.* **35**, 5642–5649 (1996).
- Schmit J., P. Unruh, A. Tavrov, W. Osten, N. Kerwien, and H. J. Tiziani, "Coherence Effects in Narrow Trench Measurements Using White Light Interferometry," *OSA Technical Digest*, **87**, MT41 (2003).
- Schneider R., A. Schick, P. Köllensperger, and T. Ninomiya, "High-Speed Optical Three-Dimensional Scanner for Automatic Solder Joint Inspection," *Opt. Eng.*, **36**, 2878–2885 (1997).
- Schneir J., J. A. Dagata, H. H. Harary, C. J. Evans, A. J. Melmed, H. B. Elswijk, and J. Sauvageau, "Scanning Tunneling Microscopy of Optical Surfaces," *Proc. SPIE*, **1164**, 112–120 (1989).
- Schulz G. and K.-E. Elssner, "Errors in Phase-Measurement Interferometry with High Numerical Aperture," *Appl. Opt.*, **30**, 4500–4506 (1991).
- Schwider J., R. Burrow, K.-E. Elssner, J. Grzanna, R. Spolaczyk, and K. Merkel, "Digital Wave-Front Measuring Interferometry: Some Systematic Error Sources," *Appl. Opt.* **22**, 3421–3432 (1983).
- Schwider J. and L. Zhou, "Dispersive Interferometric Profilometer," *Opt. Lett.*, **19**, 995–997 (1994).
- Scott P. J., "Recent Developments in the Measurement of Aspheric Surfaces by Contact Stylus Instrumentation," *Proc. SPIE*, **4927**, 199–207 (2002).
- Serry F. M. and J. Schmit, "Characterization and Measurement of Microcomponents with Atomic Force Microscopy," in W. Osten (Ed.), *Optical Inspection of Microsystems*, Taylor and Francis, Boca Raton, Florida, 2007.
- Sheppard C. J. R., "Scanning Confocal Microscopy," R. G. Driggers, Ed., *Encyclopedia of Optical Engineering*, Marcel Dekker, New York, 2003, pp. 2525–2544.
- Sheppard C. J. R., T. J. Connolly, J. Lee, and C. J. Cogswell, "Confocal Imaging of a Stratified Medium," *Appl. Opt.*, **33**, 631–640 (1994).

- Sheppard C. J. R. and M. Gu, "Aberration Compensation in Confocal Microscopy," *Appl. Opt.*, **30**, 3563–3567 (1991).
- Sheppard C. J. R. and M. Gu, "Axial Imaging Through an Aberrating Layer of Water in Confocal Microscopy," *Opt. Com.*, **88**, 180–190 (1992).
- Sheppard C. J. R. and D.K. Hamilton, "Edge Enhancement by Defocusing of Confocal Images," *Opt. Acta*, **31**, 723–727 (1984).
- Sheppard C. J. R. and K.G. Larkin, "Effect of Numerical Aperture on Interference Fringe Spacing," *Appl. Opt.*, **34**, 4731–4734 (1995).
- Sheppard C. J. R. and M. Roy, "Low-Coherence Interference Microscopy" in P. Török and F.J. Kao, Eds., *Optical Imaging Microscopy*, Springer, Berlin, 2003, pp. 257–273.
- Shiono T. and K. Setsune, "Wavelength-Independent Integrated Focus Sensor Using a Reflection Twin Micro-Fresnel Lens," *Appl. Opt.*, **28**, 5115–5121 (1989).
- Smith L. M. and C. C. Dobson, "Absolute Displacement Measurement Using Modulation of Spectrum of White Light in a Michelson Interferometer," *Appl. Opt.* **28**, 3339–3342, (1989).
- Smith P. J., C. M. Taylor, A. J. Shaw, and E. M. McCabe, "Programmable Array Microscopy with a Ferroelectric Liquid-Crystal Spatial Light Modulator," *Appl. Opt.*, **39**, 2664–2669 (2000).
- Sommargren G. E., "Optical Heterodyne Profilometry," *Appl. Opt.*, **20**, 610–616 (1981).
- Stout K. J. and L. Blunt (Eds.), *Three-Dimensional Surface Topography: Measurement, Interpretation and Application*, Penton Press, London, 1994.
- Su C., L. Huang, and P. Nielson, *In-Situ Measurement of in-Plane and out-of-Plane Force Gradient with a Torsional Resonance Mode*, AFM, *AIP Proc.* **696**, 357–364 (2003).
- Takeda M., D. S. Mehta, M. Pawlowski, and T. Kurokawa, "Recent Developments in Spectral Interference Microscopes," *Proc. SPIE*, **5776**, 22–28 (2005).
- Takeda M. and H. Yamamoto, "Fourier-Transform Speckle Profilometry: Three-Dimensional Shape Measurements of Diffuse Objects with Large Height Steps and/or Spatially Isolated Surfaces," *Appl. Opt.*, **33**, 7829–7837 (1994).
- Tanaami T., S. Otsuki, N. Tomosada, Y. Kosugi, M. Shimizu, and H. Ishido, "High-Speed 1-Frame/ms Scanning Confocal Microscope with a Microlens and Nipkow Disks," *Appl. Opt.*, **41**, 4704–4708 (2002).
- Thomas T. T., *Rough Surfaces*, Imperial College Press, London, 1999.
- Tiziani H. J., "Spectral and Temporal Phase Evaluation for Interferometry and Speckle Applications," in: P. Rastogi and D. Inaudi (Eds.), *Trends in Optical Nondestructive Testing and Inspection*, Elsevier, New York 2000, pp. 323–343.
- Tiziani H. J., B. Franze, and P. Haible, "Wavelength-Shift Speckle Interferometry for Absolute Profilometry Using Mode-Hop Free External Cavity Diode Laser," *J. Mod. Opt.*, **44**, 1485–1496 (1997).
- Tiziani H. J. and H. M. Uhde, "Three-Dimensional Analysis by a Microlens-Array Confocal Arrangement," *Appl. Opt.*, **33**, 567–572 (1994b).
- Tiziani H. J. and H.M. Uhde, "Three Dimensional Image Sensing by Chromatic Confocal Microscopy," *Appl. Opt.*, **33**, 1838–1843 (1994a).
- Tiziani H. J., M. Wegner, and D. Steudle, "Confocal Principle for Macro-and Microscopic Surface and Defect Analysis," *Opt. Eng.*, **39**, 32–39 (2000).

- Tavrov A., J. Schmit, N. Kerwien, W. Osten and H. Tiziani, "Diffraction Induced Coherence Levels," *Appl. Opt.*, **44**, 2202–2212 (2005).
- Tolomon F. R. J. G. W., "Fringe Spacing in Interference Microscopes," *J. Sci. Instrum.*, **33**, 236–238 (1956).
- Totzeck M., P. Gräupner, T. Heil, A. Göhnermeier, O. Dittmann, D. Krähmer, and V. Kamenov, "Polarization Influence on Imaging," *J. Microlith., Microfab., Microsyst.*, **4**, 031108-1-15 (2005).
- Totzeck M., N. Kerwien, A.V. Tavrov, E. Rosenthal, and H. J. Tiziani, "Quantitative Zernike Phase-Contrast Microscopy by Use of Structured Birefringent Pupil-Filters and Phase-Shift Evaluation," *Proc. SPIE*, **4777**, 1–11 (2002).
- Towers C. E., D. P. Towers, and J.D.C. Jones, "Optimum Frequency Selection in Multifrequency Interferometry," *Opt. Lett.*, **28**, 887–889 (2003).
- Towers C. E., D. P. Towers, and J.D.C. Jones, "Generalized Frequency Selection in Multifrequency Interferometry," *Opt. Lett.*, **29**, 1348–1350 (2004a).
- Towers C. E., D. P. Towers, and J.D.C. Jones, "Time Efficient Chinese Remainder Theorem Algorithm for Full-Field Fringe Phase Analysis in Multi-Wavelength Interferometry," *Optics Express*, **12**, 1136–1143 (2004b).
- Towers C. E., D. P. Towers, and J.D.C. Jones, "Absolute Fringe Order Calculation Using Optimised Multifrequency Selection in Full-Field Profilometry," *Opt. Lasers Eng.*, **43**, 788–800 (2005).
- Uhlendorf K., G. Notni, and R. Kowarschik, "Resolution Enhancement in a Reflection-Type Confocal Microscope with a Phase-Conjugate Mirror," *Appl. Opt.*, **38**, 869–873 (1999).
- van Loenen E. J., D. Dijkkamp, A.J. Hoeven, J. M. Lenssinck, and J. Dieleman, "Evidence for Tip Imaging in Scanning Tunneling Microscopy," *Applied Physics Lett.*, **56**, 1755–1757 (1990).
- Virgil B. and J. A. Gurley, *Tapping Atomic Force Microscope with Phase or Frequency Detection*, U.S. Patent RE36, 488 (2000).
- Vorburger T. V. and J. Raja, "Surface Finish and Metrology," NISTIR 89–4088, U. S. Dept. of Commerce, National Institute of Standards and Technology, Gaithersburg, MD, June, 1990.
- Wagner C. and G. Hausler, "Information Theoretical Optimization for Optical Range Sensors," *Appl. Opt.* **42**, 5418–5426 (2003).
- Wan D. S., J. Schmit, and E. Novak, "Effects of Source Shape on Numerical Aperture Factor with Geometrical-Optics Model," *Appl. Opt.*, **43**, 2023–2028 (2004).
- Wan D. S., J. Schmit, and E. Novak, "Effects of Source Shape on the Numerical Aperture Factor with a Geometrical-Optics Model," *Appl. Opt.*, **43**, 2023–2018 (2004b).
- Whitehouse D. J., "Surface Metrology," *Sci. Meas. Tech.*, **8**, 955–973 (1997).
- Whitehouse D. J., *The Handbook of Surface and Nanometrology*, IOP Publishing, Bristol, 2003.
- Wickramasinghe H. K., "Scanned-Probe Microscopes," *Sci. Amer.*, 261, No 4, (1989).
- Wiesendanger R. and H. J. Güntherodt, *Scanning Tunneling Microscopy Ii: Further Applications and Related Scanning Techniques*, vol. 28, 2nd ed., Springer-Verlag, Berlin, 1995.
- Wiesendanger R. and H. J. Güntherodt, *Scanning Tunneling Microscopy Iii: Theory of Stm and Related Scanning Probe Methods*, vol. 29, 2nd ed., Springer-Verlag, Berlin, 1996.
- Wilson T., *Confocal Microscopy*, Academic Press, London, 1990.
- Windecker R., M. Fleischer, and H.J. Tiziani, "White-Light Interferometry with an Extended Zoom Range," *J. Mod. Opt.*, **46**, 1123–1135 (1999).

- Wojtkowski M., A. Kowalczyk, R. Leitgeb and A. F. Fercher “Full Range Complex Spectral Optical–Coherence Tomography Technique in Eye Imaging,” *Opt. Lett.*, **27**, 1415–1417 (2002).
- Wyant J. C., “Testing Aspherics Using Two-Wavelength Holography,” *Appl. Opt.*, **10**, 2113–2118 (1971).
- Wyant J. C. and K. Creath, *Two-Wavelength Phase-Shifting Interferometer and Method*, U.S. Patent No. 4,832,489 (1989).
- Wyant J. C., B. F. Oreb, and P. Hariharan, “Testing Aspherics Using Two-Wavelength Holography: Use of Digital Electronic Techniques,” *Appl. Opt.*, **23**, 4020–4023 (1984).
- Xiao G. Q., T. R. Corle, and G. S. Kino, “Real Time Confocal Scanning Optical Microscope,” *Appl. Phys. Lett.*, **53**, 716–718 (1988).
- Yamaguchi I., A. Yamamoto, and S. Kuwamura, “Speckle Decorrelation in Surface Profilometry by Wavelength Scanning Interferometry,” *Appl. Opt.*, **37**, 6721–6728 (1998).
- Yamaguchi I., A. Yamamoto, and M. Yano, “Surface Topography by Wavelength Scanning Interferometry,” *Opt. Eng.*, **39**, 40–46 (2000).
- Yamamoto A., C. C. Kuo, K. Sunouchi, S. Wada, I. Yamaguchi, and H. Tashiro, “Surface Shape Measurement by Wavelength Scanning Interferometry Using an Electronically Tuned Ti:Sapphire Laser,” *Opt. Rev.*, **8**, 59–63 (2001).
- Yatagai T., “Recent Progress in White Light Interferometry,” *Proc. SPIE*, **2340**, 338–345 (1994).
- Zelenka J. S. and J. R. Varner, “Multiple-Index Holographic Contouring,” *Appl. Opt.*, **8**, 1431–1434 (1969).
- Zhang J.-H. and L. Cai, “Profilometry Using and Optical Stylus with Interferometric Readout,” *Meas. Sci. Technol.*, **8**, 546–549 (1997).
- Zhong Q., D. Inniss, K. Kjoller, and V. B. Elings, “Tapping Mode Atomic Force Microscopy,” *Surf. Sci. Lett.*, **290**, L688 (1993).
- Zuluaga A. F. and R. Richards-Kortum, “Spatially Resolved Spectral Interferometry for Determination of Subsurface Structure,” *Opt. Lett.*, **24**, 519–521 (1999).

16

Optical Metrology of Diffuse Surfaces

K. Creath, J. Schmit, and J. C. Wyant

This chapter discusses moiré, fringe projection, structured illumination, holographic interferometry, digital holography, and speckle interferometry techniques for testing diffuse surfaces. Diffuse surfaces may be ground optical surfaces; or more often than not, they are other types of engineering surfaces or human figures. The main applications of these techniques are to determine surface form and shape or to measure displacement due to stress and object motion. When measuring surface form, these techniques provide a coarser and more flexible means of testing a wider variety of surfaces than do conventional interferometers. For displacement measurement due to applied stress, static as well as time-average and dynamic displacements can be determined quantitatively. These techniques are used a lot in machine vision applications, for process control, and for specialized measurement tasks on engineering components. Applications range from measuring the shape of an airplane window to determining whether components will stay on a circuit board, to studying vibration modes of turbine blades, to monitoring the alignment of segments in a large segmented telescope, to making replicas of historic sculptures, producing a well-fitting pair of jeans, and creating animated movies or video games with realistic motion of live figures. The newest techniques pushing the limits of this technology currently focus on rapid prototyping and real-time shape observation for multimedia and security.

16.1. MOIRÉ AND FRINGE PROJECTION TECHNIQUES

16.1.1. Introduction

The term moiré is not the name of a person; in fact, it is a French word referring to “an irregular wavy finish usually produced on a fabric by pressing between engraved rollers” (*Webster’s*, 1981). In optics, it refers to a beat pattern produced between two

gratings of approximately equal spacing. It can be seen in everyday things, such as the overlapping of two window screens, the rescreening of a half-tone picture, or with a striped shirt seen on television. The use of moiré for reduced sensitivity testing was introduced by Lord Rayleigh in 1874. Lord Rayleigh looked at the moiré between two identical gratings to determine their quality even though each individual grating could not be resolved under a microscope.

Fringe projection entails projecting a fringe pattern or grating on an object and viewing it from a different direction. The first use of fringe projection for determining surface topography was presented by Rowe and Welford in 1967. It is a convenient technique for contouring objects which are too coarse to be measured with standard interferometry. Fringe projection is related to optical triangulation using a single point of light and light sectioning where a single line is projected onto an object and viewed in a different direction to determine the surface contour (Case et al., 1987). These techniques are usually used with diffuse objects; however, alternative methods have been developed to measure specular surfaces by looking at the fringe reflection (Ritter and Hahn, 1983; Hang et al., 2000).

Moiré and fringe projection interferometry complement conventional holographic interferometry, especially for testing optics to be used at long wavelengths. Although two-wavelength holography (TWH) can be used to contour surfaces at any longer-than-visible wavelength, visible interferometric environmental conditions are required. Moiré and fringe projection interferometry can contour surfaces at any wavelength longer than 10–100 μm with reduced environmental requirements and no intermediate photographic recording setup. Moiré is also a useful technique for aiding in the understanding of interferometry.

This chapter explains what moiré is and how it relates to interferometry. Contouring techniques utilizing fringe projection, projection and shadow moiré, and two-angle holography are all described and compared. All of these techniques provide the same result and can be described by a single theory. The relationship between these techniques and the holographic and conventional interferometry will be shown. Errors caused by divergent geometries are described, and applications of these techniques combined with phase-measurement techniques are presented. Further information on these techniques can be found in the following books and book chapters: Varner (1974), Vest (1979), Hariharan (1984), Gasvik (2002), Chiang (1983) Patorski and Kujawinska (1993), Post et al. (1997), Amridror (2000), and Walker (2004).

16.1.2. What is Moiré?

Moiré patterns are extremely useful to help understand basic interferometry and interferometric test results. Figure 16.1 shows the moiré pattern (or beat pattern) produced by two identical straight line gratings rotated by a small angle relative to each other. A dark fringe is produced where the dark lines are out of step one-half period, and a bright fringe is produced where the dark lines for one grating fall on top of the corresponding dark lines for the second grating. If the angle between the two gratings is increased, the separation between the bright and dark

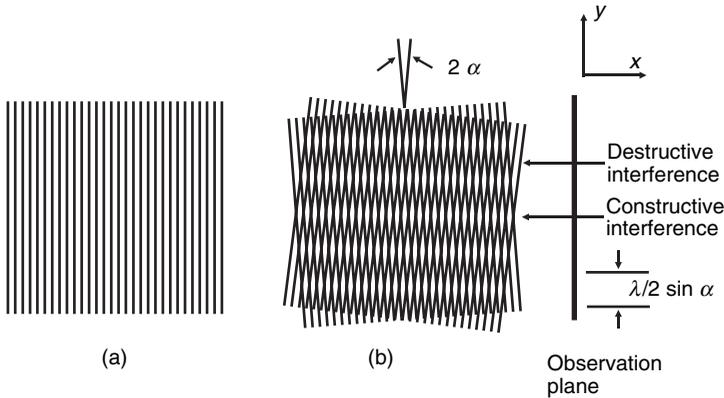


FIGURE 16.1. (a) Straight line grating. (b) Moiré between two straight line gratings of the same pitch at an angle 2α with respect to one another.

fringes decreases. (A simple explanation of moiré is given by Oster and Nishijima (1963).)

If the gratings are not identical to the straight line gratings, the moiré pattern (bright and dark fringes) will not be straight equi-spaced fringes. The following analysis shows how to calculate the moiré pattern for arbitrary gratings. Let the intensity transmission function for two gratings $f_1(x, y)$ and $f_2(x, y)$ be given by

$$\begin{aligned} f_1(x, y) &= a_1 + \sum_{n=1}^{\infty} b_{1n} \cos[n\phi_1(x, y)] \\ f_2(x, y) &= a_2 + \sum_{m=1}^{\infty} b_{2m} \cos[m\phi_2(x, y)] \end{aligned} \quad (16.1)$$

where $\phi(x, y)$ is the function describing the basic shape of the grating lines. For the fundamental frequency, $\phi(x, y)$ is equal to an integer times 2π at the center of each bright line and is equal to an integer plus one-half times 2π at the center of each dark line. The b coefficients determine the profile of the grating lines (i.e., square wave, triangular, sinusoidal, etc). For a sinusoidal line profile, b_{1i} is the only nonzero term.

When these two gratings are superimposed, the resulting intensity transmission function is given by the product

$$\begin{aligned} f_1(x, y)f_2(x, y) &= a_1a_2 + a_1 \sum_{m=1}^{\infty} b_{2m} \cos[m\phi_2(x, y)] + a_2 \sum_{n=1}^{\infty} b_{1n} \cos[n\phi_1(x, y)] \\ &+ \sum_{m=1}^{\infty} \sum_{n=1}^{\infty} b_{1n}b_{2m} \cos[n\phi_1(x, y)] \cos[m\phi_2(x, y)]. \end{aligned} \quad (16.2)$$

The first three terms of Eq. (16.2) provide information which can be determined by looking at the two patterns separately. The last term is the interesting one, and can be rewritten as

$$\begin{aligned} \text{Term 4} &= \frac{1}{2} b_{11} b_{21} \cos[\phi_1(x, y) - \phi_2(x, y)] \\ &+ \frac{1}{2} \sum_{m=1}^{\infty} \sum_{n=1}^{\infty} b_{1n} b_{2m} \cos[n\phi_1(x, y) - m\phi_2(x, y)]; n \text{ and } m \text{ both } \neq 1 \\ &+ \frac{1}{2} \sum_{m=1}^{\infty} \sum_{n=1}^{\infty} b_{1n} b_{2m} \cos[n\phi_1(x, y) + m\phi_2(x, y)] \end{aligned} \quad (16.3)$$

This expression shows that by superimposing the two gratings, the sum and difference between the two gratings is obtained. The first term of Eq. (16.3) represents the difference between the fundamental pattern making up the two gratings. It can be used to predict the moiré pattern shown in Figure 16.1. Assuming that two gratings are oriented with an angle 2α between them with the y axis of the coordinate system bisecting this angle, the two grating functions $\phi_1(x, y)$ and $\phi_2(x, y)$ can be written as

$$\phi_1(x, y) = \frac{2\pi}{\lambda_1} (x \cos \alpha + y \sin \alpha)$$

and

$$\phi_2(x, y) = \frac{2\pi}{\lambda_2} (x \cos \alpha - y \sin \alpha) \quad (16.4)$$

where λ_1 and λ_2 are the line spacings of the two gratings. Equation (16.4) can be rewritten as

$$\phi_1(x, y) - \phi_2(x, y) = \frac{2\pi}{\lambda_{\text{beat}}} x \cos \alpha + \frac{4\pi}{\bar{\lambda}} y \sin \alpha \quad (16.5)$$

where $\bar{\lambda} = (\lambda_1 + \lambda_2)/2$ is the average line spacing, and λ_{beat} is the beat wavelength between the two gratings given by

$$\lambda_{\text{beat}} = \frac{\lambda_1 \lambda_2}{\lambda_2 - \lambda_1} \quad (16.6)$$

Note that this beat wavelength is the same that was obtained for two-wavelength interferometry as described in Chapter 15, and is also referred to as the synthetic or equivalent wavelength. Using Eq. (16.3), the moiré or beat will be lines whose centers satisfy the equation

$$\phi_1(x, y) - \phi_2(x, y) = M2\pi \quad (16.7)$$

Three separate cases for moiré fringes can be considered.

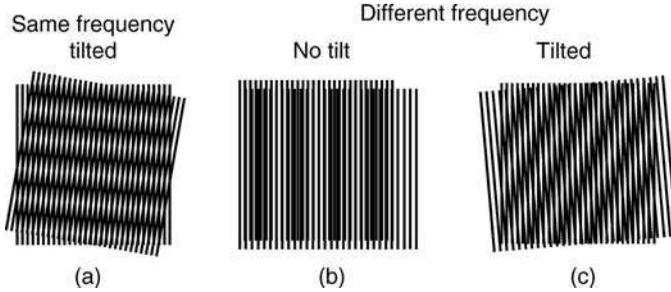


FIGURE 16.2. Moiré patterns caused by two straight line gratings with (a) the same pitch tilted with respect to one another, (b) different frequencies and no tilt, and (c) different frequencies tilted with respect to one another.

For the first case $\lambda_1 = \lambda_2 = \lambda$. The first term of Eq. (16.5) is zero, and the fringe centers are given by

$$M\lambda = 2y \sin \alpha \quad (16.8)$$

where M is an integer corresponding to the fringe order (see Fig. 16.2(a)). As was expected, Eq. (16.8) is the equation of equi-spaced horizontal lines as seen in Fig. 16.1.

For the second simple case $\lambda_1 \neq \lambda_2$ and the gratings are parallel to each other with $\alpha = 0$. This makes the second term of Eq. (16.5) vanish. The moiré will then be lines which satisfy

$$M\lambda_{\text{beat}} = x \quad (16.9)$$

These fringes are equally-spaced lines parallel to the grating lines (see Fig. 16.2(b)).

For the third and more general case where the two gratings have different line spacings $\lambda_1 \neq \lambda_2$ and the angle between the gratings is nonzero $\alpha \neq 0$, the equation for the moiré fringes will now be

$$M\bar{\lambda} = \frac{\bar{\lambda}}{\lambda_{\text{beat}}} x \cos \alpha + 2y \sin \alpha \quad (16.10)$$

This is the equation of straight lines whose spacing and orientation is dependent upon the relative difference between the two grating spacings and the angle between the gratings (see Fig. 16.2(c)).

The orientation and spacing of the moiré fringes for the general case can be determined from the geometry shown in Figure. 16.3 (Chiang, 1983). The distance \overline{AB} can be written in terms of the two grating spacings,

$$\overline{AB} = \frac{\lambda_1}{\sin(\theta - \alpha)} = \frac{\lambda_2}{\sin(\theta + \alpha)} \quad (16.11)$$

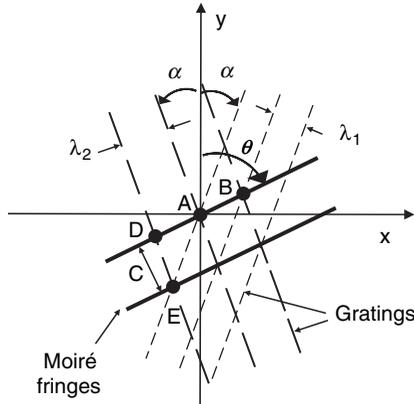


FIGURE 16.3. Geometry used to determine spacing and angle of moiré fringes between two gratings of different frequencies tilted with respect to one another.

where θ is the angle the moiré fringes make with the y axis. After rearranging, the fringe orientation angle θ is given by

$$\tan \theta = \tan \alpha \left(\frac{\lambda_1 + \lambda_2}{\lambda_2 - \lambda_1} \right) \tag{16.12}$$

When $\alpha = 0$ and $\lambda_1 \neq \lambda_2$, $\theta = 0^\circ$, and when $\lambda_1 = \lambda_2$ with $\alpha \neq 0$, $\theta = 90^\circ$ as expected. The fringe spacing perpendicular to the fringe lines can be found by equating quantities for the distance \overline{DE} ,

$$\overline{DE} = \frac{\lambda_1}{\sin 2\alpha} = \frac{C}{\sin(\theta + \alpha)} \tag{16.13}$$

where C is the fringe spacing or contour interval. This can be rearranged to yield

$$C = \lambda_1 \left[\frac{\sin(\theta + \alpha)}{\sin 2\alpha} \right] \tag{16.14}$$

By substituting for the fringe orientation θ , the fringe spacing can be found in terms of the grating spacings and angle between the gratings;

$$C = \frac{\lambda_1 \lambda_2}{\sqrt{\lambda_2^2 \sin^2 2\alpha + (\lambda_2 \cos 2\alpha - \lambda_1)^2}} \tag{16.15}$$

In the limit that $\alpha = 0$ and $\lambda_1 \neq \lambda_2$, the fringe spacing equals λ_{beat} , and in the limit that $\lambda_1 = \lambda_2 = \lambda$ and $\alpha \neq 0$, the fringe spacing equals $\lambda / (2 \sin \alpha)$. It is possible to determine λ_2 and α from the measured fringe spacing and orientation as long as λ_1 is known (Chiang, 1983).

16.1.3. Moiré and Interferograms

Now that we have covered the basic mathematics of moiré patterns, let us see how moiré patterns are related to interferometry. The single grating shown in Figure 16.1 can be thought of as a “snapshot” of plane waves (like in a collimated beam) traveling to the right where the distance between the grating lines is equal to the wavelength of light. Superimposing the two sets of grating lines in Fig. 16.1b can be thought of as superimposing two plane waves with an angle of 2α between their directions of propagation. When the two waves are in phase, bright fringes result (constructive interference) and when they are out of phase, dark fringes result (destructive interference). For a collimated plane wave, the “grating” lines are really planes (sheets) perpendicular to the plane of the figure and the dark and bright fringes are also planes perpendicular to the plane of the figure. If the plane waves are traveling to the right, these fringes would be observed by placing a screen perpendicular to the plane of the figure and to the right of the grating lines as shown in Figure 16.1. The spacing of the interference fringes on the screen is given by Eq. (16.8), where λ is now the wavelength of light. Thus, the moiré of two straight-line gratings correctly predicts the centers of the interference fringes produced by interfering two plane waves. Since the gratings used to produce the moiré pattern are binary gratings, the moiré does not correctly predict the sinusoidal intensity profile of the interference fringes. (If both gratings had sinusoidal intensity profiles, the resulting moiré would still not have a sinusoidal intensity profile because of higher-order terms.)

More complicated gratings, such as circular gratings, can also be investigated. Figure 16.4(b) shows the superposition of two identical circular grating patterns shown in Figure 16.4(a). This composite pattern indicates the fringe positions obtained by interfering two spherical wavefronts. The centers of the two circular gratings can be considered the source locations for two spherical waves. Just as for two plane waves, the spacing between the grating lines is equal to the wavelength of light. When the two patterns are in phase, bright fringes are produced, and when the patterns are completely out of phase, dark fringes result. For a point on a given fringe, the difference in the distances from the two source points and the fringe point is a constant. Hence, the fringes are hyperboloids. Due to symmetry, the fringes seen on observation plane A of Figure 16.4(b) must be circular. (Plane A is along the top of Fig. 16.4(b) and perpendicular to the line connecting the two sources as well as perpendicular to the page.) Figure 16.4(c) shows a representation of these interference fringes and represents the interference pattern obtained by interfering a non-tilted plane wave and a spherical wave. (A plane wave can be thought of as a spherical wave with an infinite radius of curvature.) Figure 16.4(d) shows that the interference fringes in plane B are essentially straight equi-spaced fringes going into the page. (These fringes are strictly speaking still hyperbolas, but in the limit of large distances, they are essentially straight lines. Plane B is along the side of Fig. 16.4(b) and parallel to the line connecting the two sources as well as perpendicular to the page.)

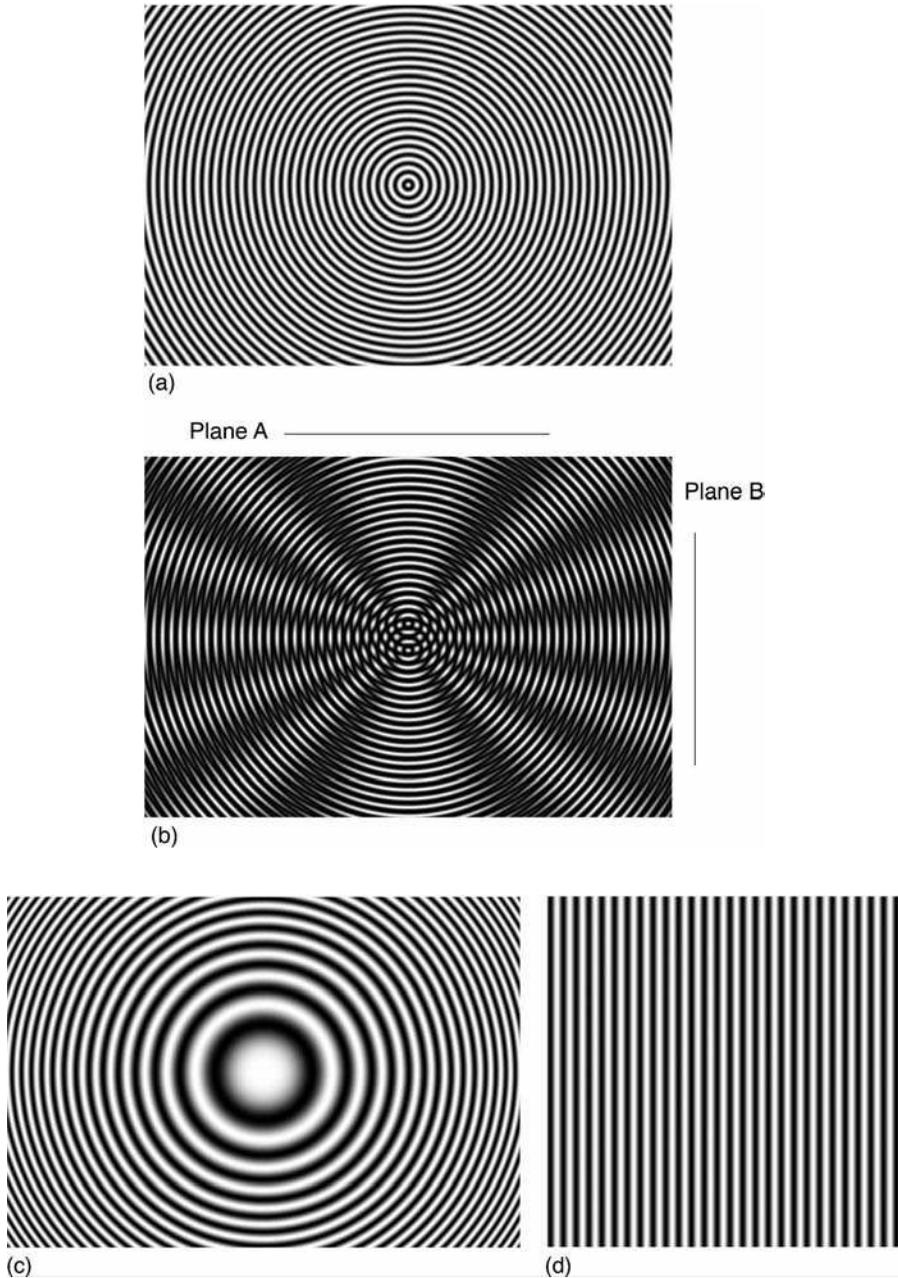


FIGURE 16.4. Interference of two spherical waves. (a) Circular line grating representing a spherical wavefront. (b) Moiré pattern obtained by superimposing two circular line patterns. (c) Fringes observed in plane A. (d) Fringes observed in plane B.

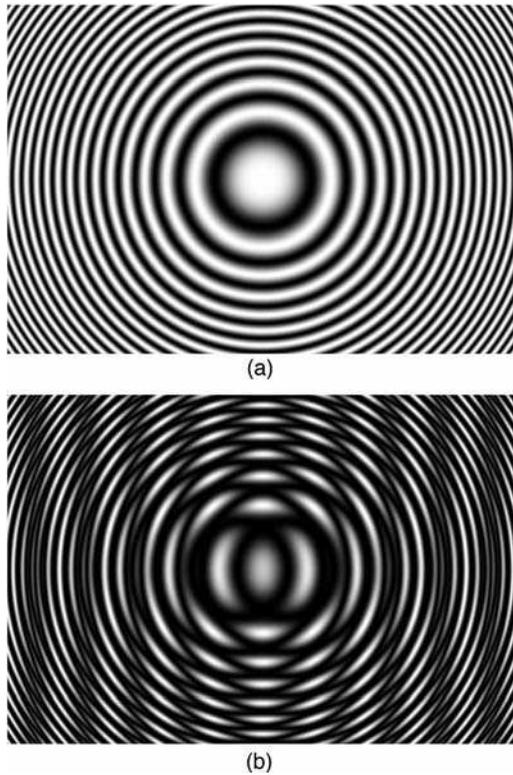


FIGURE 16.5. Moiré pattern produced by two zone plates. (a) Zone plate. (b) Straight line fringes resulting from superposition of two zone plates.

The lines of constant phase in plane B, if there were only a single spherical wave (single point source), are shown in Figure 16.5(a). (To first-order, the lines of constant phase in plane B are the same shape as the interference fringes in plane A.)

The pattern shown in Figure 16.5(a) is commonly called a zone plate. Figure 16.5(b) shows the superposition of two linearly displaced zone plates. The resulting moiré pattern of straight equi-spaced fringes illustrates the interference fringes in plane B shown in Figure 16.4(b).

Superimposing two interferograms and looking at the moiré or beat produced can be extremely useful. The moiré formed by superimposing two different interferograms shows the difference in the aberrations of the two interferograms. For example, Figure 16.6 shows the moiré produced by superimposing two computer-generated interferograms. One interferogram has 20 waves of tilt across the radius (Fig. 16.6(a)), while the second interferogram has 20 waves of tilt plus 4 waves of defocus (Fig. 16.6(b)). If the interferograms are aligned such that the tilt direction is the same for both interferograms, the tilt will cancel and only the 4 waves of defocus remain (Fig. 16.6(c)). In Figure 16.6(d), the two interferograms are rotated slightly

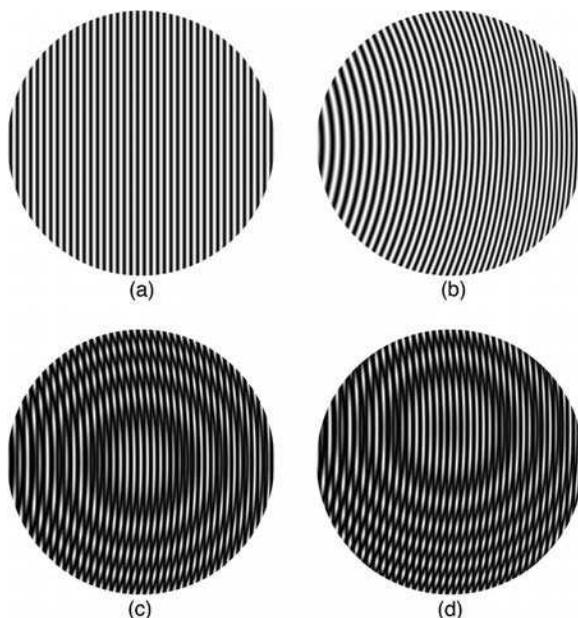


FIGURE 16.6. Moiré between two interferograms. (a) Interferogram having 20 waves tilt. (b) Interferogram having 20 waves tilt plus 4 waves of defocus. (c) Superposition of 16.6a and 16.6b with no tilt between patterns. (d) Slight tilt between patterns.

with respect to each other so that the tilt will not quite cancel. These results can be described mathematically by looking at the two grating functions:

$$\phi_1(x, y) = 2\pi(20\rho \cos \varphi + 4\rho^2)$$

and

$$\phi_2(x, y) = 2\pi[20\rho \cos(\varphi + \alpha)] \quad (16.16)$$

A bright fringe is obtained when

$$\frac{\phi_1 - \phi_2}{2\pi} = 20\rho[\cos \varphi - \cos(\varphi + \alpha)] + 4\rho^2 = M \quad (16.17)$$

If $\alpha = 0$, the tilt cancels completely and 4 waves of defocus remain; otherwise, some tilt remains in the moiré pattern.

Figure 16.7 shows similar results for interferograms containing third-order aberrations. A computer-generated interferogram having 22 waves of tilt across the radius, 4 waves of spherical and -2 waves of defocus is shown in Figure 16.7(a). Net spherical aberration with defocus and tilt is shown in Figure 16.7(d). This is the result of moiré between the interferogram in Figure 16.7(a) with an interferogram

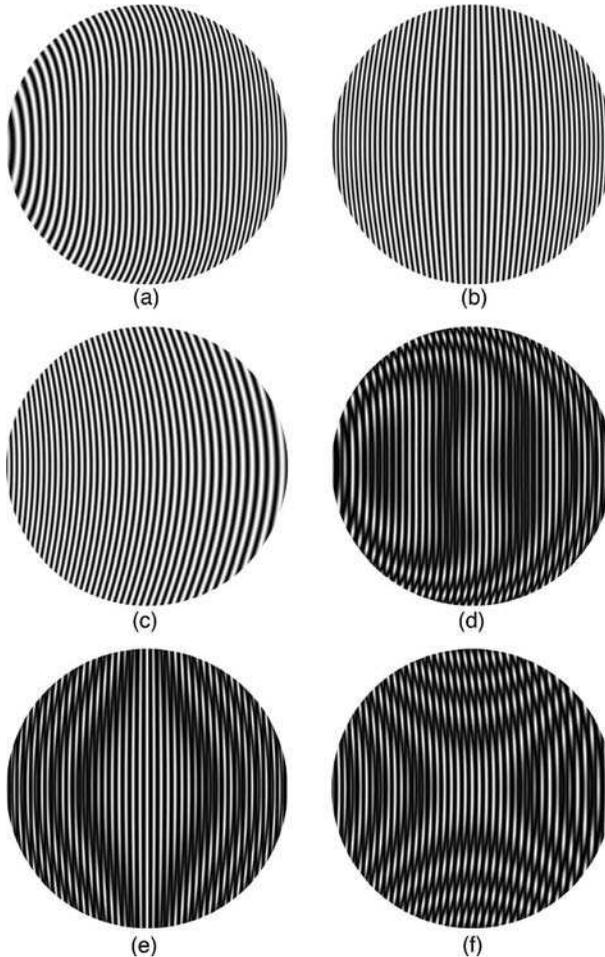


FIGURE 16.7. Moiré patterns showing third-order aberrations. Interferograms containing (a) 22 waves tilt, 4 waves of third-order spherical aberration, and -2 waves of defocus, (b) 20 waves tilt and 5 waves coma, and (c) 20 waves tilt, 7 waves astigmatism, and -3.5 waves of defocus. (d) Moiré pattern between Figure 16.6a and 16.7a. (e) Moiré pattern between Figures 16.6(a) and 16.7(b). (f) Moiré pattern between Figures 16.6(a) and 16.7(c).

having 20 waves of tilt (Fig. 16.6(a)). Figure 16.7(e) shows the moiré between an interferogram having 20 waves of tilt (Fig. 16.6(a)) with an interferogram having 20 waves of tilt and 5 waves of coma (Fig. 16.7(b)) netting 5 waves of coma in the moiré. The moiré between an interferogram having 20 waves of tilt (Fig. 16.6(a)) and one having 20 waves of tilt, 7 waves third-order astigmatism, and -3.5 waves defocus (Fig. 16.7(c)) is shown in Figure 16.7(f). Thus, it is possible to produce simple fringe patterns using moiré. These patterns can be printed or photocopied onto transparencies and used as a learning aid to understand interferograms obtained from

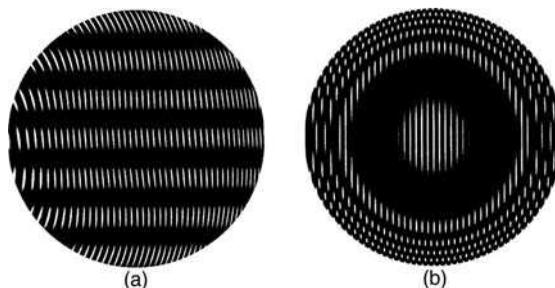


FIGURE 16.8. Moiré pattern by superimposing two identical interferograms (Figure 16.7(a)). (a) Both patterns having the same orientation. (b) One pattern is flipped.

third-order aberrations. Electronic copies are available at JC Wyant's website (Wyant, 2006) as well as on the accompanying CD to this book.

Figure 16.8(a) shows two identical interferograms superimposed with a small rotation between them. As we might by now expect, the moiré pattern consists of nearly straight equi-spaced lines. However, when one of the two interferograms is flipped over, the aberrations will add rather than subtract, and the resultant moiré is shown in Figure 16.8(b). When one interferogram is flipped, the fringe deviation from straightness in one interferogram is to the right and, in the other, to the left. Thus, the signs of the defocus and spherical aberration for the two interferograms are opposite and the resulting moiré pattern has twice the defocus and spherical of each of the individual interferograms.

When two identical interferograms given by Figure 16.7(a) are superimposed with a displacement from one another, a shearing interferogram is obtained. Figure 16.9 shows vertical and horizontal displacements with and without a rotation between the two interferograms. The rotations indicate the addition of tilt to the interferograms. These types of moiré patterns are very useful for understanding lateral shearing interferograms.

Moiré patterns are produced by multiplying two intensity distribution functions. Adding two intensity functions does not give the difference term obtained in Eq. (16.3). A moiré pattern is not obtained if two intensity functions are added. The only way to get a moiré pattern by adding two intensity functions is to use a nonlinear detector. For the detection of an intensity distribution given by $I_1 + I_2$, a nonlinear response can be written as

$$\text{Response} = a(I_1 + I_2) + b(I_1 + I_2)^2 + \dots \quad (16.18)$$

This produces terms proportional to the product of the two intensity distributions in the output signal. Hence, a moiré pattern is obtained if the two individual intensity patterns are simultaneously observed by a nonlinear detector (even if they are not multiplied before detection). If the detector produces an output linearly proportional to the incoming intensity distribution, the two intensity patterns must be multiplied to produce the moiré pattern. Since the eye is a nonlinear detector, moiré can be seen

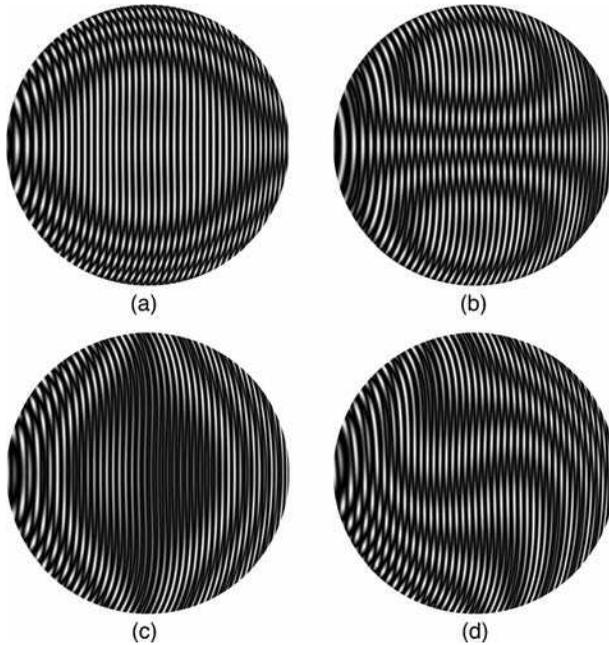


FIGURE 16.9. Moiré patterns formed using two identical interferograms (Figure 16.7(a)) where the two are sheared with respect to one another. (a) Vertical displacement. (b) Vertical displacement with rotation showing tilt. (c) Horizontal displacement. (d) Horizontal displacement with rotation showing tilt.

whether the patterns are added or multiplied. A good TV camera, on the other hand, will not see moiré unless the patterns are multiplied.

16.1.4. Historical Review

Since Lord Rayleigh first noticed the phenomena of moiré fringes, moiré techniques have been used for a number of testing applications. Righi (1887) first noticed that the relative displacement of two gratings could be determined by observing the movement of the moiré fringes. The next significant advance in the use of moiré was presented by Weller and Shepherd (1948). They used moiré to measure the deformation of an object under applied stress by looking at the differences in a grating pattern before and after the applied stress. They were the first to use shadow moiré, where a grating is placed in front of a nonflat surface, to determine the shape of the object behind it by using the shape of the moiré fringes. A rigorous theory of moiré fringes did not exist until the mid-fifties when Ligtenberg (1955) and Guild (1956, 1960) explained moiré for stress analysis by mapping slope contours and displacement measurement, respectively. Excellent historical reviews of the early work in moiré have been presented by Theocaris (1962, 1966). Books on this subject have been written by Guild (1956, 1960), Theocaris (1969), and Durelli and Parks (1970). Projection moiré techniques were introduced by Brooks and Helfinger (1969) for

optical gauging and deformation measurement. Until 1970, advances in moiré techniques occurred mostly in stress analysis. Some of the first uses of moiré to measure surface topography were reported by Meadows et al., (1970), Takasaki (1970), and Wasowski (1970). Moiré has also been used to compare an object to a master and for vibration analysis (Der Hovanesian and Yung, 1971; Gasvik, 2002). A theoretical review and experimental comparison of moiré and projection techniques for contouring is given by Benoit et al. (1975). Automatic computer fringe analysis of moiré patterns by finding fringe centers was reported by Yatagai et al. (1982). Heterodyne interferometry was first used with moiré fringes by Moore and Truax (1979), and phase measurement techniques were further developed by Perrin and Thomas (1979), Shagari (1980), and Reid (1984b). Review papers on moiré techniques include Post (1982), Reid (1984a), and Halioua and Liu (1989) and recent books include Paturski and Kujawinska (1993), Post et al. (1997), Amridror (2000), and Walker (2004).

The projection of interference fringes for contouring objects was first proposed by Rowe and Welford (1967). Their later work included a number of applications for projected fringes (Welford, 1969) and the use of projected fringes with holography (Rowe, 1971). In-depth mathematical treatments have been provided by Benoit et al. (1975) and Gasvik (2002). The relationship between projected fringe contouring and triangulation is given in a book chapter by Case et al. (1987). Heterodyne phase measurement was first introduced with projected fringes by Indebetouw (1978), and phase measurement techniques were further developed by Takeda, Ina, and Kabayashi (1982), Takeda and Mutoh (1983), and Srinivasan, Liu, and Halioua (1984, 1985). Today phase measurement techniques are the norm as described in the list of recent books listed above.

Haines and Hildebrand first proposed contouring objects in holography using two sources (Haines and Hildebrand, 1965; Hildebrand and Haines, 1966, 1967). These two sources were produced by changing either the angle of the illumination beam on the object or the angle of the reference beam. A small angle difference between the beams used to produce a double-exposure hologram creates a moiré in the final hologram which corresponded to topographic contours of the test object. Further insight into two-angle holography has been provided by Menzel (1974), Abramson (1976a,b), and DeMattia and Fossati-Bellani (1978). The technique has also been used in speckle interferometry (Winther, 1983). These holographic and speckle techniques are described more in the second half of this chapter.

Since all of these techniques are so similar, it is sometimes hard to differentiate developments in one technique versus another. MacGovern (1972) provided a theory that linked all of these techniques together. The next part of this chapter will explain each of these techniques and then show the similarities among all of these techniques and provide a comparison to conventional interferometry.

16.1.5. Fringe Projection

A simple approach for contouring is to project interference fringes or a grating onto an object and then view from another direction. Figure 16.10 shows the optical

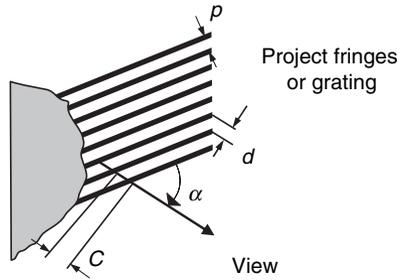


FIGURE 16.10. Projection of fringes or grating onto object and viewed at an angle α . p is the grating pitch or fringe spacing and C is the contour interval.

setup for this measurement. Assuming a collimated illumination beam and viewing the fringes with a telecentric optical system, straight equally-spaced fringes are incident on the object producing equally-spaced contour intervals. The departure of a viewed fringe from a straight line shows the departure of the surface from a plane reference surface. An object with fringes of spacing p projected onto it can be seen in Figure 16.11. When the fringes are viewed at an angle α relative to the projection direction, the spacing of the lines perpendicular to the viewing direction will be

$$d = \frac{p}{\cos \alpha} \quad (16.19)$$

The contour interval C (the height between adjacent contour lines in the viewing direction) is determined by the line or fringe spacing projected onto the surface and the angle between the projection and viewing directions;

$$C = \frac{p}{\sin \alpha} = \frac{d}{\tan \alpha} \quad (16.20)$$

These contour lines are planes of equal height and the sensitivity of the measurement is determined by α . The larger the angle α , the smaller the contour interval. If $\alpha = 90^\circ$, then the contour interval is equal to p , and the sensitivity is a maximum. The reference plane will be parallel to the direction of the fringes and perpendicular to the viewing direction as shown in Figure 16.12. Even though the maximum sensitivity can be obtained, a 90° angle between the projection and viewing directions will produce a lot of unacceptable shadows on the object. These shadows will lead to areas with missing data where the object cannot be contoured. When $\alpha = 0$, the contour interval is infinite, and the measurement sensitivity is zero. To provide the best results, an angle no larger than the largest slope on the surface should be chosen.

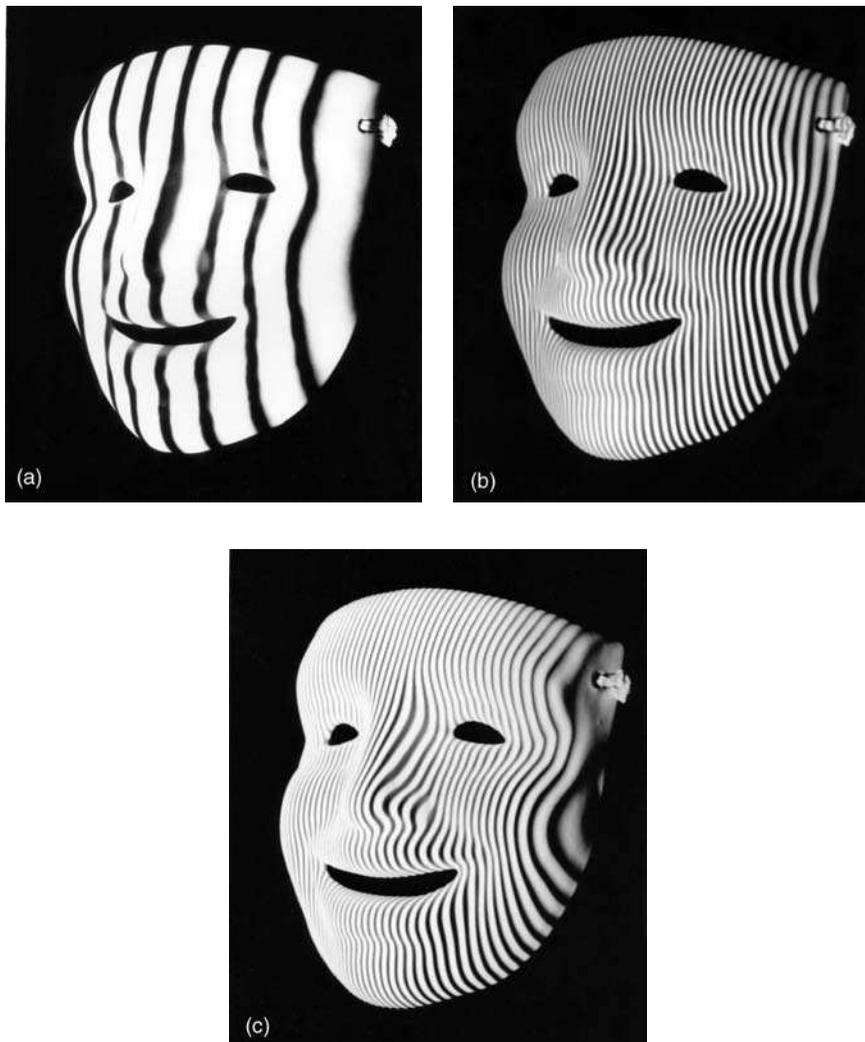


FIGURE 16.11. Mask with fringes projected onto it. (a) Coarse fringe spacing. (b) Fine fringe spacing. (c) Fine fringe spacing with an increase in the angle between illumination and viewing.

When interference fringes are projected onto a surface rather than using a grating, the fringe spacing p is determined by the geometry shown in Figure 16.13 and is given by

$$p = \frac{\lambda}{2 \sin \Delta\theta} \quad (16.21)$$

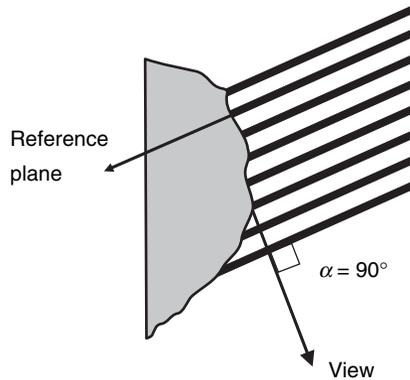


FIGURE 16.12. Maximum sensitivity for fringe projection with a 90° angle between projection and viewing.

where λ is the wavelength of illumination and $2\Delta\theta$ is the angle between the two interfering beams. Substituting the expression for p into Eq. (16.20), the contour interval becomes

$$C = \frac{\lambda}{2(\sin \Delta\theta) \sin \alpha} \quad (16.22)$$

If a simple interferometer such as a Twyman–Green is used to generate projected interference fringes, tilting one beam with respect to the other will change the contour interval. The larger the angle between the two beams, the smaller the contour interval will be. Figures 16.11(a and b) show a change in the fringe spacing for interference fringes projected onto an object. The direction of illumination has been moved away from the viewing direction between Figures 16.11(b and c). This increases the angle α and the test sensitivity while reducing the contour interval. Projected fringe contouring has been covered in detail by Gasvik (2002).

If the source and the viewer are not at infinity, the fringes or grating projected onto the object will not be composed of straight, equally-spaced lines. The height between contour planes will be a function of the distance from the source and viewer to the object. There will be a distortion due to the viewing of the fringes as well as due to the illumination. This means that the reference surface will not be a plane. As long as the object does not have large height changes compared to the illumination and viewing distances, a plane reference surface placed in the plane of the object can be

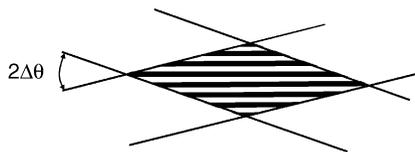


FIGURE 16.13. Fringes produced by two interfering beams.

measured first and then subtracted from subsequent measurements of the object. This enables the mapping of a plane in object space to a surface which will serve as a reference surface. If the object has large height variations, the plane reference surface may have to be measured in a number of planes to map the measured object contours to real heights. Finite illumination and viewing distances will be considered in more detail with shadow moiré in Section 16.1.6.

Fringe Projection using Microdisplays. For many years fringe projection methods relied on Ronchi gratings, which were often made as imprinted chrome lines on a glass substrate, but present-day systems employ a number of different types of microdisplays (digital light projectors); three commonly-used types of microdisplays (Armitage et al., 2002), namely micro-electro-mechanical-systems (MEMS), liquid-crystal and electroluminescent technologies, allow for active addressing of individual pixels high resolution matrix display area. The first type, which includes digital micromirror displays (DMD – Texas Instrument trademark), uses an array of individual, approximately $13\ \mu\text{m}$ square mirrors that are switched on and off at different frequencies so as to obtain different levels of projected light. The second type of microdisplays are liquid crystal displays (LCD), where standard LCDs are built of twisted nematic liquid crystal layers and work in transmission. The newer type of LCDs is based on ferroelectric crystal placed on silicon (LCoS) and works in reflection. LCD displays act as spatial light modulators (SLM) and require incident polarized light. The third type is microdisplay made of an array of organic (polymer) light emitting diodes (OLEDS or PLEDS). This type is best suited for small systems because pixels made of organic polymers are Lambertian emitters itself and do not require additional illuminator. The advantage of any fringe projection system using microdisplays controlled by computer is that they do not require a mechanical phase shifting grating and the type of projected fringes can be changed with a mouse click. Fringes are changed by addressing pixels of microdisplay. Additional advantage of microdisplays is that the period and brightness (Kowarschik et al., 2000; Proll et al., 2003) of light patterns can be adapted to the type of object and also patterns can be displayed in different colors allowing for simultaneous collection of three patterns with color CCD camera. Many authors have analyzed performance of microdisplays in fringe projection for shape measurement (Frankowski et al., 2000; Proll et al., 2003; Notni, Riehemann et al., 2004).

16.1.6. Shadow Moiré

A simple method of moiré interferometry for contouring objects uses a single grating placed in front of the object as shown in Figure 16.14. The grating in front of the object produces a shadow on the object which is viewed from a different direction through the grating. A low frequency beat or moiré pattern is seen. This pattern is due to the interference between the grating shadows on the object and the grating as viewed. Assuming that the illumination is collimated and that the object is viewed at infinity or through a telecentric optical system, the height z between the grating and

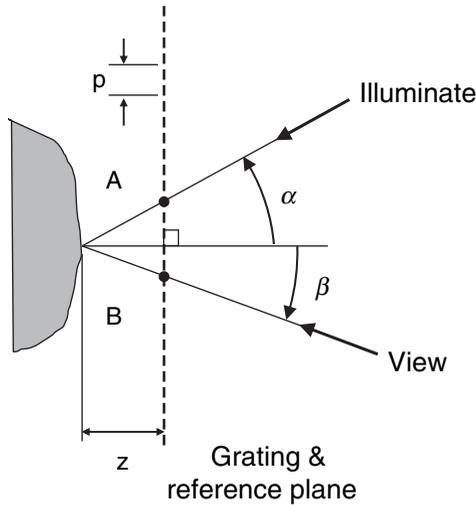


FIGURE 16.14. Geometry for shadow moiré with illumination and viewing at infinity, that is, parallel illumination and viewing.

the object point can be determined from the geometry shown in Figure 16.14 (Meadows et al., 1970; Takasaki, 1973; Chiang, 1983). This height is given by

$$z = \frac{Np}{\tan \alpha + \tan \beta} \quad (16.23)$$

where α is the illumination angle, β is the viewing angle, p is the spacing of the grating lines, and N is the number of grating lines between the points A and B (see Fig. 16.14). The contour interval in a direction perpendicular to the grating will simply be given by

$$C = \frac{p}{\tan \alpha + \tan \beta} \quad (16.24)$$

Again, the distance between the moiré fringes in the beat pattern depends upon the angle between the illumination and viewing directions. The larger the angle, the smaller the contour interval. If the high frequencies due to the original grating are filtered out, then only the moiré interference term is seen. The reference plane will be parallel to the grating. Note that this reference plane is tilted with respect to the reference plane obtained when fringes are projected onto the object. Essentially, the shadow moiré technique provides a way of removing the “tilt” term and repositioning the reference plane. The contour interval for shadow moiré is the same as that calculated for projected fringe contouring [Eq. (16.20)] when one of the angles is zero with $d = p$. Figure 16.15 shows an object which has a grating sitting in front of it. An illumination beam is projected from one direction and viewed

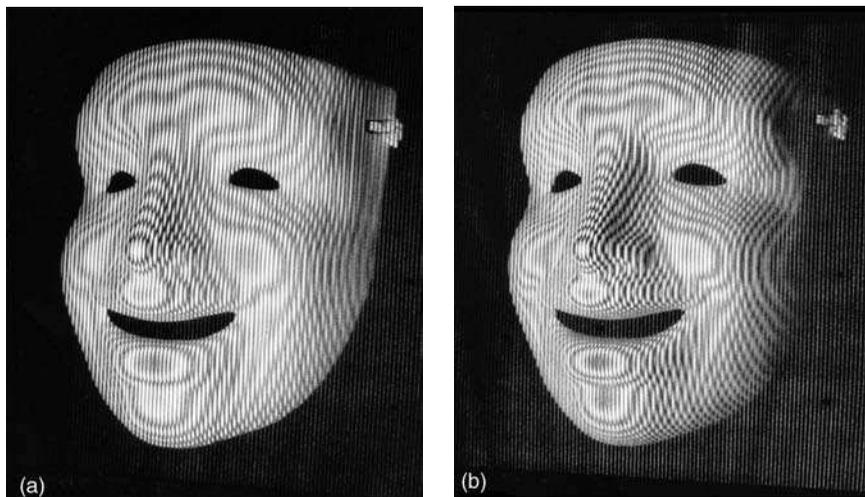


FIGURE 16.15. Mask with grating in front of it. (a) One viewing angle. (b) Larger viewing angle.

from another direction. Between Figures 16.15a and b, the angles α and β have been increased. This has the effect of decreasing the contour interval, increasing the number of fringes, and rotating the reference plane slightly away from the viewer.

Most of the time, it is difficult to illuminate an entire object with a collimated beam. Therefore, it is important to consider the case of finite illumination and viewing distances. It is possible to derive this for a very general case (Meadows, Johnson, and Allen, 1970; Takasaki, 1970; Bell, 1985); however, for simplicity, only the case where the illumination and viewing positions are the same distance from the grating will be considered. Figure 16.16 shows a geometry where the distance between the illumination source and the viewing camera is given by w , and the distance between these and the grating is l . The grating is assumed to be close enough to the object surface so that diffraction effects are negligible. In this case, the height between the object and the grating is given by

$$z = \frac{Np}{\tan \alpha' + \tan \beta'} \quad (16.25)$$

where α' and β' are the illumination and viewing angles at the object surface. These angles change for every point on the surface and are different from α and β in Figure 16.16 which are the illumination and viewing angles at the grating (reference) surface. The surface height can also be written as (Meadows et al., 1970; Takasaki, 1973; Chiang, 1983)

$$z = NC(z) = \frac{Np(l+z)}{w} = \frac{Npl}{w - Np} \quad (16.26)$$

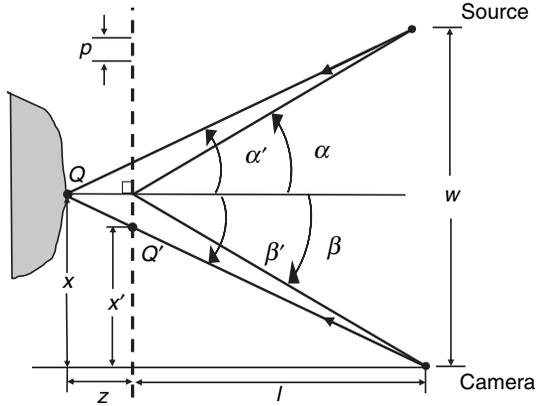


FIGURE 16.16. Geometry for shadow moiré with illumination and viewing at finite distances.

This equation indicates that the height is a complex function depending upon the position of each object point. Thus, the distance between contour intervals is dependent upon the height on the surface and the number of fringes between the grating and the object. Individual contour lines will no longer be planes of equal height. There are now surfaces of equal height. The expression for height can be simplified by considering the case where the distance to the source and viewer is large compared to the surface height variations, $l \gg z$. Then the surface height can be expressed as

$$z = \frac{Npl}{w} = \frac{Np}{\tan \alpha + \tan \beta} \quad (16.27)$$

Even though the angles α and β vary from point-to-point on the surface, the sum of their tangents remains equal to w/l for all object points as long as $l \gg z$. The contour interval will be constant in this regime and will be the same as that given by Eq. (16.24).

Because of the finite distances, there is also distortion due to the viewing perspective. A point on the surface Q will appear to be at the location Q' when viewed through the grating. By similar triangles, the distances x and x' from a line perpendicular to the grating intersecting the camera location can be related using

$$\frac{x}{z+l} = \frac{x'}{l} \quad (16.28)$$

where x and x' are defined in Figure 16.16. Equation (16.28) can be rearranged to yield the actual coordinate x in terms of the measured coordinate x' and the

measurement geometry,

$$x = x' \left(1 + \frac{z}{l} \right) \quad (16.29)$$

Likewise, the y coordinate can be corrected using

$$y = y' \left(1 + \frac{z}{l} \right) \quad (16.30)$$

This enables the measured surface to be mapped to the actual surface to correct for the viewing perspective. These same correction factors can be applied to fringe projection.

16.1.7. Projection Moiré

Moiré interferometry can also be implemented by projecting interference fringes or a grating onto an object and then viewing through a second grating in front of the viewer (see Fig. 16.17) (Brooks and Helfinger, 1969). Instead of using a second grating to observe moiré fringes, the spacing of pixels on a digital camera can be used if the pitch is close to the observed fringe spacing (Bell 1985).

The difference between projection and shadow moiré is that in projection moiré two different gratings are used in projection moiré. The orientation of the reference plane can be arbitrarily changed by using different grating pitches to view the object. The contour interval is again given by Eq. (16.24), by substituting a period of $d = p / \cos a$ for p where a is the angle of illumination direction. Fringes of spacing p or a grating of pitch p perpendicular to the direction of illumination will have a period of $d = p / \cos a$ in the y plane (see Fig. 16.16). As long as the grating pitches are matched for both illumination and viewing to have the same value of d in the y plane then the contour interval can be found using Eq. (16.24) with d substituted for p . This implementation makes projection moiré the same as shadow moiré, although

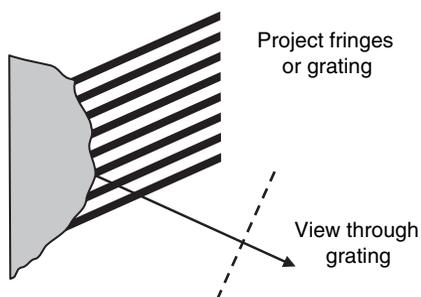


FIGURE 16.17. Projection moiré where fringes or a grating are projected onto a surface and viewed through a second grating.

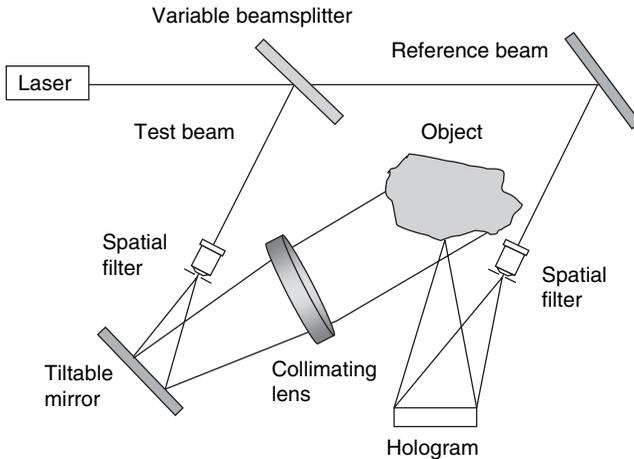


FIGURE 16.18. Setup for two-angle holographic interferometry.

projection moiré can be much more complicated than shadow moiré. A good theoretical treatment of projection moiré is given by Benoit et al. (1975).

16.1.8. Two-angle Holography

Projected fringe contouring can also be done using holography. First a hologram of the object is made using the optical setup shown in Figure 16.18. Then the direction of the beam illuminating the object is changed slightly. When the object is viewed through the hologram, interference fringes are seen which correspond to the interference between the wavefront stored in the hologram and the live wavefront with the tilted illumination. This process is depicted by Figure 16.19. These fringes are exactly what would be seen if the object were illuminated with the two illumination beams simultaneously. The beams would be tilted with respect to one another by the same amount that the illumination beam was tilted after making the hologram. These

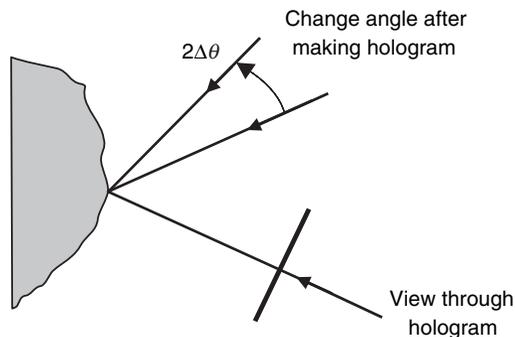


FIGURE 16.19. Two-angle holographic interferometry. Get interference fringes from shifting the illumination beam.

fringes will look the same as those produced by projected fringe contouring and shown in Figure 16.11. To produce straight, equally-spaced fringes, the object illumination should be collimated. The surface contour is measured relative to a surface which is a plane when collimated illumination is used. The theory of projected fringe contouring can be applied to two-angle holographic contouring yielding a contour interval given by Eq. (16.22), where $2\Delta\theta$ is the change in the angle of the object illumination. More detail on two-angle holographic contouring can be found in Haines and Hildebrand (1965), Hildebrand and Haines (1967, 1966), Vest (1979), and Hariharan (1984).

Surface contours can also be obtained using the digital holography and speckle interferometry techniques described in the second half of this chapter by utilizing this type of optical setup and taking data from two different angles. For those techniques, quantitative precisions of 1/100th of the contour interval are obtainable after system calibration.

16.1.9. Common Features

All of the techniques described produce fringes corresponding to contours of equal height on the object. They all have a similar contour interval determined by the fringe spacing or grating period and the angle between the illumination and viewing directions as long as the illumination and viewing are collimated.

Extracting Quantitative Information from Fringe Projection and Moiré Techniques. Phase-shifting techniques (see Chapter 14) can be applied to any of the techniques to produce quantitative height information as long as sinusoidal gratings or fringes are used. The surface heights measured are relative to a reference surface which is a plane as long as the fringes or grating lines are straight and equally spaced at the object. The only difference between the moiré techniques and the projected fringes and two-angle holography is the change in the location of the reference plane. If the fringes are digitized or phase-measuring interferometry techniques are applied, the reference plane can be changed in the computer mathematically.

The precision of these contouring techniques depends upon the number of fringes used. When the fringes are digitized using fringe following techniques, the surface height can be determined to 1/10 of a fringe. If phase-measurement is used, the surface heights can be determined to 1/100 of a fringe. Therefore it is advantageous to use as many fringes as possible. And because a reference plane can easily be changed in a computer by changing the amount of tilt subtracted, projected fringe contouring is the simplest way to contour an object interferometrically.

16.1.10. Comparison to Conventional Interferometry

The measurement of surface contour can be related to making the same measurement using a Twyman–Green interferometer assuming a long effective wavelength. The loci of the lines or fringes projected on to the surface (assuming illumination and

viewing at infinity) is given by

$$y = z \tan \alpha + nd \quad (16.31)$$

where z is the height of the surface at the point y , d is the fringe spacing measured along the y axis, and n is an integer referring to fringe order number. If the same surface were tested using a Twyman–Green interferometer, a bright fringe would be obtained whenever

$$2z - y \sin \gamma = n\lambda \quad (16.32)$$

where λ is the wavelength and γ is the tilt of the reference plane. By comparing Eqs. (16.31) and (16.32), it can be seen that they are equivalent as long as

$$d = \frac{\lambda_{\text{effective}}}{\sin \gamma} \quad (16.33)$$

and

$$\frac{2}{\sin \gamma} = \tan \alpha \quad (16.34)$$

where $\lambda_{\text{effective}}$ is the effective wavelength. The effective wavelength can then be written as

$$\lambda_{\text{effective}} = 2C = \frac{2d}{\tan \alpha} = \frac{2p}{\cos \alpha \tan \alpha} \quad (16.35)$$

where C is the contour interval as defined in Eq. (16.20). Thus, contouring using these techniques is similar to measuring the object in a Twyman–Green interferometer using a source with wavelength $\lambda_{\text{effective}}$.

16.1.11. Coded and Structured Light Projection

A method that is often used in place of moiré methods combines projection of multiple binary grey code patterns and sinusoidal fringes. Projection of grey code patterns was used in photogrammetry, and now it is used in fringe projection in order to resolve unwrapping ambiguities and extending the range of PSI methods used in fringe projection (Reich et al., 2000; Huang and Zhang, 2005) and can be utilized adaptively to follow objects in real time (Koninckx and Van Gool, 2006). The use of multiple patterns with different frequencies is analogous to using multiple wavelength sources in conventional interferometric techniques to resolve phase ambiguities. The mathematical merge of photogrammetry methods with fringe projection methods can be called phasogrammetry or phase value photogrammetry. A large number of different coding strategies for structure light projection have been proposed (Salvi et al., 2004) for applications including machine vision, industry

inspection, reverse engineering, rapid prototyping, biomedicine, and art. Advances in image processing and fringe projection techniques have enabled great strides in the measurement of object shape. These techniques provide realistic data that can follow motion almost in a real time. They have created many new applications in the multimedia industry for computer graphics and animation as well as for virtual reality and facial recognition. Recent reviews of various applications can be found in (Kujawińska and Malacara, 2001; Harding, 2005; D'Apuzzo, 2006).

16.1.12. Applications

All these techniques can be used for displacement measurement or stress analysis as well as for contouring objects. Displacement measurement is performed by comparing the fringe patterns obtained before and after a small movement of the object or before and after applying a load to the object (similar to holographic interferometry techniques described in Section 16.2). Because the sensitivity of these tests are variable, they can be used for a larger range of displacements and stresses than the holographic techniques. Differential interferometry comparing two objects or an object and a master can also be performed by comparing the two fringe patterns obtained. Finally, time-average vibration analysis can also be performed with moiré yielding results similar to those obtained with time-average holography with a much longer effective wavelength (see Section 16.2.2.2).

Using phase-measurement techniques, the surface height relative to some reference surface can be obtained quantitatively. If the contour lines are straight and equally spaced in object space, then the reference plane will be a plane. In the computer any plane (or surface) desired can be subtracted from the surface height to yield the surface profile relative to any plane. This is similar to viewing the contour lines through a grating (or deformed grating) to reduce their number. If the contour lines are not straight and equally spaced, the reference surface will be something other than a plane. The reference surface can be determined by placing a flat surface at the location of the object and measuring the surface height. Once this reference surface is measured, it can be subtracted from subsequent measurements to yield the surface height relative to a plane surface. Thus, with the use of phase-measuring interferometry techniques, the surface height can be made relative to any surface and transformed to surface heights relative to another surface. Taking this one step further, a master component can be compared to a number of test components to determine if their shape is within the specification. It should also be pointed out that this measurement is sensitive to a certain direction, and that there may be areas where data are missing because of shadows on the surface.

As an example, Figure 16.20 shows the mask of Figures 16.11 and 16.15 contoured using fringe projection and phase-measurement interferometry. The fringes are produced using a Twyman–Green interferometer with a He–Ne laser. A high-resolution camera with 1320×900 pixels and a zoom lens is used to view the fringes. Surface heights are calculated using phase-measurement techniques at each detector point. A total of five interferograms were used to calculate the surface shown in Figure 16.20(a). The best fit plane has been subtracted from the surface to yield

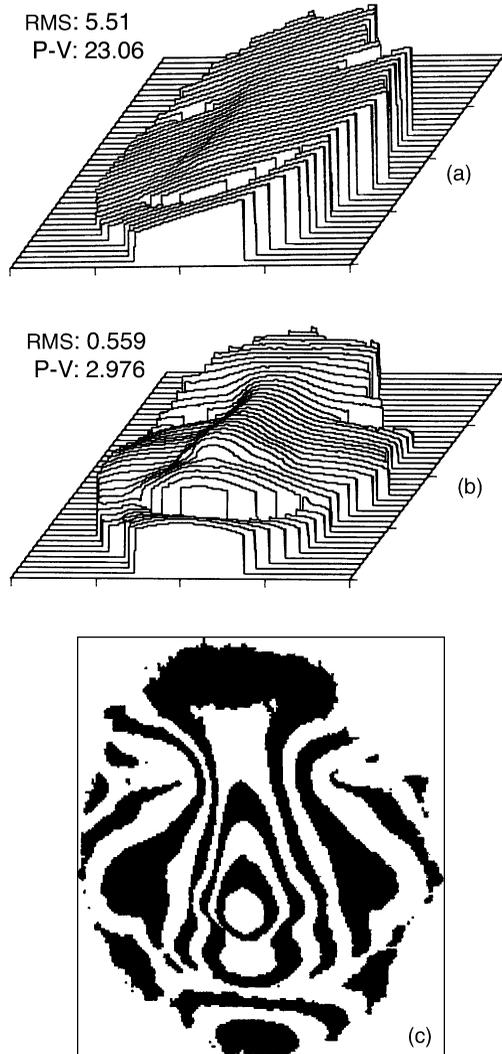


FIGURE 16.20. Mask measured with projected fringes and phase-measurement interferometry. (a) Isometric plot of measured surface height. (b) Isometric plot after best-fit plane removed. (c) Two-dimensional contour plot after best-fit plane removed. Units on plots are in numbers of contour intervals. One contour interval is approximately 10 mm. The surface is about 150 mm in diameter.

Figure 16.20(b). In this way, the reference plane has been changed. Figure 16.20(c) shows a two-dimensional contour map of the object after the best-fit plane is removed. These contours can also be thought of as the fringes which would be viewed on the object. Figure 16.20c shows the fringes with a second reference grating chosen to minimize the fringe spacing. The contour interval for this example is 10 mm, and the total peak-to-valley height deviation after the tilt is subtracted is about 30 mm.

16.1.13. Summary

The techniques of projected fringe contouring, projection moiré, shadow moiré, and two-angle holographic contouring are all similar. They all involve projecting a pattern of lines or interference fringes onto an object and then viewing those contour lines from a different direction. In the case of the moiré techniques, the contour lines are viewed through a grating to reduce the total number of fringes. In all of the techniques, the surface height is measured relative to a reference surface. The reference surface will be a plane if the projected grating lines or interference fringes are straight and equally spaced at the object and viewed at infinity or with a telecentric imaging system. The use of the second grating in the moiré techniques changes the reference plane, but does not affect the contour interval. The sensitivity of the techniques is maximum when the contour lines are viewed at an angle of 90° with respect to the projection direction. Quantitative data can be obtained from any of these techniques using phase-measurement interferometry techniques. The precision of the surface height measurement will depend upon the number of fringes present. Surface height measurements can be made with a repeatability of 1/100 of a contour interval RMS. Thus, the number of fringes used should be as many as can easily be measured. The contour interval can be changed to increase the number of fringes, and once the surface height is calculated, a reference surface can be subtracted in the computer to find the surface height relative to any desired surface.

16.2. HOLOGRAPHIC AND SPECKLE TESTS

16.2.1. Introduction

Diffusely reflecting or polished surfaces that are subject to stress can be interferometrically compared with their normal states using holographic interferometry techniques. Direct measurements as well as shearing methods are possible for static, dynamic and time-average measurements.

Traditional holographic interferometry techniques require the recording of an intermediate hologram of a previous or ideal object state to compare with the current state or shape of the object. Speckle interferometry and digital holography techniques do not require the use of an intermediate holographic recording as the phase data are directly recorded and reconstructed using electronic or digital techniques. These techniques can provide interference fringes corresponding to a change in the object shape or displacement.

This chapter describes holographic and speckle techniques useful in optical testing. This subject has such a large amount of information to cover that emphasis is placed on basic descriptions of some of the techniques, and examples showing results of measurements using these tests. Since the publication of the 2nd edition of this book, several books have been published that cover this topic more in depth (Jacquot and Fournier, 2000; Gasvik, 2002; Gastinger et al., 2003; Steinchen and Yang, 2003; Yaroslavsky, 2004; Kreis, 2005; Mix, 2005; Schnars and Jueptner,

2005). The reader may look into these references for more in-depth mathematics and descriptions of these techniques.

16.2.2. Holographic Interferometry for Nondestructive Testing

Holographic interferometry techniques have been in use for more than 4 decades in the field of stress analysis. Historical reviews can be found in Vest (1979), Gasvik (Gasvik 2002), and Kreis (Kreis 2005). There are two basic types of techniques: static and time average. Static tests, such as double-exposure holography and real-time holographic interferometry, measure an object in two different states of applied stress and find the difference between them. In these tests, the object is assumed to not be moving during the exposure time. A single exposure of the object is made before stress is applied, and a second single measurement is made after the stress is applied. Static tests can also be performed using pulsed lasers to freeze motion and obtain dynamic measurements (Vest, 1979; Gasvik, 2002; Kreis, 2005), where the length of the pulse is small compared to the change in object motion. For time-average techniques, the object is excited at some vibrational frequency providing a periodic motion. A single measurement averaged over multiple periods of vibration is evaluated.

The optical layout shown in Fig. 16.21 is used for holographic nondestructive testing to measure out-of-plane displacement. It can be modified to measure in-plane displacement as well as out-of-plane displacement and to do shearing holographic interferometry measurements. A laser beam is split into two paths using a variable-density beamsplitter to control the ratio of the amount of light in one beam with respect to the other beam. The object beam illuminates the object with a diverging beam of light. The object, usually diffuse, scatters the light, and some of this light is

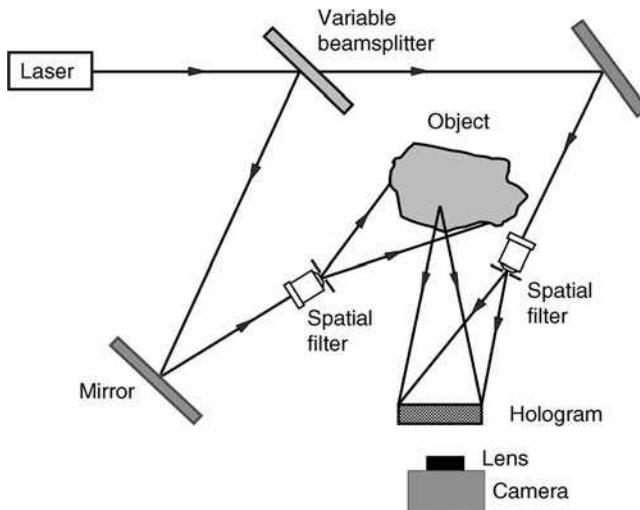


FIGURE 16.21. Holographic interferometer for nondestructive testing.

incident upon the hologram plane. The other beam is a reference beam directly incident upon the hologram plane. The angle between the object beam and the reference beam at the hologram plane will determine the spacing of the interference fringes in the hologram. Because the hologram is simply the interference of two beams of light, the difference between the path lengths of the object and reference beams must be within the coherence length of the laser being used. If an argon-ion laser with an etalon or a single-mode helium–neon laser is used, a coherence length of many meters is attainable; whereas, if a multi-mode helium–neon laser is used, the path lengths must be within a few centimeters of one another. To make an efficient hologram, the reference beam should be 6–8 times brighter than the object beam at the hologram plane, the polarizations of the two beams should be in the plane of incidence, and the angle between the object and reference beams needs to be small enough to produce resolvable interference fringes in the recording material.

Holographic interferometry techniques can be used to measure thermal changes in optical elements, changes in optical surface shape due to mounting, changes in deformable mirror shapes, and to study vibrational modes of optical elements, mounts, or entire optical systems. These techniques can be used with diffuse (ground) surfaces, specularly reflecting surfaces, and transmissive optical components. One particular application of holographic nondestructive testing is to determine the mechanical and thermal properties of large unworked mirror blanks before any time and money are spent to put optical surfaces on the blanks (Van Deelen and Nisenson 1969). Both qualitative fringe data and quantitative displacement data may be obtained.

Static Holographic Interferometry. For a static holographic measurement of an object, a hologram of the object is made while the object is in one stress state. In order for static measurements to work, the object must not move while the hologram is being made. After the object stress state is changed, either a second hologram is exposed as in double-exposure holographic interferometry, or the object is observed through the hologram as in real-time holographic interferometry. In both cases, there is a secondary interference between the wavefront generated before the change in applied stress and the wavefront generated after the change in applied stress. For double-exposure holography, both wavefronts are stored in the hologram. When the hologram is illuminated with the reference wavefront, both wavefronts are reconstructed. When these wavefronts are viewed through the hologram, cosinusoidal secondary interference fringes are visible. The secondary interference fringes correspond directly to the amount of object displacement between the two exposures. If there is no change in the object between exposures, a single interference fringe will be produced. Each fringe in the secondary interference pattern indicates one wave of displacement along a direction bisecting the illumination and viewing directions of the object.

For real-time holographic interferometry, one of the two wavefronts is stored in the hologram, and the other wavefront is produced live by the test object. Once the hologram is made and developed, care must be made to ensure that the hologram is replaced in the same location so that the wavefront stored in the hologram can be

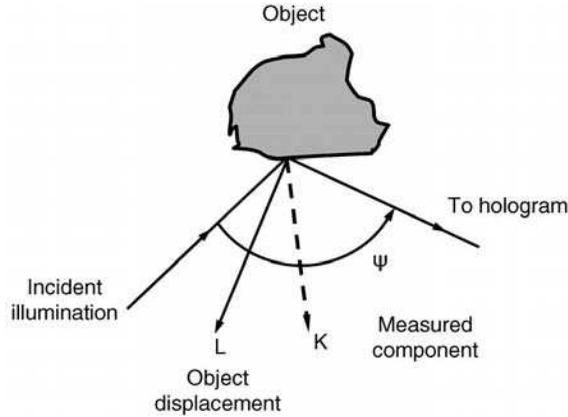


FIGURE 16.22. Recording geometry showing sensitivity vector and displacement vector.

interfered with the live wavefront. Real-time holographic interferometry is very similar to the use of the holographic test plate. Fringe location, fringe shape, and test sensitivity are the same for real-time techniques as they are with double-exposure techniques. If the object is in motion, static techniques can be used to provide cosinusoidal fringes as long as a pulsed laser or high-speed shutter is used to freeze the motion. The exposure time must be short enough that the object does not move during it.

Figure 16.22 shows the recording geometry indicating the object displacement vector \mathbf{L} , and the sensitivity vector \mathbf{K} . The sensitivity vector is defined as the direction along which the object displacement is measured. The direction of the sensitivity vector can vary across the object surface if the field of view is large, and if the source illuminating the object and the viewing position are not located at infinity. Mathematically, the secondary interference fringes for a static holographic measurement at a single point in the viewing plane can be written as

$$I = I_0(1 + \gamma \cos \Delta\phi) \quad (16.36)$$

where I_0 is the dc intensity, γ is the fringe visibility, and $\Delta\phi$ is the phase of the secondary interference fringes corresponding to the difference (displacement) between the two object states. The phase difference can be written as

$$\Delta\phi = \mathbf{K} \cdot \mathbf{L} \quad (16.37)$$

The displacement of the object at a point x, y in the direction of the sensitivity vector is given by

$$D(x, y) = \frac{\Delta\phi(x, y)\lambda}{4\pi \cos(\psi/2)} \quad (16.38)$$

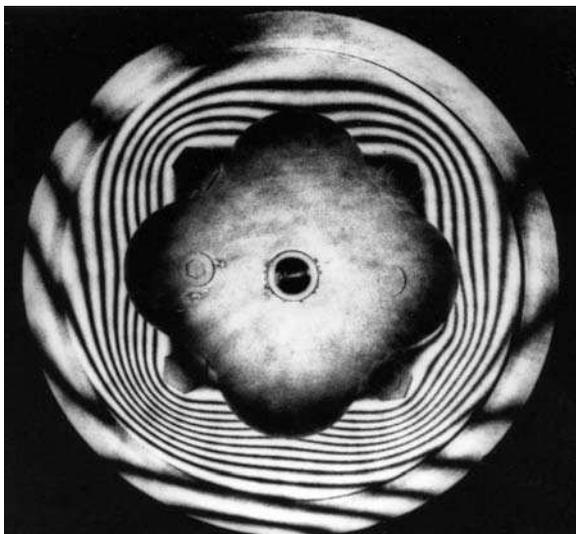


FIGURE 16.23. Double-exposure holographic interferometry fringes of helicopter coupling flange. (Courtesy K. A. Stetson, United Technologies Research Center.)

where λ is the illumination wavelength and ψ is the angle between the illumination and viewing directions. The displacement measurement is usually a combination of in-plane (along the surface of object) and out-of-plane (perpendicular to the object surface) displacement. Specific setups can be made to measure only one of these components. If all three components of displacement (x , y , and z) are desired, three measurements must be made (Pryputniewicz and Stetson, 1976; Stetson, 1979, 1990; Nakadate et al., 1981; Kakunai et al., 1985; Hariharan et al., 1987). As long as the change in the object shape is small, the secondary interference fringes will be localized at the object. If the applied stress causes the object to move (rigid-body translation or rotation) as well as deform, the secondary interference fringes may not be localized at the object (Vest, 1979; Hariharan, 1984).

An example showing the use of static holographic interferometry to test a coupling flange from a helicopter is shown in Figure 16.23. This part was stressed between exposures by applying a static force between the rim and the center of the part.

In addition to displacement measurement and stress analysis, real-time or double-exposure holographic interferometry can be used to contour objects with two wavelengths, two indices of refraction, or two different angles of object illumination (Hariharan, 1984).

Time-Average Holographic Interferometry. If the object is excited in a periodic motion, a measurement can be made using a single exposure which averages over many vibrational periods creating a *time-average hologram* (Powell and Stetson,

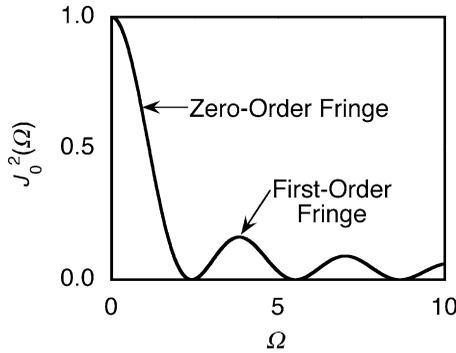


FIGURE 16.24. Time-average holography (Bessel function) fringes.

1965; Stetson and Powell, 1965; Vest, 1979; Hariharan, 1984). The secondary interference fringes at a single point in the viewing plane for this case can be written as

$$I = I_0 \gamma |M(\Omega)|^2 \quad (16.39)$$

where $M(\Omega)$ is the secondary fringe function due to the object motion. When the object is moving sinusoidally,

$$M(\Omega) = J_0(\Omega) \quad (16.40)$$

where J_0 is a zero-order Bessel function, and $\Omega = \mathbf{K} \cdot \mathbf{L} \propto A/\lambda$ for a vibration amplitude A and illumination wavelength λ . Note that this result is independent of the vibration excitation frequency. The time-average secondary interference fringes for a sinusoidal vibration become

$$I = I_0 \gamma J_0^2(\Omega) \quad (16.41)$$

These fringes are shown in Figure 16.24. For positions on the object which are stationary, the intensity is a maximum for the zero-order fringe. As the amplitude of the object movement increases, the intensity of the secondary interference fringes is reduced substantially. Figure 16.25 shows an example of a helicopter gear excited at 5812 Hz tested using time-average holography. It is obvious from the presence of four fringe orders in this example that the intensity of higher-order fringes diminishes rapidly.

Holographic Interferometry with Phase Measurement. Until the 1980s, holographic techniques provided only qualitative data in the form of fringe patterns requiring a skilled operator to interpret. With the advent of TV camera frame grabbers for personal computers and fast CPUs, quantitative data using phase-shifting interferometry techniques could be generated in less than a minute. To obtain phase data from object displacement, a phase shifter is placed in one beam

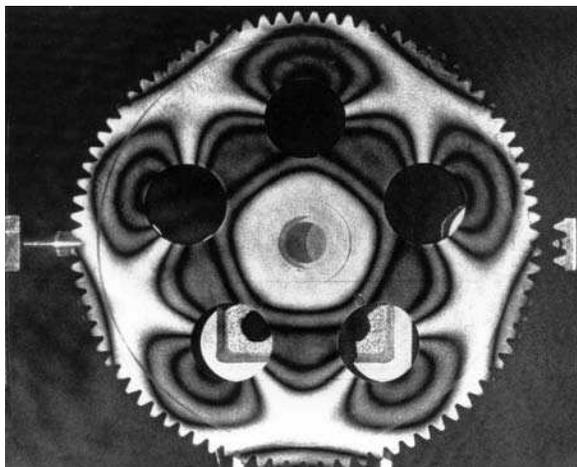


FIGURE 16.25. Time-average holographic fringes of helicopter gear excited at 5812 Hz. (Courtesy K. A. Stetson, United Technologies Research Center.)

of the interferometer shown in Figure 16.21. Standard phase-shifting techniques (as described in Chapter 14) can be used with holographic interferometry techniques which produce cosinusoidal fringes to generate a phase map of $\Delta\phi(x,y)$ corresponding to the object displacement $D(x,y)$ (Hariharan, 1984, 1985; Gasvik, 2002; Kreis, 2005).

An example of a phase measurement of out-of-plane displacement using real-time holographic interferometry to measure changes due to mechanical stress is shown in Figure 16.26. A hologram was made of a metal plate bolted on all four corners, and the phase measurement was performed after tightening a screw pushing on the back of the plate. The contour interval for Figure 16.26(a) is $0.3653\ \mu\text{m}$, and the angle between the illumination and viewing directions is 60° yielding a peak-to-valley displacement of $7.3\ \mu\text{m}$.

Quantitative data can also be extracted from time-average vibration fringes using phase-shifting techniques (Neumann et al., 1970; Stetson, 1982; Oshida et al., 1983; Nakadate and Saito, 1985; Stetson and Brohinsky, 1988). In one embodiment three separate phase measurements need to be taken (Stetson and Brohinsky, 1988). The first one is made with the object vibrating and shifting the relative phase between the object and reference beams as in standard phase-shifting interferometry. The second measurement is made by applying a vibration of the same frequency as the object vibration to a PZT in the reference beam of the interferometer. A bias phase between the object and reference vibration is added such that the relative phase difference between the vibrations is $+\pi/3$. Relative phase shifts between the object and reference beams are then applied and standard phase-shifting methods are used to calculate the phase. A third measurement is taken such that the relative phase difference between the object and reference vibrations is $-\pi/3$. Assuming a sinusoidal object vibration and 90° relative phase shifts for the phase calculations, one of the total of 12 frames of data

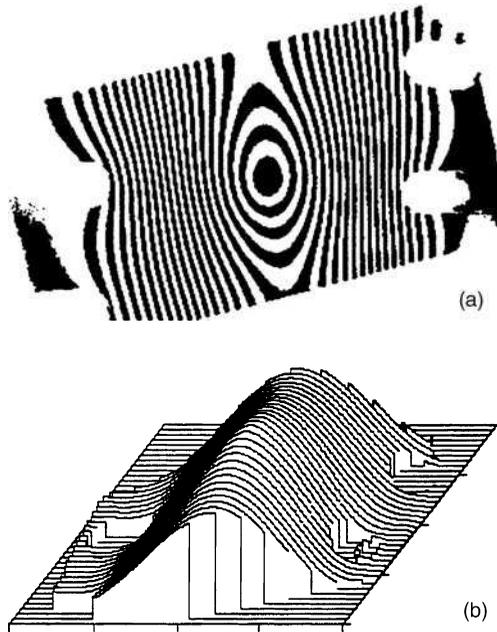


FIGURE 16.26. Out-of-plane displacement of a metal plate using phase-shifting interferometry techniques with real-time holographic interferometry. (a) Two-dimensional contour plot with a contour interval of $0.3653 \mu\text{m}$ and (b) isometric contour plot. The angle between the illumination and viewing directions is 60° yielding a peak-to-valley displacement of $7.3 \mu\text{m}$.

recorded can be written as

$$I_{ji} = I_0 [1 + \gamma \cos(\phi + \delta_i) J_0(\Omega + \beta_j)] \quad (16.42)$$

where $\delta_i = 0, \pi/2, \pi,$ and $3\pi/4$, and $\beta_j = -\pi/3, 0,$ and $\pi/3$. The amplitude of the vibration can then be calculated using

$$\Omega = \tan^{-1} \left[\frac{1}{\sqrt{3}} \left(\frac{H_1 - H_3}{2H_2 - H_1 - H_3} \right) \right] \quad (16.43)$$

where

$$H_1 = (I_{11} - I_{13})^2 + (I_{12} - I_{14})^2 = 4I_0^2 \gamma^2 J_0^2(\Omega - \pi/3) \quad (16.44)$$

$$H_2 = (I_{21} - I_{23})^2 + (I_{22} - I_{24})^2 = 4I_0^2 \gamma^2 J_0^2(\Omega) \quad \text{and} \quad (16.45)$$

$$H_3 = (I_{31} - I_{33})^2 + (I_{32} - I_{34})^2 = 4I_0^2 \gamma^2 J_0^2(\Omega + \pi/3) \quad (16.46)$$

Eq. (16.43) assumes the form of \cos^2 for the fringes. Because of this, a look-up table is necessary to find the difference between the $J_0^2(\Omega)$ and $\cos^2(\Omega)$ functions

(Stetson and Brohinsky, 1988). The error due to the difference between the $J_0^2(\Omega)$ and $\cos^2(\Omega)$ functions is dependent upon the fringe order, which can be determined from the H_2 measurement.

When the object under observation is not stable enough to capture consecutive exposures with different phase shifts, phase information can be obtained from a single interferogram using the Fourier-transform technique (Kreis, 1986, 1987) or using a multiplexing scheme encoding all the phase information for multiple phase shifts either onto multiple cameras (Koliopoulos, 1992) or onto a single camera (North-Morris et al., 2005). For the Fourier transform technique, the Fourier Transform of the fringe pattern is computed, one diffraction order is filtered out and then shifted to zero frequency, and then this single order is inverse Fourier transformed to obtain the phase distribution. This technique requires sinusoidal fringes and enough straight “tilt” fringes to separate orders. When multiple phase shifts are encoded within a single snapshot, dynamic and random motion can be followed and tracked quantitatively. These techniques are sensitive enough to be able to track thermal microbreezes related to pulse and respiration cycles in humans (Creath and Schwartz, 2005).

Digital Holographic Interferometry. In digital holographic interferometry, phase and amplitude information are captured electronically. The intermediate hologram is stored in digital memory and used to reconstruct the object after the second exposure is made (Yaroslavsky, 2004; Kreis, 2005; Schnars and Jueptner, 2005). All of the processing is done numerically and because of this these systems are more flexible and faster than using a separate holographic recording media. With the availability of megapixel cameras and sensors these types of measurements have become easier and easier. More about this topic is discussed in the next section.

16.2.3. Speckle Interferometry and Digital Holography

The first digital holography techniques were based on speckle interferometry (Dandliker, 2000; Joenathan and Tiziani, 2003). When TV cameras began being used and digital images stored and processed in hardware via electronics, these techniques were first named “electronic speckle pattern interferometry” (ESPI) and often called TV holography (Løkberg and Slettemoen, 1987; Jones and Wykes, 1983). Because of the low spatial frequency response of TV cameras these techniques measured in-line holograms (Gabor type holograms). The next step was to get quantitative information by applying phase shifting methods utilizing a computer with these ESPI techniques. These were called either “digital speckle pattern interferometry” or “phase shifting speckle interferometry” (Nakadate and Saito, 1985; Creath, 1985b, 1986; Robinson and Williams, 1986). As cameras with higher spatial resolutions became available it became possible to make the off-axis holograms that we are more familiar with. These techniques became known as digital holography (Kreis, 2005; Schnars and Jueptner, 2005) or phase shifting ESPI (Jacquot and Fournier, 2000; Gasvik, 2002). When encountering these techniques

in the literature there is not a common name, yet they are all based upon the same principles. Nowadays the term digital holography refers to any type of hologram stored and reconstructed digitally, and the term speckle interferometry also refers to the same systems where it is obvious that speckle is present. The mechanisms for creating interference are similar in both, and results are as well. What differentiates one technique from another is how well different spatial frequency structures are reconstructed, and how immune to noise the techniques are. Keep this in mind through the rest of this chapter where speckle interferometry and digital holography techniques are discussed.

Basics of Speckle Interferometry. When a laser beam is scattered off of a diffuse surface, the scattered light has a grainy appearance. The grains are an interference phenomenon known as *speckle*. The statistics of the speckle distribution depend upon the statistics of the object surface. If imaged by a lens, the speckle are said to be subjective, and the smallest speckle in the image will have a size equal to the Airy disk $2.44 \lambda/F\#$ generated by the optical system, where $F\#$ is the working F -number of the system (Goodman, 1975a). The intensity distribution and the statistics of the speckle pattern are an indication of the roughness of the surface used to generate the speckle pattern (Goodman, 1975a,b; Erf, 1978; Fujii and Lit, 1978). A speckle pattern generated by an object surface can be thought of as the object's fingerprint. When the object is perturbed in some way, the speckle pattern will change in a predictable way. Two different types of techniques are speckle photography and speckle interferometry (Ennos, 1975; Stetson, 1975; Pryputniewicz, 1985; Huntley, 1989). Both techniques involve the comparison of two or more speckle patterns. A recently published book by Goodman (2006) reviews the uses and properties of speckle.

Speckle interferometry usually includes a reference beam to enable measurement of the phase change in the speckles. In speckle interferometry, it is assumed that the speckles from one speckle pattern to another are correlated so that they do not shift by more than the diameter of a speckle between exposures. Speckle photography looks at the correlation between two speckle patterns where the fringes arise from a translation between exposures and can be used to measure larger displacements. The speckles from a small area of the two speckle patterns (translated with respect to one another) generate Young's fringes in the Fourier plane. For the rest of this discussion we are going to concentrate upon speckle interferometry techniques utilizing a reference beam.

To aid in understanding how speckle interferometry works, consider a diffuse object placed in the object beam of a Twyman–Green interferometer and a photographic plate placed at the image of the object with an $F\#$ large enough to produce visible speckle. After the speckle pattern is recorded and the photographic plate replaced in the original position, no light will get through the negative plate because the speckles line up with those stored in the photographic plate. This is similar to a single fringe using the holographic test plate. As the object is tilted, fringes will be seen which correspond to the tilt of the object. These fringes are modulated by a speckle pattern, and appear noisy. The secondary interference fringes are also known

as *speckle correlation fringes* because they correspond to the correlation between the speckles in the two fringe patterns (Jones and Wykes, 1983; Løkberg and Slettemoen, 1987). For speckle interferometry, the speckles are usually resolved by the recording system; however, this is not a necessary condition (Creath, 1985a).

Speckle interferometry techniques can be used for the same applications as holographic interferometry techniques (Dandliker, 2000; Jacquot and Fournier, 2000; Gasvik, 2002; Gastinger et al., 2003; Joenathan and Tiziani, 2003). Using a single illumination wavelength, static and dynamic measurements of object displacement provide the same results as holographic nondestructive testing. Just as in holographic interferometry, the speckle interferometry nondestructive testing techniques can measure only a single component of object displacement. This component, in the direction of the sensitivity vector, is described in the section on holographic interferometry for nondestructive testing. Two-wavelength techniques can also be used to perform a reduced-sensitivity test equivalent to two-wavelength holography or interferometry.

An in-line speckle interferometry setup for the measurement of out-of-plane displacement is shown in Figure 16.27. The optical system is similar to those described for holographic test techniques where the path lengths must be adjusted to be within the coherence length of the source, the polarizations of the two beams must be such that they are in the plane of incidence at the viewing plane, and the speckle should be recorded in a plane conjugate to the test surface so that the interference corresponds directly to changes in the object surface. Speckle interferometry techniques also utilize a small aperture which determines the size of the speckle at the recording plane.

Essentially, speckle interferometry techniques perform the holographic interferometry techniques described earlier without the need to make an intermediate

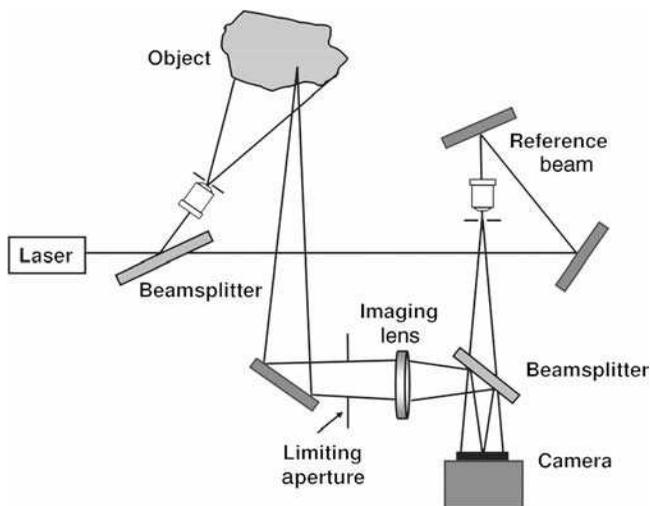


FIGURE 16.27. Speckle interferometry setup similar to that for making an in-line hologram.

hologram. The basic digital speckle interferometry technique was coined electronic speckle-pattern interferometry (ESPI) (Archbold et al., 1970; Butters and Leendertz, 1971; Macovski et al., 1971; Biedermann and Ek, 1975; Ennos, 1975; Jones and Wykes, 1983; Løkberg and Slettemoen, 1987). This technique originally used a TV camera as the recording device and processed interferograms in either electronics or a computer or array processor. It was often sometimes called *TV holography* and now is more often simply called *digital holography*. A slightly different variation on ESPI which provides improved fringe contrast has been coined electro-optic holography (EOH) by Stetson and Brohinsky (Stetson and Brohinsky, 1985, 1986, 1987; Bushman, 1989; Feit, 1989; Stetson, 1989; Stetson et al., 1989). Phase measurements can be applied to ESPI, DH and EOH to provide quantitative displacement maps of the same data obtained in holographic interferometry.

Electronic Speckle-Pattern Interferometry. ESPI uses an optical setup similar to that of Figure 16.27, where a TV camera, CCD camera, or detector array is placed at the image of the test surface (Archbold et al., 1970; Butters and Leendertz, 1971; Macovski et al., 1971; Biedermann and Ek, 1975; Ennos, 1975; Jones and Wykes, 1983; Løkberg and Slettemoen, 1987). Historically all of the processing was performed in electronic hardware and results were displayed in real-time on a TV monitor. Because standard video signals were generated, the results of a test could be stored electronically for later viewing and processing.

For ESPI the limiting aperture size is adjusted so that the speckle will be resolved by the camera. (When the speckles are reduced in size, the effect is to reduce the contrast of the secondary interference fringes.) With static ESPI measurements, a speckle interferogram of the object is recorded by the camera and stored electronically. A single interferogram can be written as

$$I = I_0(1 + \gamma \cos \phi) \quad (16.47)$$

where I_0 is the dc intensity, γ is the visibility, and ϕ is the phase of the interference between the reference beam and the speckle pattern scattered by the object. After the stress applied to the object is changed, Eq. (16.47) becomes

$$I = I_0(1 + \gamma \cos \phi') \quad (16.48)$$

where $\phi' = \phi + \Delta\phi$, and $\Delta\phi$ is the phase change. The stored interferogram is then subtracted from exposures recorded after a change in applied stress to the object, and the difference is squared to yield

$$I^2 = 4I_0^2\gamma^2(\sin^2 \phi)[\sin^2(\Delta\phi/2)] \quad (16.49)$$

This equation shows that there are fringes due to object displacement $\Delta\phi$ as well as fringes due to the phase of speckles ϕ resulting from the interference between the reference and speckled object beams. The $\sin^2\phi$ term causes modulation of the speckle and makes the fringes noticeably noisy. Much of the research effort in the last

two decades has been to reduce this noise due the inherent properties of the speckle (see for example (Federico et al., 2000; Kolenovic et al., 2003)). That which is the signal is also a noise source depending upon your point of view (Creath, 1993).

Time-average measurements can be made with ESPI using a single frame of data. For a vibrating object, the signal at the camera for a single detector point averaged over many cycles of vibration is given by

$$I = I_0[1 + \gamma \cos(\phi)M(\Omega)] \quad (16.50)$$

where ϕ is the phase difference between the object and reference beams, and Ω is proportional to the amplitude of the object displacement. In order to process this signal, the dc component is filtered out, and then the signal is rectified and squared to yield

$$I = I_0\gamma(\cos^2 \phi)|M(\Omega)|^2 \quad (16.51)$$

Note that this expression has a factor of $\cos^2\phi$, which is not included in the time-average holographic vibration fringes given by Eq. (16.39). The $\cos^2\phi$ term is due to the phase of the speckles, and causes the secondary fringes to be noisy. For a sinusoidal object motion, Eq. (16.51) can be rewritten as

$$I = I_0\gamma(\cos^2 \phi)J_0^2(\Omega) \quad (16.52)$$

Because the processing for static and dynamic measurements is performed in electronics, fringe data can be obtained with ESPI at video frame rates (25 or 30 frames per second) or faster depending upon the camera and data acquisition hardware. This speed enables measurements to be made even when the object is not very stable. This same math also can be used to describe time-average digital holography.

Example of fringes obtained using ESPI are shown in Figures 16.28 and 16.29. Fringes for a single frame of time-average data for a vibrating flat plate are shown in Figure 16.28(a). By averaging a number of statistically independent fringe patterns obtained by changing the angle of illumination on the object, the speckle noise in the fringe pattern can be reduced as shown in Figure 16.28(b) (Creath, 1985c). Figure 16.29 shows a time-average interferogram of a car body vibrating at 110 Hz (Malmo and Vikhagen, 1988). These data were taken with the car sitting in a parking lot. The body was covered with retroreflective tape and excited using a loud speaker on the seat. A commercially-available ESPI instrument with a 7 mW helium–neon laser was used for the measurements.

Electrooptic Holography. A major disadvantage of ESPI is that after processing the interference fringe data, there are terms proportional to the phase of the speckle produced by the interference between the reference beam and the speckle in the object beam. These terms cause noisy-looking fringes. By processing the data differently, Stetson and Brohinsky have shown that the speckle term can be removed

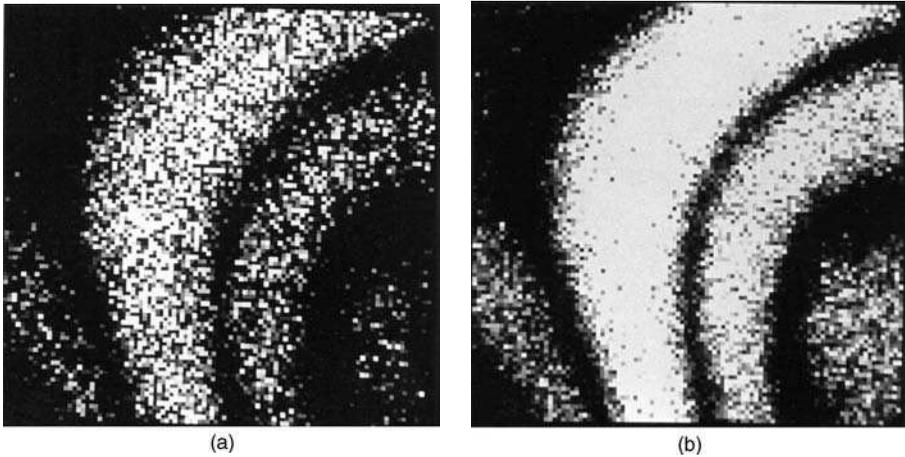


FIGURE 16.28. ESPI interferogram of vibration mode of a flat plate using (a) single interferogram and (b) with speckle averaging. (From Creath, 1985c.)

(Stetson and Brohinsky, 1985, 1986, 1987; Bushman, 1989; Feit, 1989; Stetson, 1989; Stetson et al., 1989). The electro-optic holography (EOH) system uses the same optical setup as ESPI; however, one mirror in the reference beam is mounted on a piezo-electric translator (PZT) to provide relative phase shifts between the object and reference beams. Data can be taken at TV frame rates and stored on videotape. The major difference of EOH from ESPI is the use of a pipeline image processor to handle a large number of images at once instead of the simple electronics of ESPI which process a single frame at a time.

For static measurements, this technique utilizes four frames of data with 90° relative phase shifts between the object and reference which are stored in an array processor. One of these frames of data is written as

$$I_i = I_0[1 + \gamma \cos(\phi + \delta_i)] \quad (16.53)$$

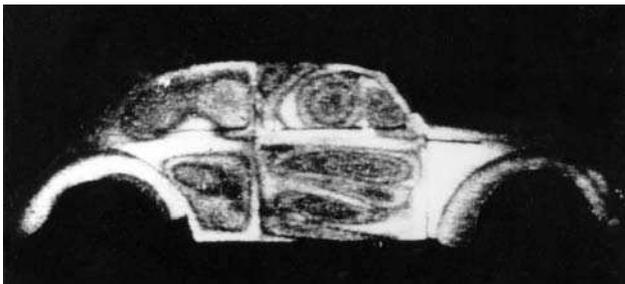


FIGURE 16.29. ESPI interferogram of car body vibrating at 110 Hz taken in parking lot. (From Malmo and Vikhagen, 1988.)

where $\delta_i = 0, \pi/2, \pi,$ and $3\pi/4$ are the applied phase shifts for $I_1, I_2, I_3,$ and I_4 . After the applied stress on the object is changed, four more frames of data are recorded with 90° relative phase shifts:

$$I'_i = I_0[1 + \gamma \cos(\phi' + \delta_i)] \quad (16.54)$$

where $\phi' = \phi + \Delta\phi$. The eight frames of data are combined in a pipeline image processing system to yield

$$\begin{aligned} I &= \sqrt{[(I_1 - I_3) + (I'_1 - I'_3)]^2 + [(I_2 - I_4) + (I'_2 - I'_4)]^2} \\ &= 4I_0\gamma \cos(\Delta\phi/2) \end{aligned} \quad (16.55)$$

This calculation produces fringes proportional to the phase change due to the object displacement. There is no extra term due to the speckles. The calculation involves only addition, subtraction, and the application of a look-up table to find the square root of the sum of the squares of two numbers. This is easily accomplished using an array processor. As the applied stress to the object changes, four new frames need to be obtained. As long as the applied stress is slowly varying, the secondary interference fringes can be calculated after each video frame is recorded and 90° relative phase shifts are applied between consecutive video frames. Thus, these fringes can be calculated and displaced at TV frame rates using an image-processing system.

For a time-average measurement using EOH, four time-average frames of data are taken while the object is vibrating with 90° relative phase shifts. One of these frames is written as

$$I_i = I_0[1 + \gamma \cos(\phi + \delta_i)M(\Omega)] \quad (16.56)$$

where the phase shifts δ_i are $0, \pi/2, \pi,$ and $3\pi/4$. These four frames can then be combined to obtain secondary fringes using

$$I = \sqrt{(I_1 - I_3)^2 + (I_2 - I_4)^2} = 2I_0\gamma|M(\Omega)| \quad (16.57)$$

which can be written as

$$I = \sqrt{(I_1 - I_3)^2 + (I_2 - I_4)^2} = 2I_0\gamma|J_0(\Omega)| \quad (16.58)$$

for a sinusoidal object vibration. Note that the speckle modulation term is not present with this technique.

An example of fringes produced with EOH is shown in Figure 16.30. Figure 16.30(a) shows a rectangular plate undergoing static deflection and Figure 16.30(b) shows a time-average vibration mode of the same plate. EOH can also be used for shearography (shearing speckle interferometry) by utilizing a different optical system (Stetson, 1989).

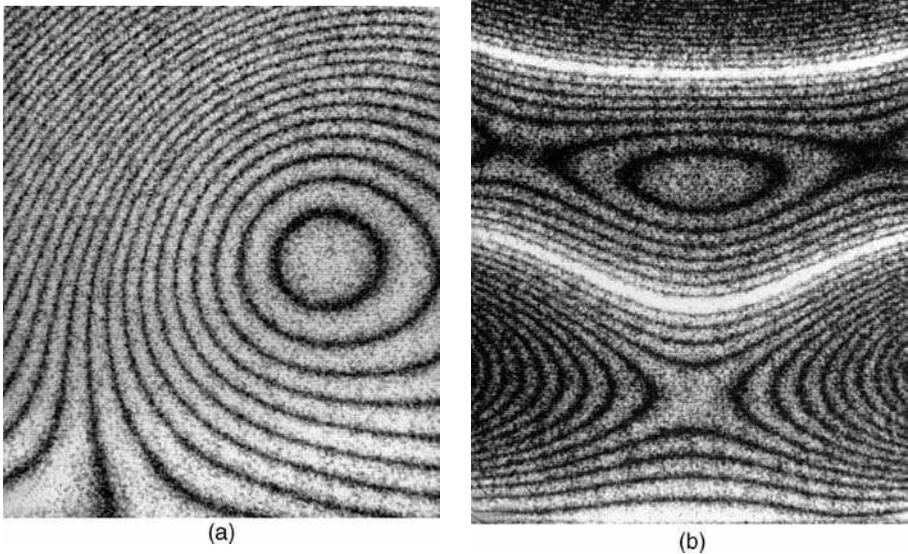


FIGURE 16.30. EOH interferograms of (a) rectangular plate undergoing static deflection, and (b) time-average vibration mode of same plate. (Courtesy K. A. Stetson, United Technologies Research Center.)

Extracting Quantitative Data Using Phase Measurement. A straight-forward technique similar to the two-wavelength interferometry method can be used for quantitative measurement of object displacement or deformation in EPSI (Nakadate and Saito, 1985; Creath, 1985b, 1986; Robinson and Williams, 1986). Using a single measurement wavelength, intensity data are taken for an object, and the phase of the interference pattern is calculated using standard phase-shifting techniques. This is essentially recording the intermediate hologram digitally in the computer by resolving the interference data present on the detector plane. After the stress applied to the object is changed, a second set of intensity data are taken. Modulo 2π phases are calculated for each object state, and the phase due to the change in the object state is found by taking the difference of these two phases

$$\Delta\phi = \phi - \phi' \quad (16.59)$$

where ϕ and ϕ' are the phases before and after the applied force, and $\Delta\phi$ is the phase due change in the object position and shape. $\Delta\phi$ can also be thought of as the phase due to the secondary interference fringes produced in double-exposure holographic interferometry.

A more direct way of determining the phase is to take all the frames of data for both object states and then calculate the phase directly from the intensity data. This

can be written as

$$\Delta\phi = \tan^{-1} \left[\frac{\sin(\Delta\phi)}{\cos(\Delta\phi)} \right] = \tan^{-1} \left[\frac{\sin(\phi - \phi')}{\cos(\phi - \phi')} \right] = \tan^{-1} \left[\frac{\sin\phi \cos\phi' - \cos\phi \sin\phi'}{\cos\phi \cos\phi' + \sin\phi \sin\phi'} \right] \quad (16.60)$$

which for the four-frame method becomes

$$\Delta\phi = \tan^{-1} \left[\frac{(I_4 - I_2)(I'_1 - I'_3) - (I_1 - I_3)(I'_4 - I'_2)}{(I_1 - I_3)(I'_1 - I'_3) + (I_4 - I_2)(I'_4 - I'_2)} \right] \quad (16.61)$$

where I_i and I'_i are the i th data frames taken for each object state. This process will work for any phase measuring algorithm where the phase ϕ_i can be written in terms of a numerator function N_i and a denominator function D_i as

$$\phi_i = \tan^{-1} \left[\frac{\sin\phi_i}{\cos\phi_i} \right] = \tan^{-1} \left[\frac{N_i}{D_i} \right] \quad (16.62)$$

The phase of the difference between measurement 1 and measurement 2 is then given by

$$\Delta\phi = \tan^{-1} \left[\frac{\sin(\phi_1 - \phi_2)}{\cos(\phi_1 - \phi_2)} \right] = \tan^{-1} \left[\frac{N_1 D_2 - D_1 N_2}{D_1 D_2 + N_1 N_2} \right] \quad (16.63)$$

Another type of phase calculation requires data at only two different phase shifts for displacement/deformation measurement. A total of four frames of data, two before and two after are all the data required in a method developed by Kerr et al. (1990).

Another variation utilizes EOH. The eight frames of data taken for a static measurement [Eqs. (16.53) and (16.54)] can be used to calculate the phase change $\Delta\phi$ due to the displacement of the object,

$$\Delta\phi = \tan^{-1} \left(\frac{C_3 - C_4}{C_1 - C_2} \right) \quad (16.64)$$

where the quantities C_i are given by

$$C_1 = [(I_1 - I_3) + (I'_1 - I'_3)]^2 + [(I_2 - I_4) + (I'_2 - I'_4)]^2 \quad (16.65)$$

$$C_2 = [(I_1 - I_3) - (I'_1 - I'_3)]^2 + [(I_2 - I_4) - (I'_2 - I'_4)]^2 \quad (16.66)$$

$$C_3 = [(I_1 - I_3) + (I'_2 - I'_4)]^2 - [(I_2 - I_4) + (I'_1 - I'_3)]^2 \quad (16.67)$$

$$C_4 = [(I_1 - I_3) - (I'_2 - I'_4)]^2 - [(I_2 - I_4) - (I'_1 - I'_3)]^2 \quad (16.68)$$

These calculations only involve simple calculations and the use of lookup tables which can be done in an array processor. The modulo 2π phase corresponding to the object displacement can be calculated in real time at video frame rates. Quantitative measurements of time-average interferograms can be made using the techniques described for holographic nondestructive testing in section 16.2.2 and Eqs. (16.42) thru (16.46) (Pryputniewicz and Stetson, 1989).

In practice there are numerous other ways that have been developed to extract phase information for speckle interferometry and holographic interferometry. Most of the algorithms are tailored to a specific application and a specific type of measurement and optical system.

Dynamic Techniques. In the real world of testing diffuse objects using speckle interferometry and digital holography, it can be difficult to ensure that an object is stable and not moving relative to the instrument. It is desirable to have instrumentation that is either insensitive to vibration or can freeze motion to make measurements. For this, many different techniques have been developed to cope with these situations. This essentially enables taking data of arbitrarily moving objects and being able to track motion, modes of vibration, flows, and even watch shock waves travel. For dynamic measurements, there are many different commercial systems available that fit in a briefcase size or are mounted on a tripod. Most of the work published on dynamic techniques in the past decade has originated in Europe and been initiated to find solutions for automotive applications.

To freeze motion, some sort of pulsed laser is usually implemented. Pulsed lasers can provide enough energy in a single pulse to get a decent exposure on the return with a standard silicon detector array. Fast shutters can also be used, but then more overall power is needed from the source to get a sufficient exposure unless low-light level imaging systems are implemented. Any of the techniques mentioned so far in this chapter can be implemented using a pulsed laser or fast shutter including all of the quantitative techniques that determine phase. It is also possible to unwrap phase in time as well as space to follow motion in time (Ruiz et al., 2003). Amplitude and phase information can both be extracted and added to shape information to determine complex motions and perform active metrology measurements (Osten, 2003). Modifications of the optical system utilizing endoscopes and fibers can enable getting measurements in tight spaces of otherwise hidden surfaces (Pedrini et al., 2003) or biological objects (Pedrini et al., 2005). These types of developments drastically increase the possible applications and usefulness of digital holography and speckle interferometry.

Shearography as described in the next subsection is one way to ensure that vibrations and relative motions between the interferometer and object don't hamper the data. Another way to get quantitative data from a single fast snapshot is to use phase measuring techniques that either require only a single interferogram such as Fourier Transform techniques (Kreis, 1986, 1987) or those that multiplex phase data using polarization techniques to take all necessary data at once (Koliopoulos, 1992) (North-Morris et al., 2005).

Shearography. Just as there are many ways to test optical components with shearing interferometry, so there are many ways to test diffuse surfaces with speckle shearing interferometry also known as *shearography* (Joenathan and Tiziani, 2003; Steinchen and Yang, 2003). Shearography tests a surface relative to itself. Because it is common path, it is insensitive to vibrations and relative motion between the part under test and the interferometer. But, like all shearing interferometers, it does not directly measure displacement or shape. Instead it measures the slope of the displacement or shape in the direction of the shear. It can be used with phase measurement techniques as described earlier in this chapter to quantitatively extract the slope. To get to displacement or shape from the slope data, the techniques are similar to those described in the shearing interferometer chapter of this book. Steinchen and Yang outline numerous techniques in their book *Digital Shearography* (Steinchen and Yang, 2003).

Low Coherence Techniques. Low coherence techniques combine a longer depth-resolved measurement range with speckle interferometry and digital holography techniques (Gastinger et al., 2003; Martinez-Leon et al., 2004; Gastinger and Winther, 2005). Like low coherence interferometry (also known as white light interferometry, or vertical scanning interferometry) low coherence speckle techniques follow the high coherence fringe and map out where the greatest fringe contrast is. These techniques require scanning in depth in order to make up a composite phase map and are similar to the interference microscope techniques described in Chapter 15 of this book.

REFERENCES

- Abramson N., "Sandwich Hologram Interferometry. 3: Contouring," *Appl. Opt.*, **15**, 200–205 (1976a).
- Abramson N., "Holographic Contouring by Translation," *Appl. Opt.*, **15**, 1018–1022 (1976b).
- Amridror I., *The Theory of the Moiré Phenomenon*. Kluwer Academic Publishers, Dordrecht, 2000.
- Archbold E., J. M. Burch, and A. E. Ennos, "Recording of In-Plane Surface Displacement by Double-Exposure Speckle Photography," *Opt. Acta*, **17**, 883–898 (1970).
- Armitage D., I. Underwood and S.-T. Wu, Microdisplays Move Outside the Box, *OE magazine*, 2002, November pp. 28–30.
- Bell B., *Digital Heterodyne Topography*, Ph.D. Dissertation, Optical Sciences Center, University of Arizona, Tucson, AZ, 1985.
- Benoit P., E. Mathieu, J. Hormier, and A. Thomas, "Characterization and Control of Three Dimensional Objects Using Fringe Projection Techniques," *Nouv. Rev. Opt.*, **6**, 67–86 (1975).
- Biedermann K. and L. Ek, "A Recording and Display System for Hologram Interferometry with Low Resolution Imaging Devices," *J. Phys. E. Sci. Instrum.*, **8**, 571–576 (1975).
- Brooks R. E. and L. O. Helfinger, "Moiré Gauging Using Optical Interference Patterns," *Appl. Opt.*, **8**, 935–939 (1969).

- Bushman T., "Development of a Holographic Computing System," *Proc. SPIE*, **1162**, 66–77 (1989).
- Butters, J. N. and J. A. Leendertz, "Speckle Pattern and Holographic Techniques in Engineering Metrology," *Opt. Laser Tech.*, **3**, 26–30 (1971).
- Case S. K., J. A. Jalkio, and R. C. Kim, "3-D Vision System Analysis and Design," in *Three-Dimensional Machine Vision*, Takeo Kanade, Ed., Kluwer Academic Publishers, Norwell, MA, 1987, pp. 63–95.
- Chiang, F.-P., "Moiré Methods of Strain Analysis," in *Manual on Experimental Stress Analysis*, A. S. Kobayashi, Ed., Soc. for Exp. Stress Anal., Brookfield Center, CT, 1983, pp. 51–69.
- Creath K., "Averaging Double-Exposure Speckle Interferograms," *Opt. Lett.*, **10**, 582–584 (1985c).
- Creath K., "Phase-Shifting Speckle Interferometry," *Appl. Opt.*, **24**, 3053–3058 (1985b).
- Creath K., *Digital Speckle Pattern Interferometry*, Ph.D. Dissertation, Optical Sciences Center, University of Arizona, Tucson AZ, University Microfilms, Ann Arbor MI, 1985a.
- Creath K., "Speckle: Signal or Noise?" *Proceedings FRINGE 93 Bremen 1993*, W. Jueptner and W. Osten (Eds.), Akademie Verlag, Berlin, pp. 97–102.
- Creath, K., "Measuring Step Heights Using an Optical Profiler," *Proc. SPIE*, **661**, 296–301 (1986).
- Creath K. and G. E. Schwartz, "The Dynamics of Life: Imaging Temperature and Refractive Index Variations Surrounding Material and Biological Systems with Dynamic Interferometry", *Fringe 2005: The 5th International Workshop on Automatic Processing of Fringe Patterns*. W. Osten (Eds.), Springer-Verlag, Berlin, 2005, pp. 588–596.
- Dandliker R., "The Story of Speckles in Interferometry," *Interferometry in Speckle Light*, P. Jacquot and J.-M. Fournier (Eds.), Springer Verlag, Berlin, 2000, pp. 3–10.
- D'Apuzzo N., "Overview of 3D Surface Digitization Technologies in Europe," *Proc. SPIE* 6056, 605605:1–13, (2006).
- DeMattia P. and V. Fossati-Bellani, "Holographic Contouring by Displacing the Object and the Illumination Beam," *Opt. Commun.*, **26**, 17–21 (1978).
- Der Hovanesian J. and Y. Y. Yung, "Moiré Contour-Sum Contour-Difference, and Vibration Analysis of Arbitrary Objects," *Appl. Opt.*, **10**, 2734–2738 (1971).
- Durelli A. J. and V. J. Parks, *Moiré Analysis of Strain*, Prentice-Hall, Englewood Cliffs, NJ, 1970.
- Ennos A., "Speckle Interferometry," in *Laser Speckle and Related Phenomena*, J. C. Dainty, Ed., Springer-Verlag, New York, 1975, pp. 203–253.
- Erf R. K., *Speckle Metrology*, Academic Press, New York, 1978.
- Federico A., G. H. Kaufmann, E. P. Serrano, "Speckle Noise Reduction in ESPI Fringes Using Wavelet Shrinkage," *Interferometry in Speckle Light*, P. Jacquot and J.-M. Fournier (Eds.), Springer Verlag, Berlin, 2000, pp. 397–404.
- Feit E., "Electronic Holography for Non-Destructive Testing," *Advanced Imaging*, January, 42–45, (1989).
- Frankowski G., M. Chen, T. Huth, "Real-Time Shape Measurement With Digital Stripe Projection by Texas Instruments Micromirror Devices DMDTM," *Proc. SPIE* **3958**, 90–105 (2000).
- Fujii H. and J. W. Lit, "Surface Roughness Measurement Dichromatic Speckle Patterns: An Experimental Study," *Appl. Opt.*, **17**, 2690–2694 (1978).

- Gastinger K. and S. Winther, "Optimisation of Low Coherence Speckle Interferometry (LSCI) for Characterisation of Multi-layered Materials," *Proc. SPIE* **5858**, 58580K:1-12 (2005).
- Gastinger K., O. J. Lokberg, and S. Winther, *Speckle Metrology 2003*, SPIE Proc, Bellingham, WA, SPIE, 2003.
- Gastinger K., S. Winther, and K. Hinsch, "Low-Coherence Speckle Interferometry (LCSI) for Characterisation of Adhesion in Adhesive Bonded Joints," *Proc. SPIE*, **4933**, 59–65 (2003).
- Gasvik K. J., *Optical Metrology*, 3rd Ed., John Wiley and Sons, Chichester, England, (2002).
- Goodman J. W., "Dependence of Image Speckle Contrast on Surface Roughness," *Opt. Commun.*, **14**, 324–327 (1957b).
- Goodman J. W., "Statistical Properties of Laser Speckle Patterns," in *Laser Speckle and Related Phenomena*, J. C. Dainty, Ed., Springer Verlag, Berlin, 1975a, pp. 9–75.
- Guild J., *The Interference Systems of Crossed Diffraction Gratings*, Clarendon Press, Oxford, 1956.
- Guild J., *Diffraction Gratings as Measuring Scales*, Oxford University Press, London, 1960.
- Haines K. and B. P. Hildebrand, "Contour Generation by Wavefront Reconstruction," *Phys. Lett.*, **19**, 10–11 (1965).
- Halioua M. and H.-C. Liu, "Optical Three-Dimensional Sensing by Phase Measuring Profilometry," *Opt. Lasers Eng.*, **11**, 185–215 (1989).
- Harding K., "Latest Optical Methods for Industrial Dimensional Metrology," *Proc. SPIE* 6000, 600001:1–14 (2005).
- Hariharan P., "Quasi-Heterodyne Hologram Interferometry," *Opt. Eng.*, **24**, 632–638 (1985).
- Hariharan P., *Optical Holography*, Cambridge University Press, Cambridge, 1984.
- Hariharan P., B. F. Oreb, and C. H. Freund, "Stroboscopic Holographic Interferometry: Measurements of Vector Components of a Vibration," *Appl. Opt.*, **26**, 3899–3903 (1987).
- Hildebrand B. P. and K. A. Haines, "Interferometric Measurements Using the Wavefront Reconstruction Techniques," *Appl. Opt.*, **57**, 155–162 (1967).
- Huang P. S. and S. Zhang, "3D Optical Measurement Using Phase Shifting Based Methods," *Proc. SPIE* 6000, 600002:1–10 (2005).
- Hung Y. Y., L. Lin, and B. G. Park, "Practical Three-Dimensional Computer Vision Techniques for Full-Field Surface Measurement," *Opt. Eng.*, **39**, 143–149 (2000).
- Huntley J. M. "Speckle Photography Fringe Analysis: Assessment of Current Algorithms," *Appl. Opt.*, **28**, 4316–4322 (1989).
- Indebetouw G., "Profile Measurement Using Projection of Running Fringes," *Appl. Opt.*, **17**, 2930–2933 (1978).
- Jacquot P. and J.-M. Fournier, *Interferometry in Speckle Light: Theory and Applications*, Springer Verlag, Berlin, 2000.
- Joenathan C. and H. J. Tiziani, "Speckle and Speckle Metrology," *The Optics Encyclopedia*, T. G. Brown, K. Creath, H. Kogelnik (Eds.), Wiley-VCH, Berlin, 2003, pp. 2709–2773.
- Jones R. and C. Wykes, *Holographic and Speckle Interferometry*, Cambridge University Press, Cambridge, 1983.
- Kakunai S., K. Iwata, R. Nagata, and H. Sekiguchi, "Measurement of Three Components of a Displacement Vector using Heterodyne Holographic Interferometry," *Opt. Lasers Eng.*, **6**, 213–223 (1985).

- Kerr D., F. Mendoza-Santoyo, and J. R. Tyrer, "Extraction of Phase Data from Electronic Speckle Pattern Interferometric Fringes Using a Single-Step Method: A Novel Approach," *J. Opt. Soc. Am. A*, **7**, 820–826 (1990).
- Kolenovic E., W. Osten, and W. P. Juptner, "Improvement of Interferometric Phase Measurements by Consideration of the Speckle Field Topology," *Proc. SPIE*, **4933**, 206–211 (2003).
- Koliopoulos C. L., "Simultaneous Phase-shift Interferometer," *Advanced Optical Manufacturing and Testing II*, vol. 1531, V. J. Doherty (Ed.), SPIE, Bellingham, WA, 1992, pp. 119–127.
- Koninckx T. P. and L. Van Gool, "Real-Time Acquisition by Adaptive Structured Light," *IEEE Trans Pattern Analysis and Machine Intelligence*, **28**, 432–445 (2006).
- Kowarschik R., P. Kuhmstedt, J. Gerber, W. Schreiber, and G. Notni, "Adaptive Optical Three-Dimensional Measurement with Structural Light," *Opt. Eng.*, **39**, 150–158 (2006).
- Kreis T. M., "Digital Holographic Interference-Phase Measurement Using the Fourier Transform Method," *J. Opt. Soc. Am. A*, **3**, 847–855 (1986).
- Kreis T. M., "Fourier-Transform Evaluation of Holographic Interference Patterns," *Proc. SPIE*, **814**, 365–371 (1987).
- Kreis T., *Handbook of Holographic Interferometry*. Wiley-VCH, Weinheim, Germany, 2005.
- Kujawińska M. and D. Malacara, "Optical Metrology of Diffuse Objects: Full-Field Methods" in: *Handbook of Optical Engineering*, D. Malacara (Ed.), Marcel Dekker Inc, New York, 2001, pp. 493–522.
- Ligtenberg F. K., "The Moiré Method," *Proc. Soc. Exp. Stress Anal. (SESA)*, **12**, 83–98 (1955).
- Løkberg O. J. and G. A. Slettemoen, "Basic Electronic Speckle-pattern Interferometry," in *Applied Optics and Optical Engineering*, Vol. X, R. R. Shannon and J. C. Wyant, Eds., Academic Press, San Diego, 1987, pp. 455–504.
- MacGovern A. J., "Projected Fringes and Holography," *Appl. Opt.*, **11**, 2972–2974 (1972).
- Macovski A., S. D. Ramsey, and L. F. Schaefer, "Time-Lapse Interferometry and Contouring Using Television Systems," *Appl. Opt.*, **10**, 2722–2727 (1971).
- Malmo J. T. and E. Vikhagen, "Vibration Analysis of a Car Body by Means of TV Holography," *Exp. Tech.*, **12**, 28–30 (1988).
- Martinez-Leon L., G. Pedrini, and W. Osten, "Short-Coherence Digital Holography for the Investigation of 3D Microscopic Samples," *Proc. SPIE*, **5457**, 528–537 (2004).
- Meadows D. M., W. O. Johnson and J. B. Allen, "Generation of Surface Contours by Moiré Patterns," *Appl. Opt.*, **9**, 942–947 (1970).
- Menzel E., "Comment to the Methods of Contour Holography," *Optik*, **40**, 557–559 (1974).
- Mix P. E., *Introduction to Nondestructive Testing: A Training Guide, Second Edition*, John Wiley and Sons, Hoboken, NJ, 2005.
- Moore D. T. and B. E. Truax, "Phase-Locked Moiré Fringe Analysis for Automated Contouring of Diffuse Surfaces," *Appl. Opt.*, **18**, 91–96 (1979).
- Nakadate S. and H. Saito, "Fringe Scanning Speckle-Pattern Interferometry," *Appl. Opt.*, **24**, 2172–2180 (1985).
- Nakadate S., N. Magome, T. Honda, and J. Tsujiuchi, "Hybrid Holographic Interferometer for Measuring Three-Dimensional Deformations," *Opt. Eng.*, **20**, 246–252 (1981).
- Neumann D. B., C. F. Jacobson, and G. M. Brown, "Holographic Technique for Determining the Phase of Vibrating Objects," *Appl. Opt.*, **9**, 1357–1362 (1970).
- North-Morris M., J. Millerd, N. Brock, J. Hayes, and B. Saif, "Dynamic Phase-Shifting Electronic Speckle Pattern Interferometer," *Proc SPIE* **5869**, 58691B:1–9 (2005).

- Notni G., S. Riehemann, P. Kuehmstedt, L. Heidler, and N. Wolf, "OLED Microdisplays – A New Key Element for Fringe Projection Setups," *Proc. SPIE* **5532**, 170–177 (2004).
- Oshida Y., K. Itawa, and R. Nagata, "Optical Heterodyne Measurement of Vibration Phase," *Opt. Lasers Eng.*, **4**, 67–69 (1983).
- Osten W., "Active Metrology by Digital Holography," *Proc. SPIE* **4933**, 96–110 (2003).
- Oster G. and Y. Nishijima, "Moiré Patterns," *Sci. Am.*, **208**, May 54–63 (1963).
- Patorski K. and M. Kujawinska, *Handbook of the Moiré Fringe Technique*, Elsevier Science Publishers, Amsterdam, 1993.
- Pedriani G., I. Alexeenko, and H. J. Tiziani and W. Oskey, "Digital Holographic Interferometry for Investigations in Biomechanics," *Proc. SPIE* **5776**, 325–332 (2005).
- Pedriani G., I. Alexeenko, and H. J. Tiziani, "Pulsed Endoscopic Digital Holographic Interferometry for Investigation of Hidden Surfaces," *Proc. SPIE* **4933**, 123–128 (2003).
- Perrin J. C. and A. Thomas, "Electronic Processing of Moiré Fringes: Application to Moiré Topography and Comparison with Photogrammetry," *Appl. Opt.*, **18**, 563–574 (1979).
- Post D., "Developments in Moiré Interferometry," *Opt. Eng.*, **21**, 458–467 (1982).
- Post D., B. Han and P. Ifju, *High Sensitivity Moiré: Experimental Analysis for Mechanics and Materials*, Springer, 1997.
- Powell R. L. and K. A. Stetson, "Interferometric Vibration Analysis by Wavefront Reconstruction," *J. Opt. Soc. Am.*, **55**, 1593–1598 (1965).
- Proll K.-P., J.-M. Nivet, V. Voland and H. J. Tiziani, "Application of Liquid Crystal Spatial Light Modulator for Brightness Adaptation in Microscopic Topometry," *Applied Optics* **39**, 6430–6435 (2003).
- Pryputniewicz R. J. and K. A. Stetson, "Holographic Strain Analysis: Extension of Fringe-Vector Method to Include Perspective," *Appl. Opt.*, **15**, 725–728 (1976).
- Pryputniewicz R. J. and K. A. Stetson, "Measurement of Vibration Patterns Using Electro-Optic Holography," *Proc. SPIE*, **1162**, 456–467 (1989).
- Pryputniewicz R. J., "Speckle Metrology Techniques and their Applications," *Proc. SPIE*, **556**, 90–98 (1985).
- Reich C., R. Ritter and J. Thesing, "3D Shape Measurement of Complex Objects by Combining Photogrammetry and Fringe Projection," *Optical Eng.* **39**, 224–231 (2000).
- Reid G. T., "Moiré Fringes in Metrology," *Opt. Lasers Eng.*, **5**, 63–93 (1984a).
- Reid G. T., R. C. Rixon, and H. I. Messer, "Absolute and Comparative Measurements of Three-Dimensional Shape by Pulse Measuring Moiré Topography," *Opt. Laser Tech.*, **16**, 315–319 (1984b).
- Righi, A., "Sui Fenomeni Che si Producono colla Sovrapposizione dei Due Reticoli e sopra Alcune Lora Applicazioni: I," *Nuovo Cim.*, **21**, 203–227 (1887).
- Ritter R. and R. Hahn, "Contribution to Analysis of the Reflection Grating Method," *Optics Lasers Eng.* **4**, 13–24 (1983).
- Robinson D. W. and D. C. Williams, "Digital Phase Stepping Speckle Interferometry," *Opt. Commun.*, **57**, 26–30 (1986).
- Rowe S. H. "Projected Interference Fringes in Holographic Interferometry," *J. Opt. Soc. Am.*, **61**, 1599–1603 (1971).
- Rowe S. H. and W. T. Welford, "Surface Topography of Non-Optical Surfaces by Projected Interference Fringes," *Nature*, **216** (5117), 786–787 (1967).

- Ruiz P. D., J. M. Huntley, and G. Kaufmann, "Recent Advances in Temporal Phase-Shifting Speckle Interferometry," *Proc. SPIE*, **4933**, 297–304 (2003).
- Salvi J., J. Pages and J. Battle, "Pattern Codification Strategies in Structured Light Systems," *Pattern Recognition* **37**, 827–849 (2004).
- Schnars U. and W. Jueptner, *Digital Holography: Digital Hologram Recording, Numerical Reconstruction, and Related Techniques*, Springer Verlag, Berlin, 2005.
- Shagam R., "Heterodyne Interferometric Method for Profiling Recorded Moiré Interferograms," *Opt. Eng.*, **19**, 806–809 (1980).
- Srinivasan Y., H. C. Liu, and M. Halioua, "Automated Phase-Measuring Profilometry of 3-D Diffuse Objects," *Appl. Opt.*, **23**, 3015–3108 (1984).
- Srinivasan Y., H. C. Liu, and M. Halioua, "Automated Phase-Measuring Profilometry: A Phase Mapping Approach," *Appl. Opt.*, **24**, 185–188 (1985).
- Steinchen W. and L. Yang, *Digital Shearography: Theory and Application of Digital Speckle Pattern Shearing Interferometry*, SPIE, Bellingham, WA, 2003.
- Stetson K. A. and R. L. Powell, "Interferometric Hologram Evaluation and Real-Time Vibration Analysis of Diffuse Objects," *J. Opt. Soc. Am.*, **55**, 1694–1695 (1965).
- Stetson K. A. and W. R. Brohinsky, "Electro-Optic Holography and Its Application to Hologram Interferometry," *Appl. Opt.*, **24**, 3631–3637 (1985).
- Stetson K. A. and W. R. Brohinsky, "Measurement of Phase Change in Heterodyne Interferometry: A Novel Scheme," *Appl. Opt.*, **25**, 2643–2644 (1986).
- Stetson K. A. and W. R. Brohinsky, "Electro-Optic Holography System for Vibration Analysis and Nondestructive Testing," *Opt. Eng.*, **26**, 1234–1239 (1987).
- Stetson K. A. and W. R. Brohinsky, "Fringe-Shifting Technique for Numerical Analysis of Time-Average Holograms of Vibrating Objects," *J. Opt. Soc. Am. A*, **5**, 1472–1476 (1988).
- Stetson K. A. and W. R. Brohinsky, J. Wahid and T. Bushman, "An Electro-Optical Holography System with Real-Time Arithmetic Processing," *J. Nondestructive Evaluation*, **24**, 3631–3637 (1989).
- Stetson K. A., "A Review of Speckle Photography and Interferometry," *Opt. Eng.*, **14**, 482–489 (1975).
- Stetson K. A., "Electro-Optic Holography for Real-Time Display and Quantitative Analysis of Interference Fringes," Fringe'89, Automatic Processing of Fringe Patterns (E. Berlin, April 25–28, 1989) and Optical Sensing and Measurement, *Proc. ICALEO, Laser Int. Am.*, **70**, 78–85 (1989).
- Stetson K. A., "Use of Projection Matrices in Hologram Interferometry," *J. Opt. Soc. Am.*, **69**, 1705–1710 (1979).
- Stetson K. A., "Use of Sensitivity Vector Variations to Determine Absolute Displacements in Double Exposure Hologram Interferometry," *Appl. Opt.*, **29**, 502–504 (1990).
- Takasaki H., "Moiré Topography," *Appl. Opt.*, **12**, 845–850 (1973).
- Takasaki H., "Moiré Topography," *Appl. Opt.*, **9**, 1467–1472 (1970).
- Takeda M. and K. Mutoh, "Fourier Transform Profilometry for the Automatic Measurement of 3-D Object Shapes," *Appl. Opt.*, **22**, 3977–3982 (1983).
- Takeda M., H. Ina, and S. Kabayashi, "Fourier-Transform Method of Fringe-Pattern Analysis for Computer-Based Topography and Interferometry," *J. Opt. Soc. Am.*, **72**, 156–160 (1982).
- Theocarlis P. S., "Moiré Fringes: A Powerful Measuring Device," *Appl. Mech. Rev.*, **15**, 333–339 (1962).
- Theocarlis P. S., *Moiré Fringes in Strain Analysis*, Pergamon Press, Oxford, 1969.

- Van Deelen W. and P. Nisenson, "Mirror Blank Testing by Real-Time Holographic Interferometry," *Appl. Opt.*, **8**, 951–955 (1969).
- Varner J. R., "Holographic Moiré Surface Contouring," in *Holographic Nondestructive Testing*, R. K. Erf, Ed., Academic Press, Orlando, 1974.
- Vest C. M., *Holographic Interferometry*, Wiley, New York, 1979.
- Walker C. A. (Ed.), *The Practical Application of the Moire Fringe Method in the Measurement of Strain and Shape*, IOP Publishing, Bristol, UK, 2004.
- Wasowski J., "Moiré Topographic Maps," *Opt. Commun.*, **2**, 321–323 (1970).
- Welford, W. T., "Some Applications of Projected Interference Fringes," *Opt. Acta*, **16**, 371–376 (1969).
- Weller R. and B. M. Shepperd, "Displacement Measurement by Mechanical Interferometry," *Proc. Soc. Exp. Anal., (SESA)*, **6**, 35–38 (1984).
- Winther S. and G. A. Slettemoen, "An EPSI Contouring Technique in Strain Analysis," *Proc. SPIE*, **473**, 44–47 (1983).
- Wyant J. C., 2006, Moiré Patterns. <http://www.optics.arizona.edu/jcwyant/MoirePatterns/MoirePatterns.htm>.
- Yaroslavsky L., *Digital Holography and Digital Image Processing: Principles, Methods, Algorithms*, Kluwer Academic Publishers, Boston, 2004.
- Yatagai T., M. Idesawa, Y. Yamaashi, and M. Suzuki, "Interactive Fringe Analysis System: Applications to Moiré Contourgram and Interferogram," *Opt. Eng.*, **21**, 901–906 (1982).

17

Angle, Prisms, Curvature, and Focal Length Measurements

Z. Malacara

17.1. INTRODUCTION

Geometrical measurements of a lens in the optical workshop are necessary to obtain precise lens and optical system parameters. On doing so, one is not only looking for precision but also for speed and simplicity. These measurements should be made following international units and in some cases, must be certified by a laboratory from a secondary standard. A review of this topic may be found in the article by Geiser (1965). Geometrical measurements in optical testing may be classified mainly in angle, distance, curvature, and focal length measurements.

In the production run, it is preferred not to make measurements but comparisons or null measurements. This calls for gauges and templates that are more appropriate for the shop rather than expensive and precise equipment.

17.2. ANGLE MEASUREMENTS

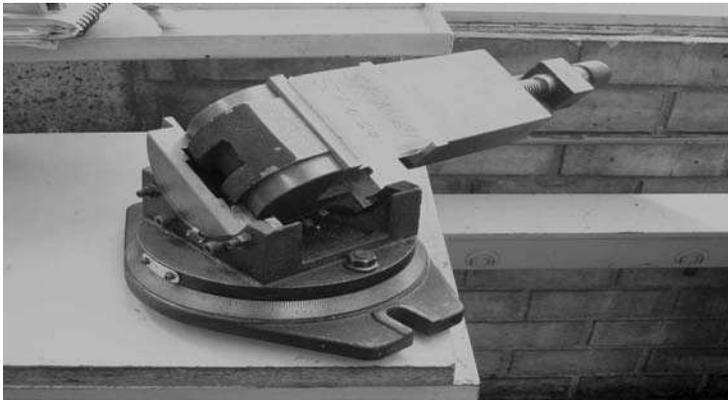
Angle measurements in the optical shop require different levels of accuracy. While for glass cutting the accuracy can be several degrees, for standard test plates, less than a second of arc may be required. For every case, different measurement methods are developed. For angle measurements, the most common unit remains the degree, which divides the circle in 360 parts. A radian is an angle such that the arc length is equal to the radius of the circle. A gradient which is scarcely used is an attempt to use decimal fractions of a quadrant.

17.2.1. Divided Circles and Goniometers

Protractors are used as main angle measuring devices. Although a typical protractor has an accuracy of about 30 min, (Fig. 17.1(a)) modern electronic protractors can



(a)



(b)

FIGURE 17.1. (a) A shop protractor; (b) an angle vise.

attain up to 0.5 min of precision. During the manufacturing process, several methods are used to cut a block of glass at a given angle. An angle vise (Fig. 17.1(b)) and a sine plate are mechanisms to support the glass and to measure the angle precisely. For these supporting heads, the table is tilted or raised at some known distance for a known hypotenuse length. For glass cutting machine tools, a rotary table is frequently used. An accuracy of about 30 min is attainable for these instruments. Horne (1972) describes a serrated table for angle measurements up to 0.1 seconds. For a semi-finished glass wedge, the angle can be compared to an angle block set. Angle blocks are templates commercially available to a degree and accuracy of ± 20 s, to virtually obtain any angle between 1° and 90° . By reversing the blocks, negative angles are obtained as described by Horne (1974). Angle blocks are made from hardened steel with a flat precision-polished face. This permits angle blocks in conjunction with a goniometer to be used as an angle standard. Angle blocks are sometimes certified up to ± 0.1 s of arc. Similar to angle blocks, polygons are available from 3 up to 12 faces, although Horne (1974) reports a 72-faces polygon. Polygons are mostly used with an



FIGURE 17.2. A goniometer.

autocollimator for divided circles and goniometer calibration. Modern numerically controlled machines can cut a piece of glass to a prism or a polygon within a precision of a few minutes.

A goniometer (Fig. 17.2) is a precision spectrometer table with a fixed collimator and a moving telescope. Microscopes or magnifiers are located at opposite ends in the divided circle; the mean value of the readings is calculated to compensate for centering errors. Goniometers' accuracy can typically yield up to 20 s arc, although Geiser (1965) reviews a system for checking such a circle to better than one arc second. For some precise angle measurements, a theodolite could be used, reaching accuracy up to a second arc. Goniometers are used to check prisms angles; in this case the telescope with illuminated reticle is used as an autocollimator. The glass surface must be semipolished and wet; for a precision work it must be fully polished and flat.

17.2.2. Autocollimator

An autocollimator, is essentially a telescope focused to infinity. An illuminated reticle located at the focal plane is sent to infinity by the objective. A reflecting surface perpendicular to the beam, images the reticle on itself. When the reflecting surface fails to be perpendicular to the beam, the reticle image is displaced at the focal plane. The amount of displacement d is

$$d = 2\alpha f \quad (17.1)$$

with α being the tilt angle for the mirror and f the focal length of the objective (Fig. 17.3).

Objective lenses are usually a corrected doublet, although sometimes a negative lens is included to form a telephoto lens to increase effective focal length, maintaining compactness at the same time. The collimating lens adjustment is

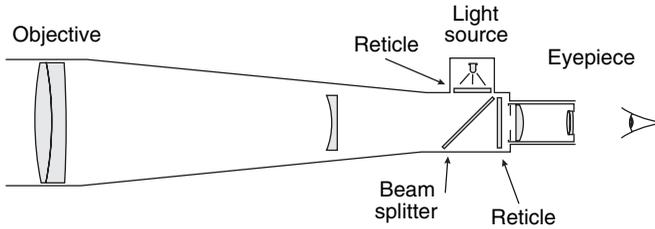


FIGURE 17.3. An autocollimator.

critical for the final accuracy. A means to adjust the collimator has been widely described (De Vany, 1976), with the aid of a lateral shear interferometry (Murty, 1964 and Choi et al., 1995). Talbot interferometry can also be used for a precise adjustment (Kothiyal and Sirohi, 1987) (Sriram et al., 1993).

For observation at the focal plane, several means have been developed. An illuminated reticle at the eyepiece, with a calibrated reticle is used to measure the displacement. Rank (1946) modified a Gauss eyepiece to give dark field. In others systems, a drum micrometer displaces a reticle to be positioned at the image plane. Some autocollimators have a microscope to observe the returning image; such system is called a microptic autocollimator.

Autocollimators are used for angle measurements in prisms and polygons, but are also used in other applications such as measurement of parallelism in optical flats, manufacturing of divided circles (Horne, 1974) and, by integrating values for a scan in position, flatness for machine tool and optical beds can also be measured (Young, 1967).

Some variants to the basic autocollimator include dual-axis reticle or micrometer for measurements for both axes. A moving slit can be included in the focal plane and an electronic synchronous detection system serves for the purpose of indicating a null position for the return beam. Also, a position-measuring detector at the focal plane and a display indicates the angular deviation. Particularities about electronic autocollimators are described by Thurston (1986). Some new computerized versions include software for on-line data reducing.

Direct-reading autocollimators can measure over a field of about one degree, and this field is reduced with the distance to be measured. Precision in an autocollimator is limited by the method for measuring the image's centroid. In a diffraction-limited visual system, diffraction image sets the limit for the precision. Usually the effective f-number is small and the image odd-shaped, like in a prism. For a precision electronic measuring system centroid measurement is limited by the electronic detector, independent from the diffraction image itself, and can go beyond the diffraction limit. In some photo-electric system, precision is improved up to more than 10 times.

Reflecting surface in autocollimator measurements must be kept close to the objective, otherwise the reflected beam fails to enter the system, with a subsequent decrease in the range. The reflecting surface must be a high-quality one. A curved surface has the effect of introducing another lens in the system with a change in the effective focal length (Young, 1967). Irregular surfaces affect the shape of the reflecting beam too.

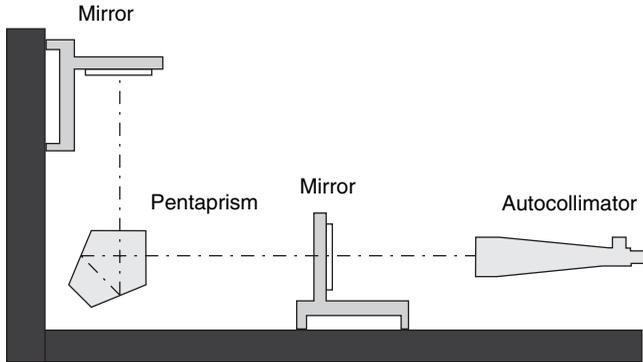


FIGURE 17.4. Perpendicularity measurement with an autocollimator.

Several accessories are part of an autocollimator. For a single-axis angle measurement, a pentaprism is used; an optical square permits angle measurements for surfaces at right angles. Perpendicularity is measured with a pentaprism and a mirror (Fig. 17.4). A handy horizontal reference can be produced with an oil pool; care must be taken with the surface stability. A complete description for the autocollimator is found in Hume (1965).

17.2.3. Interferometric Measurements of Angles

Interferometric methods find their main applications in measuring very small wedge angles in glass slabs (Met, 1966, Leppelmeier and Mullenhoff, 1970) and in parallelism measurements (Wasilik et al., 1971). Refer to Chapter 1 for further details.

Interferometric measurements of large angles may also be performed. In one method a collimated laser beam is reflected from the surfaces from a rotating glass slab; the produced fringes can be considered as coming from a Murty lateral shear interferometer (Malacara and Harris, 1970). This device can be used as a secondary standard to produce angles from 0° to 360° with accuracy within a second of arc. Further analysis of this method is done by Tentori and Celaya (1986). In another system a Michelson interferometer is used with an electronic counter to measure within a range of $\pm 5^\circ$ with a resolution of 10^{-4} degrees (Stijns, 1986) and (Shi and Stijns, 1988). An interferometric optical sine bar for angles in the millisecond of arc was built by Chapman (1974).

A precise angle measurement interferometer has been built by measuring the phase shift of a total internal reflection in a prism. This method, originally proposed by Huang (Huang et al., 1992; Huang and Ni, 1996; Huang and Li, 1998), has been improved by Chiu (Chiu and Su 1997a, 1997b).

17.3. TESTING OF PRISMS

A problem frequently encountered in the optical shop is precise angle measurement in the manufacturing of prisms. In most cases, prism angles are 90° , 45° , and 30° .

These angles are easily measured by comparison from a standard, other angles need an autocollimator or a goniometer to be accurately measured. Prism angle testing becomes very involved for precision production. Extensive work has been done and is described mainly by Twyman (1957) De Vany (1968) and Fantone (1991).

To accurately measure angles, it is important that the prism be free of pyramidal error. Suppose a prism with angles A, B and C (Fig. 17.5(a)), and let OA be perpendicular to plane ABC. If line AP is perpendicular to segment BC, then the angle AOP is a measure for the pyramidal error. In a prism with pyramidal error, the angles between the faces, as measured in planes perpendicular to the edges between these faces, adds to over 180° . To simply detect pyramidal error in a prism, Johnson (1947) and Martin (1924) suggest looking at both the refracted and the reflected image from a straight line (Fig. 17.5(b)). When pyramidal error is present, the line appears to be broken. A far target could be graduated to measure directly in minutes. A sensitivity of up to 3 minutes could be obtained.

During the milling process of a prism in a production run, a glass blank is mounted on a jig collinear to a master prism (Fig. 17.6). An autocollimator pointing to the master prism accurately sets the position for each prism faces (Twyman, 1957; DeVany, 1971). With a carefully set diamond lap, pyramidal error is minimized. In

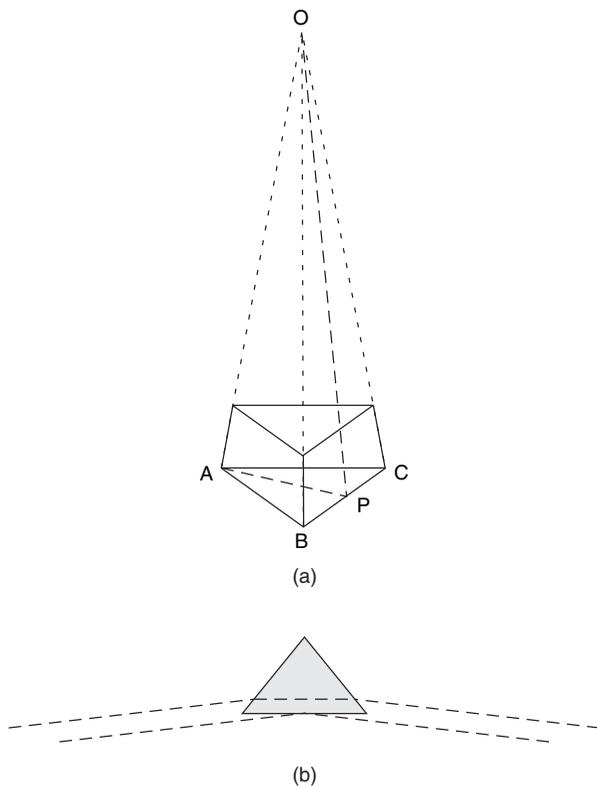


FIGURE 17.5. Pyramidal error in a prism. (a) Nature of the error; (b) Test for the error.

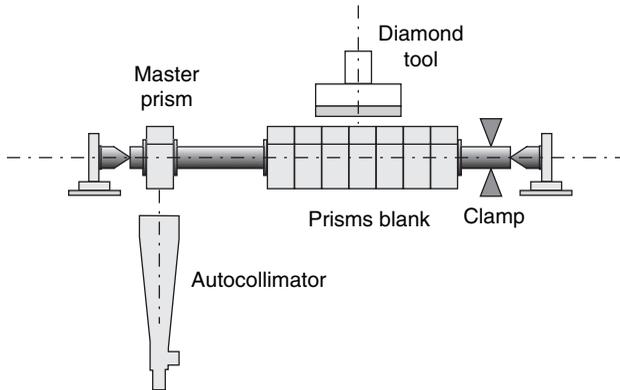


FIGURE 17.6. Milling prisms by replication.

a small quantity run, angles can be checked with a bevel gauge. Visual tests for a prism in a bevel gauge can measure an error less than a minute of arc (Noble, 1978).

A 90° angle in a prism can be measured with an autocollimator by internal reflection (Fig. 17.7(a)). At the autocollimator, two images are seen and their angular separation is being the prism angle error, although it is not possible to know the sign for the angle error. Since the hypotenuse face has to be polished and the glass must be homogeneous, the measurement of the external angle with respect to a reference flat is preferred (Fig. 17.7(b)). In this case, the angle error is determined by a change in the angle by tilting the prism. If the external angle is decreased and the images separate further, then the external angle is less than 90° . Conversely, if the images separate by tilting in such a way that the external angle increases, then the external angle is larger than 90° .

Several methods to determine the sign of the error have been proposed. DeVany (1978) suggested that when looking at the double image from the autocollimator, the image should be defocused inward; if the images tend to separate, then the angle in the prism is greater than 90° . Conversely, an outward defocusing will move the images closer to each other for an angle greater than 90° . Another way to eliminate the sign of the error in the angle is by introducing between the autocollimator and the prism a glass plate with a small wedge whose orientation is known. The wedge should cover only one half of the prism aperture. Ratajczyk and Bodnar (1966) suggested a different method using polarized light.

Right angle prisms can be measured with an autocollimator with acceptable precision (Tareev, 1985). With some practice, perfect cubes with angles more accurate than 2 s of arc can be obtained (DeVany, 1979).

An extremely simple test for the 90° angle in prisms (Johnson, 1947) is performed by looking to the retro-reflected image of the observer's pupil without any instrument. The shape of the image of the pupil determines the error, as shown in Figure 17.8. The sensitivity of this test is not very high, and may be used only as a coarse qualitative test. As shown by Malacara and Flores (1990), a small improvement in the sensitivity of this test may be obtained if a screen with a small hole is placed in front of the eye, as in Figure 17.9(a). A cross-centered on the small hole is

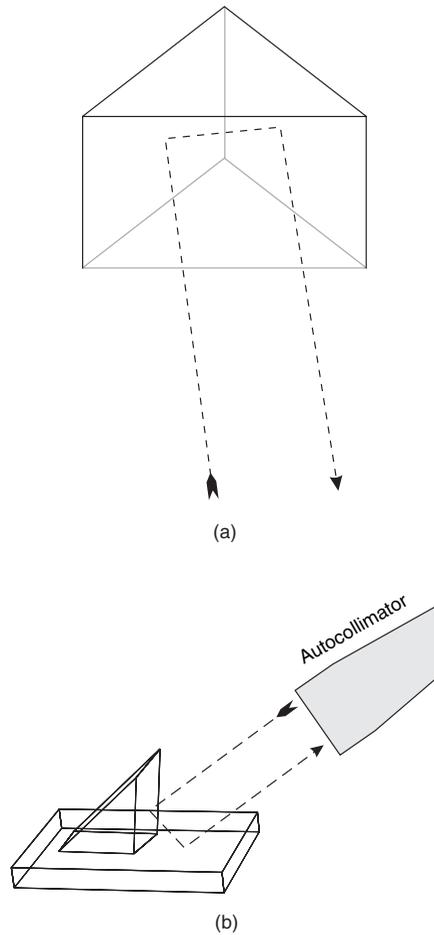


FIGURE 17.7. Right angle measurement in prisms. (a) Internal and (b) external measurements.

painted on the front face of the screen. The observed images are as shown in the same Figure 17.9(b). It is easy to see the similarity between this test and the Placido disc used some years ago in optometry and ophthalmology, for observing irregularities in the cornea of a patient.

It is interesting to notice that as opposed to the collimator test, there is no error sign uncertainty in the tests just described. The reason is that the observed plane is

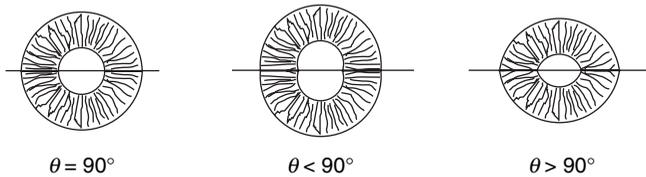


FIGURE 17.8. Retroreflected images of the observer's pupil on a 90 prism.

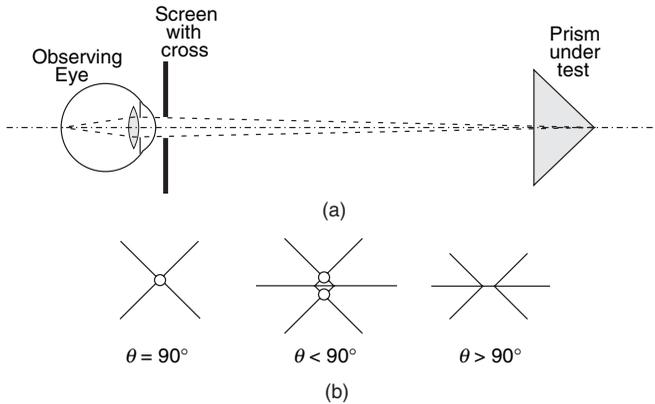


FIGURE 17.9. Testing a right angle prism. (a) Screen in front of the eye and (b) its observed images.

located where the two prism surfaces intersect. An improvement described by Malacara and Flores (1990), combining these simple tests with an autocollimator, is obtained with the instrument in Figure 17.10. In this system, the line defining the intersection between the two surfaces is out of focus and barely visible, while the reticle is in perfect focus at the eyepiece.

The bevel gauge consists of two hinged straight bars hinged at their edges by a pivot, as shown in Figure 17.11. This device may be used to generate a master prism whose angles are of 45° to about 20 arc second (Deve, 1945). To accomplish this, first, the bevel gauge is approximately set to 45° . The second step is to find two right-angle prisms with the same size, for which an angle on each of them is close to 45° , but almost equal, as far as the bevel gauge can measure. During the measurements, the gauge is held between the eye and a bright light, in order to see very small spaces.

Let us now assume that the two angles that were found are equal to 46° . These two prisms are cemented with the two equal angles together, as shown in Figure 17.11. The cemented surface is the bisector of a 92° angle. The next step is to regrind the large face until it becomes flat and the angles at the ends appear equal as measured with the bevel gauge. Under these conditions, when the two prisms are separated, the right angle should be exactly 90° to the accuracy available with the bevel gauge. The final step is to regrind the large face of each prism until the end angles are equal. Then, these angles will be exactly 45° .

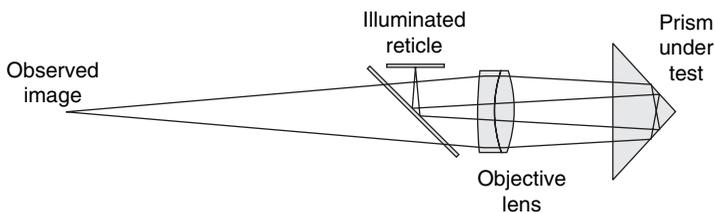


FIGURE 17.10. Modified autocollimator for testing the right angle in prisms without sign uncertainty.

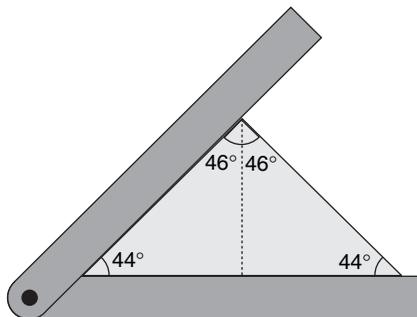


FIGURE 17.11. A bevel gauge to measure a 45° prism.

If the measured prism has a 50-mm hypotenuse, a space of $5\ \mu\text{m}$ at one end represents an angle of 0.0001 rad or 20 arc s.

17.4. RADIUS OF CURVATURE MEASUREMENTS

The curvature of a spherical optical surface or the local curvature of an aspherical surface may be measured by mechanical or optical methods, as described next. Some methods measure the sagitta, some the surface slope, and some others directly the position of the center of curvature.

17.4.1. Mechanical Measurement of Radius of Curvature

Templates. The easiest way to measure the radius of curvature is by comparing it with metal templates with different radius of curvature until the best fit is obtained. The template is held against the optical surface, with a bright light source in front of the observer and behind the template and the optical surface. If the surface is polished, openings close to one wavelength may be detected. If the opening is very narrow, the light becomes blue due to diffraction.

Test Plates. Another method for curvature measurement is to use a test plate with opposite curvature as template; Newton rings reflect the curvature difference. The use of an interferometer increases its accuracy. For this method, to be practical the surface has to be polished. (See Chapter 1).

Spherometers. This is the most popular mechanical device for measuring the radius of curvature. The value of the radius is calculated by measuring the sagitta (See Fig. 17.12). A classical spherometer consists of three equally spaced feet with a central moving plunger. The spherometer is first placed on top of a flat surface and then on top of the surface to be measured. The difference in the position of the central plunger is the sagitta of the spherical surface. Several practical problems may arise. One is that sharp legs may scratch the surface; thus, a steel ball is placed at the end of the legs as well as at the end of the plunger (Aldis spherometer). In this case if the

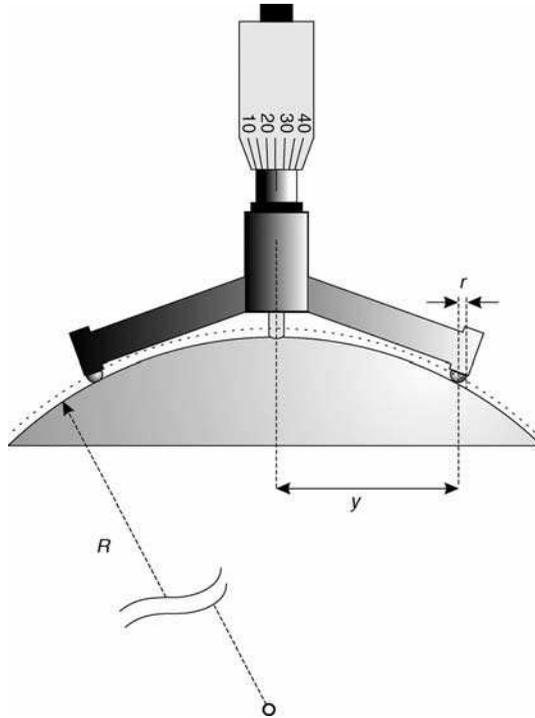


FIGURE 17.12. Three leg spherometer.

measured sagitta is z , the radius of curvature R of the surface is given by: (Cooke, 1964)

$$R = \frac{z}{2} + \frac{y^2}{2z} \pm r \quad (17.2)$$

where r is the radius of curvature of the balls. The plus sign is used for concave surfaces, and the minus sign for convex surfaces.

The precision of this instrument may be obtained by differentiating Eq. (17.2), as follows:

$$\frac{dR}{Dz} = \frac{1}{2} - \frac{y^2}{2z^2} \quad (17.3)$$

to obtain:

$$\Delta R = \frac{\Delta z}{2} \left(1 - \frac{y^2}{z^2} \right) \quad (17.4)$$

This result is valid assuming that the spherometer is perfectly built and that the dimensional parameters y and r are well known. The uncertainty comes only from the measurement of the sagitta. Noble (1978) has made an evaluation of this precision for a spherometer with $y = 50$ mm. and a sagitta reading uncertainty equal to 0.005 mm. and found the results in Table 17.1. We may see that the precision is better than 2%.

TABLE 17.1. Spherometer precision.

Radius of sphere R (mm)	Sagitta Z (mm)	Precision ΔR (mm)	Fractional precision $\Delta R/R$
10 000	0.125	-400	-0.040
5 000	0.250	-100	-0.020
2 000	0.625	-16	-0.008
1 000	1.251	-4	-0.004
500	2.506	-1	-0.002
200	6.351	-0.15	-0.0008

An extensive analysis of the precision and accuracy of several types of spherometers is given in the book by Jurek (1977).

Another type of spherometer is the so-called ring spherometer, which has a cup instead of the three legs. The cup is flat in the upper part and has its outside and its external walls with a cylindrical shape, as in Figure 17.13. A concave surface touches the external edge of the cup, whereas the convex surface touches the internal edge of the ring. Thus, Eq. (17.2) may be used if a different value of y is used for concave and convex surfaces, and r is taken as zero. In this instrument the cups may be interchangeable with different ring diameters, for different surface diameters and radii of curvature. There is an averaging effect because the ring strikes the high spots only. The main advantage is that an astigmatic deformation of the surface is easily detected, but it cannot be measured. With the three-leg spherometer, the astigmatic deformations cannot be even detected.

The spherometer accuracy may be improved in many ways, by different methods of taking the readings of the sagitta. One method is the Steinheil spherometer (Martin, 1924) in which a mechanical device is employed to indicate the pressure between the central plunger and the surface to be measured. In the Abbe spherometer (Martin, 1924), the displacement of the central plunger is measured with the aid of a

**FIGURE 17.13.** Ring spherometer.

scale and a reading microscope. The dial spherometer is very popular for quick measurements in industrial processes. The movement of the plunger activates the hand of a circular measuring dial.

Some modern spherometers use a differential transformer as a transducer to measure the plunger displacement. This transformer coupled to an electronic circuit produces a voltage linear with the plunger displacement. This voltage is then analyzed by a microprocessor. The microprocessor then calculates the radius or curvature in any desired units and displays it. In another case, a precise laser distance measuring equipment is used for sagitta measurement (Soares and Fernández, 1988).

The bar spherometer permits the measurement of toric or cylindrical surfaces, since it measures the curvature along any diameter. However, the accuracy of the measurement may not be as high as in the previous devices, due to tilting of the instrument. A commercial version of a small bar spherometer for the specific application in optometric work is the Geneva gauge (Fig. 17.14). In this gauge the scale is directly calibrated in diopters since the power is linear with the sagitta, assuming a nominal refractive index of 1.53 for the glass.

A problem indirectly related is the calculation of the sagitta from a knowledge of the radius of curvature, in order to mount a lens properly in its mechanical holder. Different graphical and algebraic procedures have been described (Zanker, 1981; Foote, 1981; Brixner, 1982).

17.4.2. Optical Measurement of Radius of Curvature

Foucault Test. One of the easiest optical methods to find the radius of curvature of a concave surface is by means of the knife-edge test (See Chapter 8) in order to locate the center of curvature. Then, the distance from the knife-edge to the optical surface is measured with a scale. The accuracy of this method depends on the distance measurement tool. A measuring tape may be a bad tool because the scale may sag due to gravity. Foucault test can be used to determine the radius of curvature of both



FIGURE 17.14. A Geneva gauge.

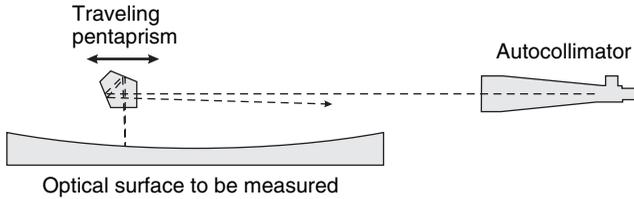


FIGURE 17.15. Autocollimator and pentaprism used to determine the radius of curvature by measuring surface slopes.

surfaces in a convergent lens as suggested by Hugo and Lessing, (1964). In this method – suitable for long radius of curvature – a light from a Foucault tester is refracted in the first lens surface and then reflected in the second surface. From this, a distance from the tester-lens is obtained; then the lens is reversed and a new tester-lens distance is obtained. If the glass refractive index and lens thickness is known, both curvatures can be known.

Autocollimator. In the autocollimator technique, (Horne, 1972) the radius of curvature is determined through the measurements of the slopes of the optical surface. The well-known property that a pentaprism deflects a light beam 90° , independently of small errors in its orientation, is used. This method, appropriate for large aperture optical surfaces, is illustrated in Figure 17.15, where we may see that the pentaprism travels over the optical surface to be measured, along one diameter. The first step is to center the light on the reticle of the autocollimator when the vertex of the surface is being examined. Then, the pentaprism is moved outside the central part of the surface in order to measure the slope variations. From these measurements, the radius of curvature and even its shape may be calculated. This method is used only for large radii of curvature, and it is equally applicable for both, concave and convex surfaces.

Confocal Cavity Technique. The optical cavity technique (Gerchman and Hunter, 1979 and 1980) permits the interferometric measurement of very long radii of curvature with an accuracy of 0.01%. The method consists of forming the cavity of a Fizeau interferometer (See Chapter 1) as illustrated in Figure 17.16. This is a confocal cavity of n th order, where n is the number of times the path is folded. The radius of curvature is approximately $2n$ times the cavity length Z_n , but for a higher accuracy values listed in Table 17.2 should be used.

Traveling Microscope. This is one of the most popular methods for measuring the radius of curvature of small concave optical surfaces, with short radius of curvature. As shown in Figure 17.17 a point light source is produced at the front focus of a microscope objective. This light source illuminates the concave optical surface to be measured, near its center of curvature. Then, this concave surface forms an image, also close to its center of curvature. This image is observed with the same microscope used to illuminate the surface. During this procedure the microscope is focused both at the center of curvature and at the surface to be measured. A sharp image of the light source is observed at both places. The radius of curvature is the distance between these two positions for the microscope.

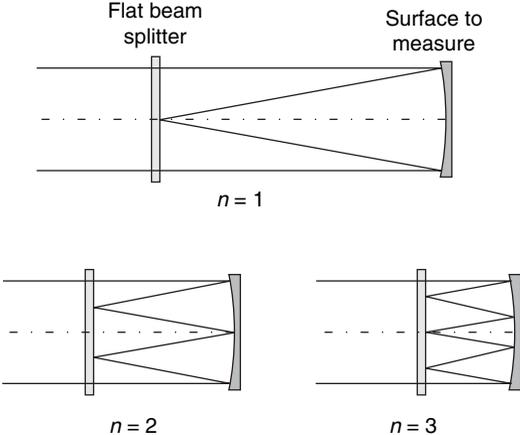


FIGURE 17.16. Confocal cavity arrangements used to measure radius of curvature.

This distance traveled by the microscope may be measured on a scale, to obtain a precision of about ± 0.1 mm. If a bar micrometer is used, the precision may be increased by an order of magnitude. In this case two small convex buttons are required, one fixed to the microscope carriage and the other to the stationary part of the bench. They must face each other when the microscope carriage is close to the optical bench fixed component.

Carnell and Welford (1971) describe a method that requires only one measurement. The microscope is focused only at the center of curvature. Then, the radius of curvature is measured by inserting a bar micrometer with one end touching the vertex of the optical surface, and the other end is adjusted until it is observed in focus on the microscope. Accuracies of a few microns are obtained with this method.

In order to focus the microscope properly, the image of an illuminated reticle must fall after reflection on the same reticle itself.

A convex surface may also be measured with this method, if a well-corrected lens with a conjugate longer than the radius of curvature of the surface under test is used (O’Shea

TABLE 17.2. Constants relating cavity length to radius of curvature $Z_0 = C_n R$.

N	C_n
1	0.5
2	0.25
3	0.1464466
4	0.0954915
5	0.0669873
6	0.0495156
7	0.00380603
8	0.0301537

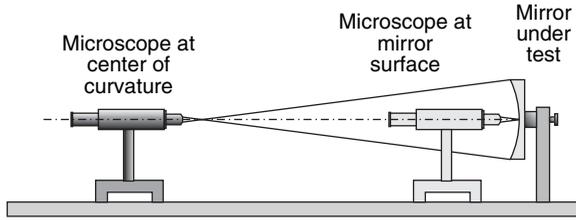


FIGURE 17.17. Traveling microscope to measure radii of curvature.

and Tilstra, 1988). Another alternative for measuring convex surfaces is by inserting an optical device with prisms in front of the microscope, as described by Jurek (1977).

Some practical aspects of the traveling microscope are examined by Rank (1946), who obtained a dark field at focus with an Abbe eyepiece, which introduces the illumination with a small prism. This method has been implemented using a laser light source by O'Shea and Tilstra (1988).

A traveling microscope device for curvature measurement can be improved for high-precision by bringing the focus determination up to the diffraction limit and the position determination by the use of a coordinate measuring machine (Jiaming et al., 1995).

Additional optical methods to measure the radius of curvature of a spherical surface have been described. Evans (1971, 1972a, and 1972b) determines the radius by measuring the lateral displacements on a screen of a laser beam reflected on the optical surface, when this optical surface is laterally displaced. Cornejo-Rodríguez and Cordero-Dávila (1980), Klingsporn (1983) and Díaz-Urbe et al. (1986) rotate the surface about its center of curvature in a nodal bench.

17.5. FOCAL LENGTH MEASUREMENTS

Two distances from the focal point are commonly used; one is the *back focal length*, defined as the distance from the last surface of the system to the focus. The other is the *focal distance* defined as the distance from the principal plane to the focal point. The back focal length is easily measured following the same procedure used for measuring the radius of curvature, using a microscope and the lens bench (Chandler, 1974). Additionally, for photo and video camera lenses, the flange focal distance is defined as the distance from the focal point to mounting flange in an optical system. Flange focal distance can be measured with an autocollimator (Foote and Townsley, 1947).

17.5.1. Nodal Slide Bench

For an optical system in air, the principal points (intersection of the principal plane and the optical axis) coincide with the nodal points. Thus, to locate this point we may use the well-known property that small rotations of the lens about an axis, perpendicular to the optical axis and passing through the nodal point, do not produce any

lateral shift of the image. The instrument used to perform this procedure, is called an optical nodal slide bench (Kingslake, 1932). The use of a nodal slide is described in Chapter 11: *Star test* in this book. Besides a precise measurement of the focal distance, one can measure of others parameters like aberrations. A nodal slide bench can be used also for measurements of reflecting surfaces (Cornejo-Rodríguez and Cordero-Dávila, 1980) and even toric surfaces (Díaz Uribe et al., 1986).

17.5.2. Focimeters

A focimeter is an instrument designed to measure the focal length of lenses in a simple manner. It is a commonly used instrument by optometrists. Actually, a focimeter or vertometer measures the back focal length or vertex power. A typical focimeter is shown in Figure 17.18(a). When there is no lens in the measuring position (Zero dioptric power), a light source illuminates a target (reticle) and a convergent lens with focal length f is placed at a distance x from the reticle. The reticle is imagined at infinity and a telescope permits the observation of the target at that position. The lens to be measured is placed at a distance d from the converging lens (Fig. 17.18b). The position of the target x is variable, and is adjusted until the light beam going out from the lens under test becomes collimated. This collimation is verified by means of the small telescope in front of this lens, focused to infinity. The values of d and the focal length f are chosen to be equal. Then, the back focal length f_b of the lens under test is given by thin lens formula:

$$\frac{1}{f_b} = \frac{1}{d} - \frac{x}{d^2} = P_v \quad (17.5)$$

where P_v is the vertex power defined as $P_v = 1/f_b$.

From here we see that the power of the lens being measured is linear with the distance x . Also, we can see that the measurement range of the focimeter goes from a maximum positive power $P_{max} = 1/d$ without a limit for negative powers. There are

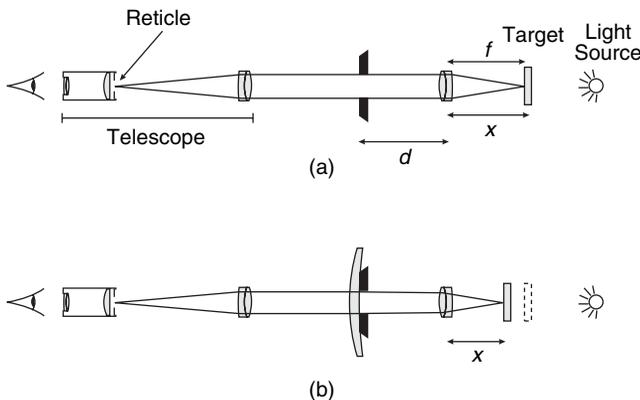


FIGURE 17.18. Focimeter (a) without test lens and (b) with test lens.

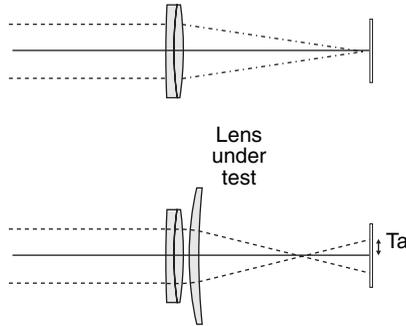


FIGURE 17.19. Focal length determination by transverse aberration measurements.

many variations of this instrument. Some modern focimeters measure the lateral deviation of a light ray from the optical axis, (transverse aberration) as in Figure 10.21 (Chapter 10), when a defocus is introduced (Evans 1971, 1972a, 1972b and Bouchaud and Cogno, 1982). This method is mainly used in some modern automatic focimeters for optometric applications.

The power error determination in a focimeter can be easily obtained by derivation of Eq. 17.5:

$$\delta P_v = -\frac{\delta x}{f_c^2} \quad (17.6)$$

hence, the power error is linear with the target position error (Martínez-Corral et al., 1998) and decreases with the square of the collimating lens focal distance.

To measure the transverse aberration, a position-sensing detector is frequently used. An array of holes in a Hartmann-like test can be used. In particular an array of four holes and a position sensing detector can be used to measure the lens vertex power (Malacara and Malacara, 1992) as shown in Figure 17.19. With this method it is possible to measure Vertex power, prism and cylinder for an ophthalmic lens. A full Hartman screen can be used to obtain a power map of an ophthalmic lens (Salas-Peimbert et al., 2004). Power map is obtained for a progressive lens by scanning the lens aperture with a laser beam (Castellini et al., 1994), (Gnavo et al., 2002).

17.5.3. Other Focal Length Measurements

- (a) *Moiré deflectometry*. Moiré deflectometry uses collimated light to impinge in two Ronchi rulings (Fig. 17.20). Convergent or divergent beams of light produce a rotation in the fringe orientation of a moiré pattern. This rotation of magnitude α is related to the focal distance by (Kafri and Glatt, 1990)

$$f \approx \frac{d}{\theta \tan \alpha} \quad (17.7)$$

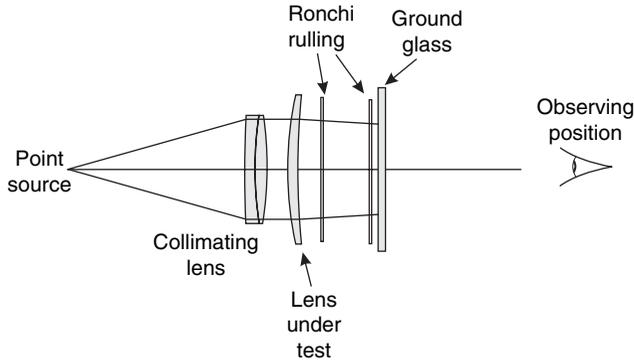


FIGURE 17.20. Set-up for a moiré deflectometry lens power measurement.

where d is the ruling's line pitch, θ is the angle between the ruling's lines, and α is the rotation angle of the moiré pattern (Fig. 17.21). This method is useful for long focal distance lenses (> 1 m). A variant of this method is shown in Figure 17.22 (Glatt and Kafri, 1987). This setup looks for a collimated beam of light that is achieved through the mirror displacement. Under perfect collimation, the mirror position defines the focal point position. Back focal distance could be measured by measuring the distance between the mirror position and the last lens surface. This new set-up allows the focal distance measurement for a lens of any focal distance, not only long focal distances.

- (b) *Talbot Autoimages.* Talbot autoimages are produced by periodic Ronchi rulings in a coherent beam of light (see Chapter 9.). Talbot method is done with a coherent collimated beam of light illuminating a periodic amplitude or phase grating. Periodically and equally spaced, an image of the grating will be produced as shown in Figure 17.23. Every Talbot autoimage is considered an object for the lens that produces an image of the same autoimage at the lens image space. Two methods are used to find the focal lens: One is to put another ronchi ruling coincident with one of the autoimages (Malacara-Doblado and Malacara-Hernández, 1996) (Nakano and Murata, 1985).

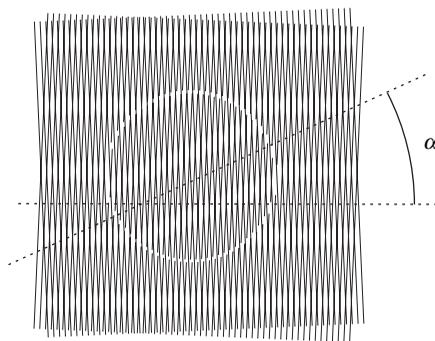


FIGURE 17.21. Moiré pattern for a lens in a moiré deflectometer.

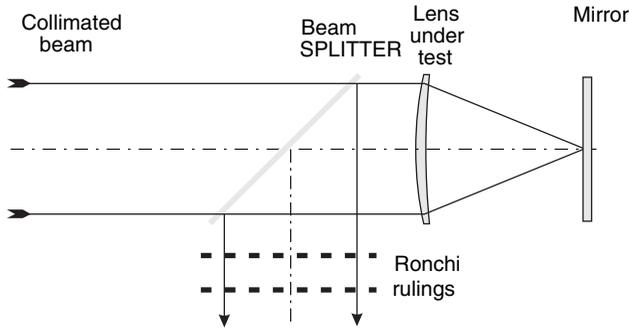


FIGURE 17.22. Moiré deflectometry set-up for long and short focal length measurement.

A moiré interferometer results with the same procedure described in the moiré deflectometer case. Another method consists in measuring the position of several autoimages. In this case, the lens power can be measured, not just the back focal length (Bernardo and Soares, 1988)

- (c) *Fourier Transforms.* Fourier transforming properties of a lens can be used to determine the focal length of a lens. In one case (Horner, 1989), a slit is placed in the object plane and the diffraction pattern is analyzed on the image plane. By measuring the diffraction pattern maxima and spacing, the focal length can be determined. One advantage of this method is that the incoming beam of light does not have to be perfectly collimated.
- (d) *Microlenses.* Microlenses impose several practical problems in focal distance measurement. The main problem is related to the small distances to be measured. A proposed solution is by the analysis of the propagation of a Gaussian beam of a laser (Camacho et al., 2000) (Camacho et al., 2002). A lens is placed in the beam waist. From here, it is possible to know the focal length of a lens when the shape of the propagating is measured.
- (e) *Fiber Optics Bundle.* A clever method used to automatically find the position of the focus has been described by Howland and Proll (1970). They used a fiber optics bundle mixed to illuminate at the focal plane of a lens, mixed to another bundle to collect the retroreflected beam of light in an autocollimating configuration.

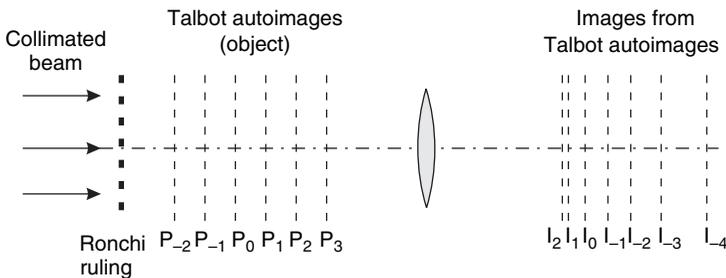


FIGURE 17.23. Talbot autoimages in a lens.

REFERENCES

- Bernardo L. M., and O.D.D. Soares, "Evaluation of the Focal Distance of a Lens by Talbot Interferometry," *Appl Opt*, **27**, 296–301 (1988).
- Bouchaud P. and J. A. Cogno, "Automatic Method for Measuring Simple Lens Power," *Appl. Opt.*, **21**, 3068–3068 (1982).
- Brixner B., "Easier Way to Find the Sagittal Depth: Comments," *Appl. Opt*, **21**, 976–976 (1982).
- Camacho A. A., C. Solano, M. Cywiak, G. Martínez-Ponce, and R. Baltasar, "Method for the Determination of the Focal Length of a Microlens," *Opt. Eng.*, **39**, 2149–2152 (2000).
- Camacho A. A., C. Solano, G. Martínez-Ponce, and R. Baltasar, "Simple Method to Measure the Focal Length of a Lenses," *Opt. Eng.*, **41**, 2899–2902 (2002).
- Carnell K. H. and W. T. Welford, "A Method for Precision Spherometry of Concave Surfaces," *J. Phys.*, **E4**, 1060–1062 (1971).
- Castellini C., F. Francini, and B. Tiribilli, "Hartmann test modification for measuring ophthalmic progressive lenses," *Appl. Opt.*, **33**, 4120–4124 (1994).
- Chandler C. H., "A Lens-Measuring Instrument-Engineered for People," *Opt. Eng.*, **13**, 275–278 (1974).
- Chapman G. D., "Interferometric Angular Measurement," *Appl. Opt.*, **13**, 1646–1651 (1974).
- Chiu M.-H. and D.-C. Su, "Angle Measurement Using Total-Internal-Reflection Heterodyne Interferometry," *Opt. Eng.*, **36**, 1750–1753 (1997a).
- Chiu M.-H. and D.-C. Su, "Improved Technique for Measuring Small Angles," *Appl. Opt.*, **36**, 7104–7105 (1997b).
- Choi J., G. M. Perera, M. D. Aggarwal, R. P. Shukla, and M. V. Mantravadi, "Wedge-Plate Shearing Interferometers for Collimation Testing: Use of a Moiré Technique," *Appl. Opt.*, **34:19**, 3628–3638 (1995)
- Cooke F., "The Bar Spherometer," *Appl. Opt.*, **3**, 87–88 (1964).
- Cornejo-Rodriguez A. and A. Cordero-Davila, "Measurement of Radii of Curvature of Convex and Concave Surfaces Using a Nodal Bench and a He-Ne Laser," *Appl. Opt.*, **19**, 1743–1745 (1980).
- DeVany A. S., "Making and Testing Right Angle and Dove Prisms," *Appl. Opt.*, **7**, 1085–1087 (1968)
- DeVany A. S., "Reduplication of a Penta-Prism Angle Using Master Angle Prisms and Plano Interferometer," *Appl. Opt.*, **10**, 1371–1375 (1971).
- DeVany A. S., "Setting the Optimum Focus of the Pinhole or Reticle of a Collimator Lens," *Appl. Opt.*, **15:5**, 1104 (1976).
- DeVany A. S., "Testing Glass Reflecting-Angles of Prisms," *Appl Opt.*, **17**, 1661–1662 (1978).
- DeVany A. S., "Near Perfect Optical Square," *Appl. Opt.*, **18**, 1284–1286 (1979).
- Deve C., *Optical Workshop Principles* (Translated by T. L. Tippell), Hilger and Watts, London 1945.
- Diaz-Urbe R., J. Pedraza-Contreras, O. Cardona-Núñez, A. Cordero-Dávila, and A. Cornejo Rodriguez, "Cylindrical Lenses: Testing and Radius of Curvature Measurement," *Appl. Opt.*, **25**, 1707–1709 (1986).
- Evans J. D., "Method for Approximating the Radius of Curvature of Small Concave Spherical Mirrors Using a He-Ne Laser," *Appl. Opt.*, **10**, 995–996 (1971).

- Evans J. D., "Equations for Determining the Focal Length of On-Axis Parabolic Mirrors by He-Ne Laser Reflection," *Appl. Opt.*, **11**, 712–714 (1972a).
- Evans J. D., "Error Analysis to: Method for Approximating the Radius of Curvature of Small Concave Spherical Mirrors Using a He-Ne Laser," *Appl. Opt.*, **11**, 945–946 (1972b).
- Fantone, S. D., *Optics Cooke Book*, 2nd ed. Optical Society of America, Washington, D. C., 1991.
- Foote V. S., "Easier Way to Find the Sagitta Depth," *Appl. Opt.*, **20**, 2605–2605 (1981).
- Foote P. C. and Townsley, M. G., "Autocollimator for Precise Measurement of the Flange Focal Distance of Photographic Lenses," *J. Opt. Soc. Am.*, **37**, 42–42 (1947).
- Geiser R. D., "Precision and Accuracy," in: R. Kingslake (Ed), *Applied Optics and Optical Engineering*, vol. I, Academic Press, New York, 1965, Chapter 11.
- Gerchman M. C. and G. C. Hunter, "Differential Technique for Accurately Measuring the Radius of Curvature of Long Radius Concave Optical Surfaces," *Proc. SPIE*, **192**, 75–84 (1979).
- Gerchman M. C. and G. C. Hunter, "Differential Technique for Accurately Measuring the Radius of Curvature of Long Radius Concave Optical Surfaces," *Opt. Eng.* **19**, 843–848 (1980).
- Glatt I. and O. Kafri, "Determination of the Focal Length of Nonparaxial Lenses by Moiré-Deflectometry," *Appl. Opt.*, **26**, 2507–2508 (1987).
- Gnavo G., Z. Y. Wu, J. L. de Bougrenet de la Tocnaye and L. Liu, "Large-Aperture Automatic Focimeter for the Measurement of Optical Power and Other Optical Characteristics of Ophthalmic Lenses," *Appl. Opt.*, **41**, 5997–6005 (2002).
- Horne D. F., *Optical Production Technology*, Adam Hilger, London and Crane Russak, New York, 1972, Chapter XII.
- Horne D. F., *Dividing, Ruling and Mask Making*, Adam Hilger, London, 1974, Chapter VII.
- Horner J. L. "Collimation Invariant Technique for Measuring the Focal Length of a Lens," *Appl. Opt.*, **28**, 1047–1047 (1989).
- Howland, B., and A. F. Proll, "Apparatus for the Accurate Determination of Flange Focal Distance," *Appl. Opt.*, **11**, 1247–1251 (1970).
- Huang P. S., S. Kiyono, and O. Kamada, "Angle Measurement Based on the Internal-Reflection Effect: A New Method," *Appl. Opt.*, **31**, 6047–6055 (1992).
- Huang P. S. and J. Ni, "Angle Measurement Based on the Internal-Reflection Using Elongated Critical-Angle Prisms," *Appl. Opt.*, **35**, 2239–2241 (1996).
- Huang P. S. and Y. Li, "Small-Angle Measurement by Use of a Single Prism," *Appl. Opt.*, **37**, 6636–6642 (1998).
- Hugo T. J. and N. V. D. W. Lessing, "Determination of Long Radii of Curvature of Positive Lenses," *Appl. Opt.*, **3**, 483–486 (1964).
- Hume K. J., *Metrology with Autocollimators*, Hilger and Watts, London, 1965.
- Johnson B. K., *Optics and Optical Instruments*, Dover, New York, 1947, Chapters II and VIII.
- Jiaming L, S. Datu, Y. Fen and S. Dingguo, "Research on High Precision Measuring Technique for the Curvature Radius of Concave Spherical Surface," *Proc. SPIE*, **2536**, 489–497 (1995).
- Jurek B., *Optical Surfaces*, Elsevier Scient. Publ. Co., New York, 1977.
- Kafri, O. and I. Glatt, *The Physics os Moire Metrology*, Wiley Interscience, New York, 1990.

- Kingslake R., "A New Bench for Testing Photographic Lenses," *J. Opt. Soc. Am.*, **22**, 207–222 (1932).
- Klingsporn P. E., "Use of a Laser Interferometric Displacement-Measuring System for Non-contact Positioning of a Sphere on a Rotation Axis Through its Center and for Measuring the Spherical Contour," *Appl. Opt.*, **18**, 2881–2890 (1983).
- Kothiyal M. P. and R. S. Sirohi, "Improved Collimation Testing using Talbot Interferometry," *Appl. Opt.*, **26**, 4056–4057 (1987).
- Leppelmier G. W. and D. J. Mullenhoff, "A Technique to Measure the Wedge Angle of Optical Flats," *Appl. Opt.*, **9**, 509–510 (1970).
- Malacara D. and O. Harris, "Interferometric Measurement of Angles," *Appl. Opt.*, **9**, 1630–1633 (1970).
- Malacara D. and R. Flores, "A Simple Test for the 90 Degrees Angle in Prisms," *Proc. SPIE*, **1332**, 36–40 (1990).
- Malacara D. and Z. Malacara, "Testing and centering of lenses by means of a Hartmann test with four holes," *Opt. Eng.*, **31**, 1551–1555 (1992).
- Malacara-Doblado D. and D. Malacara-Hernández, "Measuring Convergence or Divergence Power with Moiré Fringes," *Proc SPIE*, **2860**, 390–393 (1996).
- Martin L. C., *Optical Measuring Instruments*, Blackie and Sons Ltd., London, 1924.
- Martínez Corral, M, W. D. Furlan, A. Pons and G Saavedra, *Instrumentos Ópticos y Optométricos. Teoría y Prácticas*, Universitat de Valencia, Valencia, 1998.
- Met V., "Determination of Small Wedge Angles Using a Gas Laser," *Appl. Opt.*, **5**, 1242–1244 (1966).
- Murty M. V. R. K., "The Use of a Single Plane Parallel Plate as a Lateral Shearing Interferometer with a Visible Gas Laser Source," *Appl. Opt.* **3:4**, 531–534 (1964)
- Nakano Y. and K. Murata, "Talbot Interferometry for Measuring the Focal length of a Lens," *Appl. Opt.*, **24:19** (1985).
- Noble R. E. "Some Parameter Measurements," in: D. Malacara (Ed) *Optical Shop Testing*, 1st Edition, John Wiley and Sons, New York, 1978.
- O'Shea D. C. and S. A. Tilstra, "NonContact Measurements of Refractive Index and Surface Curvature," *Proc. SPIE*, **966**, 172–176 (1988).
- Rank D. H., "Measurement of the Radius of Curvature of Concave Spheres," *J. Opt. Soc. Am.*, **36**, 108–110 (1946).
- Ratajczyk F. and Z. Bodnar, "An Autocollimation Measurement of the Right Angle Error with the Help of Polarized Light," *Appl. Opt.*, **5**, 755–758 (1966).
- Salas-Peimbert D. P., G. Trujillo-Schiaffino, D. Malacara-Hernández, D. Malacara-Doblado and S. Almazán-Cuellar, "Ophthalmic Lenses Measurement Using Hartmann Test," *Proc. SPIE*, **5622**, 102–106 (2004).
- Shi P. and E. Stijns, "New Optical Method for Measuring Small Angle Rotations," *Appl. Opt.*, **27**, 4342–4344 (1988).
- Soares O. D. D. and J. F. Fernández, "Laser-Aided Spherometer," *Proc. SPIE*, **954**, 234–240 (1988).
- Sriram, K. V., P. Senthikumar, M. P. Kothiyal, and R. S. Sirohi, "Double wedge plate shearing interferometer for collimation testing: new configurations," *Appl. Opt.* **32**, 4199–4203 (1993).

- Stijns E., "Measuring Small Rotation Rates with a Modified Michelson Interferometer," *Proc. SPIE*, **661**, 264–266 (1986).
- Tareev A. M., "Testing the Angles of High-Precision Prisms by Means of an Autocollimator and a Mirror Unit," *Sov. J. Opt. Technol.*, **52**, 50–52 (1985).
- Tentori D. and M. Celaya, "Continuous Angle Measurement with a Jamin Interferometer," *Appl. Opt.*, **25**, 215–220 (1986).
- Thurston T. H., "Specifying Electronic Autocollimators," *Proc. SPIE*, **661**, 399–401 (1986).
- Twyman F., *Prisms and Lens Making*, 2nd ed., Hilger and Watts, London, 1957.
- Wasilik, J. H., T. V. Bloomquist, and C. S. Willett, "Measurement of Parallelism of the Surfaces of a Transparent Sample Using Two-Beam Nonlocalized Fringes Produced by a Laser," *Appl. Opt.*, **10**, 2107–2112 (1971).
- Young A. W., "Optical Workshop Instruments," in: R. Kingslake, (Ed.), *Applied Optics and Optical Engineering*, vol 4, Academic Press, New York, 1967, Chapter 7.
- Zanker A. "Easy Way to Find the Saggita Depth," *Appl. Opt.*, **20**, 725–726 (1981).

18

Mathematical Representation of an Optical Surface and Its Characteristics

D. Malacara

18.1. DEFINITION OF AN OPTICAL SURFACE

Aspherical surfaces are extremely important in optical systems and have been studied and described by many authors, for example, by Schulz (1988) and Shannon (1980). Of special interest are surfaces with rotational symmetry, which may be defined by means of the following relation, taking the z axis as the axis of revolution:

$$z = \frac{cS^2}{1 + [1 - (K + 1)c^2S^2]^{1/2}} + A_1S^4 + A_2S^6 + A_3S^8 + A_4S^{10} \quad (18.1)$$

where $S^2 = x^2 + y^2$ and $c = 1/r = 1/\text{radius of curvature}$. Also, $A_1, A_2, A_3,$ and A_4 are the aspheric deformation constants, and K is a function of the eccentricity of a conic surface ($K = -e^2$), called the conic constant. If the A_i are all zero, the surface is a conic surface of revolution, according to the following Table 18.1:

In an ellipsoid with rotational symmetry about the z axis, the excentricity is defined as

$$e = \frac{\sqrt{a^2 - b^2}}{a} \quad (18.2)$$

where a is the major semi-axis, along the z axis and b is the minor semi-axis, along the y axis. For the case of the oblate spheroid, the same definition (18.2) holds but then a is the minor semiaxis, along the z axis and b is the major semi-axis, along the y axis. Thus, the excentricity is an imaginary number and $K > 0$, since $a < b$.

As it is easy to see, the conic constant is not defined for a flat surface ($c = 0$). An alternative expression for the conic constant in terms of the vertex radius of curvature

TABLE 18.1. Values of conic constants for conic surfaces.

Type of conic	Conic constant value
Hyperboloid	$K < -1$
Paraboloid	$K = -1$
Prolate spheroid or ellipsoid: (ellipse rotated about its major axis)	$-1 < K < 0$
Sphere	$K = 0$
Oblate spheroid (ellipse rotated about its minor axis)	$K > 0$

r and the distance L from the center of curvature to the focus of the conic optical surface located between the center of curvature and the optical surface is

$$K = -\frac{L^2}{(L - r)^2} \tag{18.3}$$

The sign of L and the corresponding values of K for different values of L and r are graphically described in Figure 18.1.

For conic surfaces of revolution there is an expression for z somewhat simpler than the general expression (18.1)

$$z = \frac{1}{K + 1} \left[r - \sqrt{r^2 - (K + 1)s^2} \right] \tag{18.4}$$

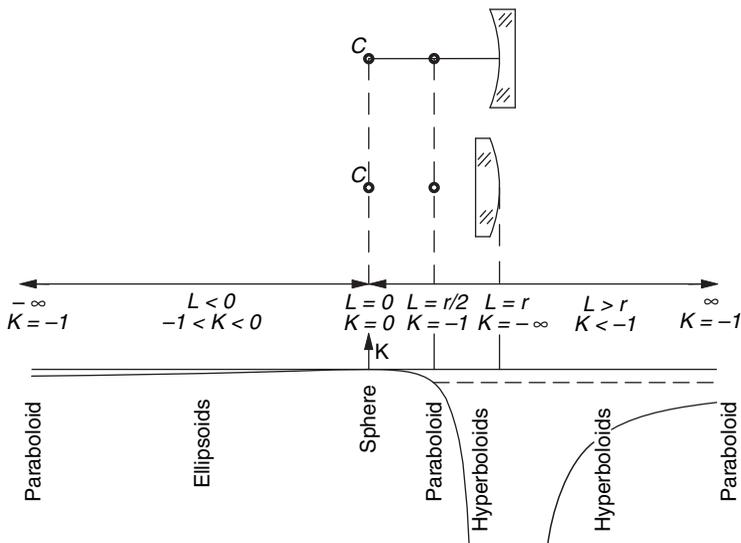


FIGURE 18.1. Values of the conic constant K as a function of the vertex radius of curvature and the distance L from the center of curvature of the conic surface to one of the foci.

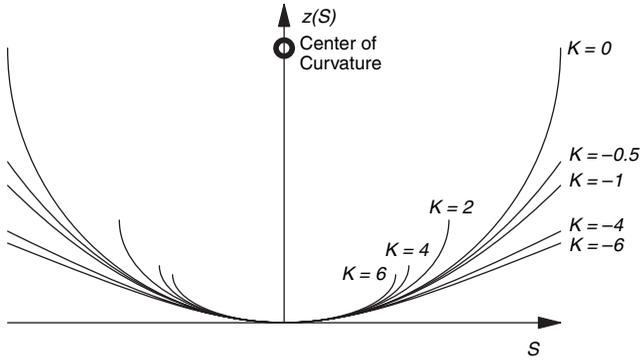


FIGURE 18.2. Profiles for aspheric surfaces with the same curvature but different conic constants.

This serves for all conics except the paraboloid, where

$$z = \frac{S^2}{2r} \tag{18.5}$$

An illustration of the profiles of some conic optical surfaces are in Figure 18.2. A surface of the axicon type (McLeod, 1960) that has the shape of a cone may be represented by means of a hyperboloid with an extremely large curvature, as shown in Figure 18.3, thus obtaining

$$K = -(1 + \tan^2 \theta) < -1 \tag{18.6}$$

and

$$c = \frac{1}{(K + 1)b} \tag{18.7}$$

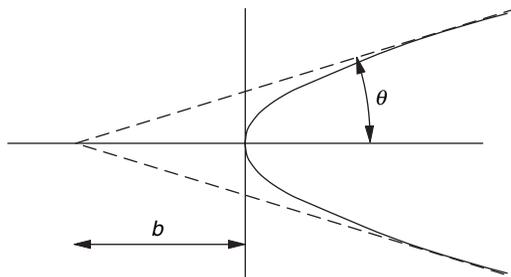


FIGURE 18.3. Axicon surface.

18.1.1. Parameters for Conic Surfaces

The positions of the foci for the conic surfaces are functions of r and K and are given by the following relations, as illustrated in Figure 18.4:

$$d_1 = \frac{r}{K + 1} \tag{18.8}$$

$$d_2 = \frac{r}{K + 1} (2\sqrt{K}) \tag{18.9}$$

$$d_3, d_4 = \frac{r}{K + 1} (1 \pm \sqrt{-K}) \tag{18.10}$$

$$d_5 = \frac{r}{2}, \tag{18.11}$$

$$d_6, d_7 = \frac{r}{K + 1} (\sqrt{-K} \pm 1) \tag{18.12}$$

18.1.2. Some Useful Expansions of z

Sometimes it is convenient to consider an aspheric or conic optical surface as the sum of its osculating sphere plus some deformation terms, as follows:

$$z = \frac{cS^2}{1 + [1 - c^2S^2]^{1/2}} + B_1S^4 + B_sS^6 + B_3S^8 + B_4S^{10} \tag{18.13}$$

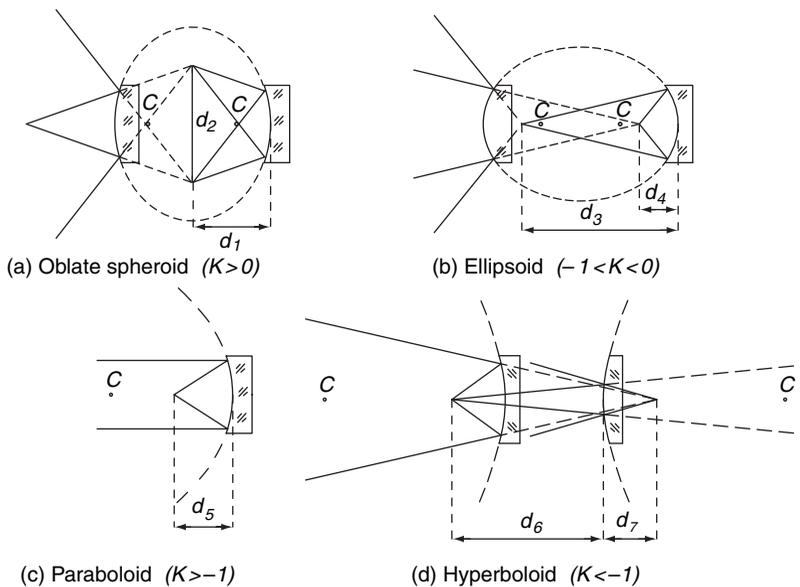


FIGURE 18.4. Some parameters for conic surfaces. The vertex center of curvature is labeled as C.

where

$$\begin{aligned} B_1 &= A_1 + \frac{[(K+1) - 1]c^3}{8}, \\ B_2 &= A_2 + \frac{[(K+1)^2 - 1]c^5}{16}, \\ B_3 &= A_3 + \frac{5[(K+1)^3 - 1]c^7}{128}, \\ B_4 &= A_4 + \frac{7[(K+1)^4 - 1]c^9}{256}, \end{aligned} \quad (18.14)$$

or some other times this expression is preferred:

$$z = D_2S^2 + D_4S^4 + D_6S^6 + D_8S^8 + D_{10}S^{10}, \quad (18.15)$$

where

$$D_2 = \frac{c}{2}, \quad (18.16)$$

$$D_4 = \frac{c}{2} \left(\frac{c}{2}\right)^2 + B_1 = \frac{c^3}{8} + B_1, \quad (18.17)$$

$$D_6 = c \left(\frac{c}{2}\right)^4 + B_2 = \frac{c^5}{16} + B_2, \quad (18.18)$$

$$D_8 = \frac{5c}{2} \left(\frac{c}{2}\right)^6 + B_3 = \frac{5c^7}{128} + B_3, \quad (18.19)$$

$$D_{10} = \frac{14c}{2} \left(\frac{c}{2}\right)^8 + B_4 = \frac{7c^9}{256} + B_4 \quad (18.20)$$

18.1.3. Aberration of the Normals to the Surface

A normal to the optical surface intersects the optical axis at a distance Z_n from the vertex of the surface. To compute this distance, we need to know the value of the derivative of z with respect to S , which is given by

$$\frac{dz}{dS} = \frac{cs}{[1 - (K+1)c^2S^2]^{1/2}} + 4A_1S^3 + 6A_2S^5 + 8A_3S^7 + 10A_4S^9 \quad (18.21)$$

The distance Z_n is then given by

$$Z_n = \frac{S}{dz/dS} + z \quad (18.22)$$

and as shown by Buchroeder et al. (1972), for conic surfaces, this expression becomes

$$Z_n = \frac{1}{c} - Kz \quad (18.23)$$

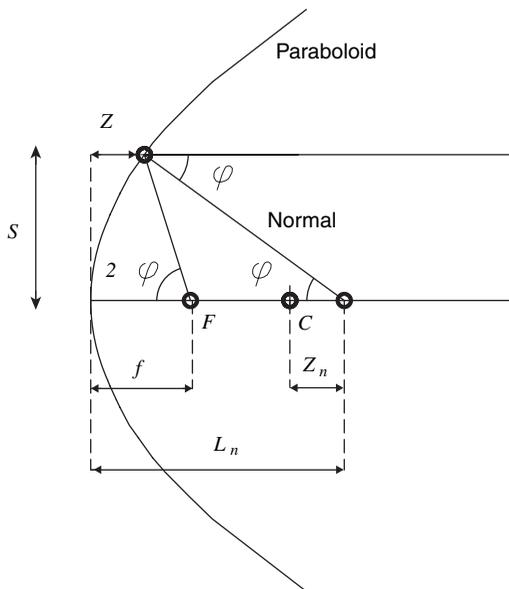


FIGURE 18.5. Rays reflection in a reflecting concave paraboloid.

It is interesting to see that for the case of a paraboloid ($K = -1$) as shown in Figure 18.5 this aberration of the normals becomes

$$Z_n = \frac{1}{c} + Z = \frac{1}{c} + \tan^2 \varphi \tag{18.24}$$

where the angle φ is the angle between the normal to the surface and the optical axis, as illustrated in this figure, and f is the focal length of the paraboloid. We may see that for this case of the paraboloid, the distance Z_n from the center of curvature to the intersection of the normal with the optical axis is equal to the sagitta Z as shown in Figure 18.5. In the general case of aspheric surfaces, the intersection of the normals may be approximated by

$$Z_n = \frac{1}{c} - \frac{(Kc^3 + 8A_1)S^2}{2c^2} \tag{18.25}$$

18.2. CAUSTIC PRODUCED BY AN ASPHERIC SURFACE

When testing an aspheric surface at its center of curvature, it is useful at times to know the pertinent dimensions for the caustic, which may be derived by the method to be explained. The envelope of the caustic produced by the light reflected from the aspheric surface when illuminated with a point source at the center of curvature is

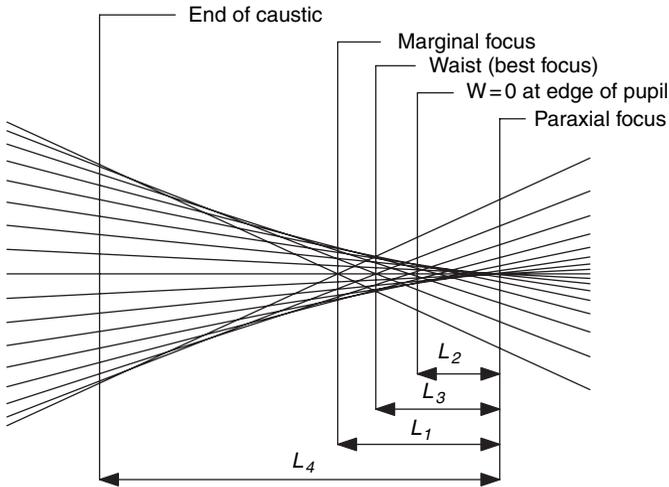


FIGURE 18.6. Caustic produced by a reflective concave aspheric surface tested at the center of curvature.

called the *evolute* in analytic geometry. The wavefront $W(S)$ reflected from the aspheric surface can be written as (see Fig. 18.6)

$$W(S) = \frac{Kc^3}{4} S^4 + \frac{(\Delta L)c^2}{2} S^2 \quad (18.26)$$

where ΔL is the distance of the plane under consideration from the paraxial focus. The distance ΔL from the paraxial to the marginal focus can be found by means of the condition

$$\left(\frac{dW}{dS} \right)_{S=S_{\max}} = 0 \quad (18.27)$$

where S_{\max} is the semidiameter of the surface under test. Thus we obtain

$$(\Delta L)_{\text{marginal focus}} = L = -KcS_{\max}^2 \quad (18.28)$$

The distance ΔL to the end of the caustic from the paraxial focus is found with the condition

$$\left(\frac{d^2W}{dS^2} \right)_{S=S_{\max}} = 0 \quad (18.29)$$

thus obtaining

$$(\Delta L)_{\text{end of caustic}} = 3L \quad (18.30)$$

The distance ΔL from the waist of the caustic to the paraxial focus is obtained with the condition

$$\left(\frac{dW}{dS}\right)_{S=S_{\max}} = -\left(\frac{dW}{dS}\right)_{S=\sigma} \quad (18.31)$$

where σ is the value of S that gives a maximum or minimum for dW/dS or, equivalently, $(d^2W/dS^2)_S = \sigma = 0$. Then it is possible to obtain

$$(\Delta L)_{\text{waist of caustic}} = \frac{3}{4}L \quad (18.32)$$

The diameter w of the waist, or circle of least confusion, is given by

$$w = \frac{2}{c} \left(\frac{dW}{dS}\right)_{S=S_{\max}} \quad (18.33)$$

using the value of ΔL for the waist of the caustic. The result is

$$w = -\frac{1}{2}Kc^2S_{\max}^3 \quad (18.34)$$

The diameter at the beginning of the caustic (paraxial focus) is equal to $4w$, and at the end of the caustic, it is equal to $8w$.

18.3. PRIMARY ABERRATIONS OF SPHERICAL SURFACES

The usual expressions for the primary aberrations of spherical refracting and reflecting optical surfaces will be described in the next sections.

18.3.1. Spherical Aberration of and Aspherical Surface

A system of k -centered reflective or refractive aspherical surfaces has a transverse third-order spherical aberration given by

$$Sph^T = \sum_{j=1}^k Sph_j \quad (18.35)$$

with

$$Sph_j = -\frac{y(N_{-1} - N)}{2N_k u_k} \left\{ (8A_1 + Kc^3)y^3 + \frac{n-1}{n^2} (n + n_{-1})u_{-1} + n_{-1}yc \right\} (yc + u_1)^2 \quad (18.36)$$

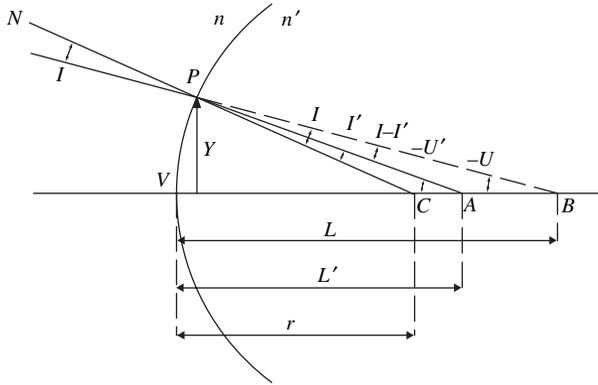


FIGURE 18.7. Sign convention (all shown quantities are positive).

where u_{-1} and u are the slopes of the incident and refracted or reflected rays, u_{-1} and n are the refractive indices, and y is the height of the ray on the surface.

The sign convention is in Figure 18.7. For a reflective surface we may write $u_{-1} = -n = 1$, hence the transverse spherical aberration becomes

$$Sph_j = \frac{y}{N_k u_k} [(8A_1 + Kc^3)y^3 + (yc + u_{-1})yc] \tag{18.37}$$

Considering now the particular case of a mirror tested at its center of curvature ($n_k = -1, u_{-1} = 0$, and $u = -yc$), the transverse spherical aberration is given by

$$Sph = \frac{-(8A_1 + Kc^3)y^3}{c} \tag{18.38}$$

while with a point source at infinity and the image at the focus ($u = -2yc$) we have

$$Sph_f = \frac{-(8A_1 + Kc^3 + c^3)y^3}{2c} \tag{18.39}$$

18.3.2. Coma of a Concave Mirror

The expression for the third-order sagittal coma with the pupil at the optical surface is independent of the asphericity of the mirror and is given by

$$Coma_s = (yc + u)cyh \tag{18.40}$$

where h is the image height (or deviation from the optical axis). It is quite interesting that if a surface is tested near the center of curvature ($u = -yc$), no third-order coma is introduced if the light source is placed slightly off axis.

18.3.3. Astigmatism of a Concave Mirror

The Petzval surface of a concave reflecting optical surface depends only on the curvature of the surface and has a curvature $1/\rho_p$ given by

$$\frac{1}{\rho_p} = -2c \quad (18.41)$$

We restrict ourselves to the case of an optical surface and its pupil at the same place since this is the most interesting configuration in optical testing.

Then we can show that the sagittal surface is always flat and that the tangential surface has a curvature $1/\rho_T$ given by

$$\frac{1}{\rho_T} = -4c \quad (18.42)$$

The expression for the third-order transverse sagittal astigmatism (pupil at the optical surface) as measured on the Petzval surface is

$$Ast_T = cuh^2 \quad (18.43)$$

The tangential astigmatism on the Petzval surface has three times the magnitude of the sagittal astigmatism. The difference between these two aberrations is a residual transverse aberration, given by

$$TA_{ast} = Ast_t - Ast_s = 2Ast_s \quad (18.44)$$

Thus, when testing with the light source slightly off axis, the apparent astigmatic difference between the tangential and the sagittal wavefront profiles is

$$W = -\frac{2}{l} \int_0^y Ast_s dy \quad (18.45)$$

where l is the distance from the surface to the image ($y = ul$). It may be shown that

$$W = -cu^2h^2 = -\frac{ch^2y^2}{l^2} \quad (18.46)$$

Therefore an apparent astigmatism equal to $W/2$ will be found on the surface under test. The tangential curvature appears stronger than the sagittal curvature. Notice that, if the surface is tested at the center of curvature, l is the radius of curvature and h is half the separation between the point source and the image.

18.4. ASTIGMATIC SURFACES

An astigmatic surface is one that has two different curvatures in two perpendicular directions, called normal curvatures. These surfaces do not have rotational symmetry, but have symmetry at about two orthogonal planes that intersect at the optical axis. An

astigmatic surface may be generated in several manners, for example, as described by Malacara and Malacara, (1971). Sasian (1997) has shown that an astigmatic surface can sometimes be replaced by an off-axis paraboloid, which is relatively easier to manufacture. The curvature at the vertex in any direction θ is related to the two normal curvatures by the Euler formula as follows:

$$c_\theta = c_x \cos^2 \theta + c_y \sin^2 \theta \quad (18.47)$$

There is an infinite number of astigmatic surfaces that have these properties (Malacara-Doblado *et al.*, 1996), but the main ones used in optics will now be briefly described.

18.4.1. Toroidal Surface

A toroidal surface frequently has the shape of a doughnut, but the central hole may also be missing. The expression for a toroidal surface, with its axis of symmetry parallel to the y axis, is shown in Figure 18.8 is given by

$$z = r_x - \left[(\sqrt{r_y^2 - y^2} + r_x - r_y)^2 - x^2 \right]^{\frac{1}{2}} \\ = \frac{c_x x^2 + c_y y^2 + 2 \left(\frac{(c_y - c_x) y^2}{1 + \sqrt{1 - c_y^2 y^2}} \right)}{1 + \left\{ 1 - c_x \left[c_x x^2 + c_x y^2 + 2 \left(\frac{(c_y - c_x) y^2}{1 + \sqrt{1 - c_y^2 y^2}} \right) \right] \right\}^{1/2}} \quad (18.48)$$

Here, r_x is the radius of curvature on the x - z plane (semidiameter of the doughnut), and r_y is the radius of curvature in the y - z plane. The cross sections in the y - z plane as well as in all other planes containing the axis of symmetry are circles with curvature c_y . The cross section in the plane x - z is a circle with curvature c_x . The cross sections in all other planes containing the optical axis have complicated shapes that are not circles neither ellipses. If the toroid has a hole like a doughnut ($r_y < r_x/2$), it is called a *ring toroid*, if the hole is not present ($r_y > r_x/2$), it is called a *spindle toroid*.

This expression is not invariant if we rotate the toroid 90° about the optical axis. However, the expression for a toroid with its axis of symmetry along the x axis can be found from Eq. (18.49) by just interchanging x and y , as well as c_x and c_y . Depending on the orientation of the toroid as well as on the relation between c_x and c_y , the toroid becomes either a ring or a spindle toroid, as illustrated in Table. 18.2.

18.4.2. Astigmatic Ellipsoidal and Oblate Spheroidal Surfaces

An astigmatic ellipsoidal surface is generated by rotating an ellipse about an axis that we call an axis of symmetry. This axis is perpendicular to the optical axis and passes

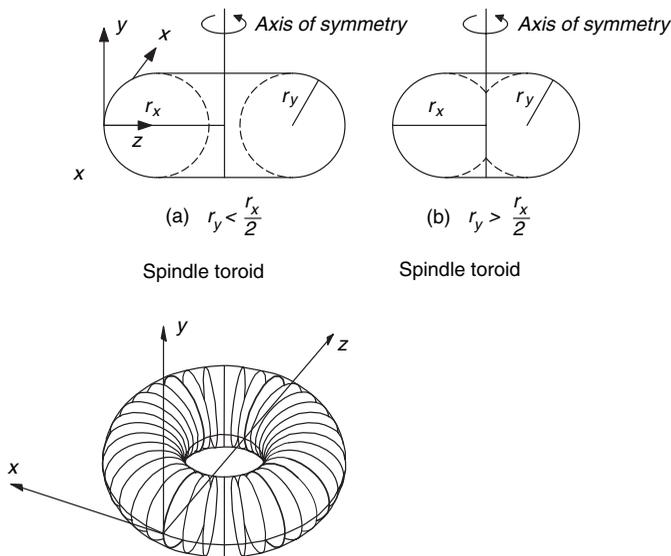


FIGURE 18.8. Astigmatic toroidal surface.

TABLE 18.2. Cross sections for a toroidal surface for two orthogonal orientations of the axis of symmetry, and the relation between the two normal curvatures.

Axis of symmetry parallel to:	Curvatures	Surface type	Cross-sections	
			Plane $x-z$	Plane $y-z$
y axis	$c_y < 2c_x$	Spindle toroid		
	$c_y > 2c_x$	Ring toroid		
x axis	$c_x < 2c_y$	Spindle toroid		
	$c_x > 2c_y$	Ring toroid		

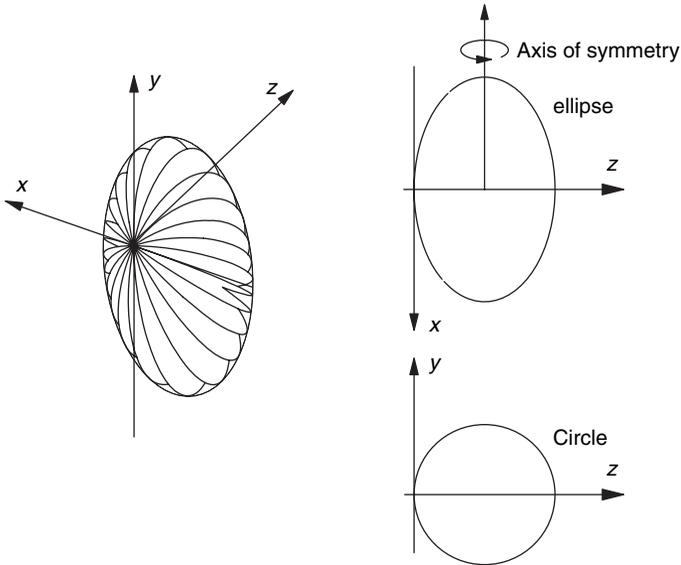


FIGURE 18.9. Astigmatic ellipsoidal surface.

through it. When the axis of symmetry is parallel to the y axis (see Fig. 18.9), the expression for the sagitta z is

$$z = \frac{1 - [1 - (c_x c_y x^2 + c_y^2 y^2)]^{1/2}}{c_y} = \frac{(c_x x^2 + c_y y^2)}{1 + [1 - c_y(c_x x^2 + c_y y^2)]^{1/2}} \tag{18.49}$$

If the major semi-axis of the generating ellipse is along the y axis, the generated surface is an ellipsoid. On the contrary, if the major semi-axis is along the z axis, the surface is an oblate spheroid. As in the case of the toroid, this expression is not invariant to rotations about the optical axis. The expression for z when the axis of symmetry is along the x axis can be found from Eq. (18.49) by interchanging x and y , as well as c_x and c_y .

The two cross sections on the planes x - z and y - z will depend on the orientation for the axis of symmetry as well as in the relative values of curvatures c_x and c_y , as illustrated in Table 18.3.

18.4.3. Sphero-Cylindrical Surface

A sphero-cylindrical surface, illustrated in Figure 18.10 may be represented by the expression (Menchaca and Malacara, 1986)

$$z = \frac{c_x x^2 + c_y y^2}{1 + \left[1 - \frac{(c_x x^2 + c_y y^2)^2}{x^2 + y^2} \right]^{1/2}} \tag{18.50}$$

TABLE 18.3. Cross sections for an ellipsoidal or oblate spheroidal surface for two different orientations of the axis of symmetry, and the relation between the two normal curvatures.

Axis of symmetry parallel to:	Curvatures	Semi-axis along optical axis (Z)	Surface type	Cross-sections	
				Plane x-z	Plane y-z
y axis	$c_y < c_x$	Minor semi-axis	Ellipsoid		
	$c_y > c_x$	Major semi-axis	Oblate spheroid		
x axis	$c_x < c_y$	Minor semi-axis	Ellipsoid		
	$c_x > c_y$	Major semi-axis	Oblate spheroid		

If the curvatures c_x and c_y , are equal to c , this expression becomes identical to Eq. (18.1). This expression is symmetrical with respect to x and also with respect to y . As a result, if x and y as well as c_x and c_y are interchanged, the form of this expression is preserved. Another interesting property of this surface is that the cross section in a plane containing the optical axis is always a circle, with a curvature c_θ given by the Euler formula.

The sphero-cylindrical and the toroidal surfaces have in common that their cross sections in the x - z and y - z planes are circles.

An important difference between the sphero-cylindrical and the toroidal surface is that the first has only two possible solutions for z since there is only one square root. For the toroidal surface we have two square roots, one inside the other. For a spindle toroid two of the four solutions may be imaginary.

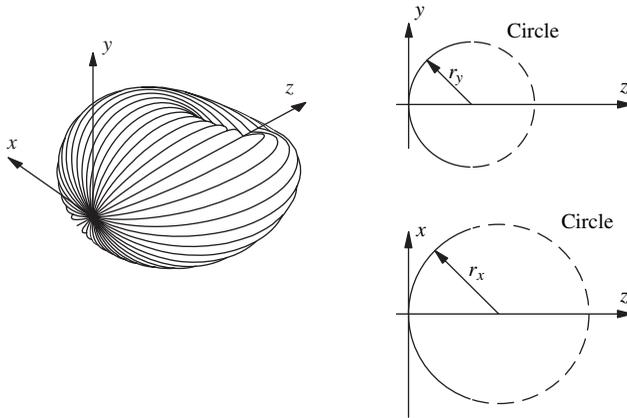


FIGURE 18.10. Astigmatic sphero-cylindrical surface.

18.4.4. Testing Astigmatic Surfaces and Reference Astigmatic Surface

If we test the preceding astigmatic surfaces with a point light source at one of the two normal centers of curvature, we would expect a perfect image with the shape of a thin line formed at the same plane by the reflected light, as described in Chapter 12. This is the basis of some tests used for astigmatic surfaces. This line image is not always a perfectly thin image from the geometrical point of view since some transverse aberration in the perpendicular direction to the line image may be present in the reflected wavefront. However, it is possible to have a perfect line image. The requirement is that the surface under test has an axis of rotational symmetry perpendicular to the optical axis, as in the cases of the toroidal, ellipsoidal and oblate spheroidal surfaces, and that the illuminating point source is located at the point where the axis of symmetry and the optical axis intersect. This condition guarantees that the reflected wavefront has no transverse aberration TA_x in the x direction. Then, the image is a perfectly thin and straight line oriented along the axis of symmetry of the optical surface.

The only problem with the surfaces previously described with an axis of symmetry is that the reflected light rays in the plane containing this axis of symmetry may have some transverse spherical aberration TA_y , which, of course, does not broaden the thin line image; only the light distribution along this line is not evenly distributed. If it is necessary to eliminate this remaining aberration, a perfect astigmatic reflecting surface is produced if the surface generating ellipse is rotated not about its semi-axis but about a line passing through one of the focii of the ellipse and parallel to the y axis. The astigmatic surface is illustrated in Figure 18.11. The cross section in the plane y - z is a section of an ellipse. As in the toroidal, ellipsoidal and oblate spheroidal surfaces, the cross sections in any plane parallel to the plane x - z are circles whose centers of curvature are all aligned on the axis of symmetry in this case, passing through one of the ellipse focii. This surface could be considered as an ellipsoidal spindle toroidal surface. For simplicity, we will call it a *reference astigmatic surface*.

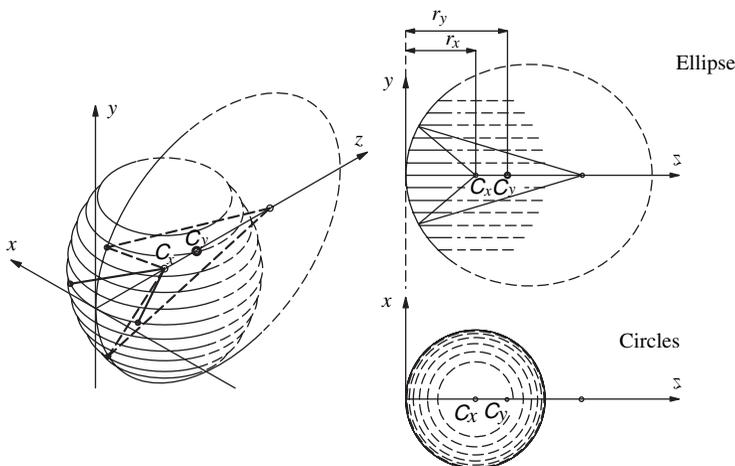


FIGURE 18.11. Perfect astigmatic reflecting surface.

The general expression for a surface generated by rotating an ellipse about any axis of symmetry parallel to the y axis and passing through the optical axis can be found to be

$$z = r_x - \left\{ \left[\left(a^2 - \frac{ay^2}{r_y} \right)^{1/2} + r_x - a \right]^2 - x^2 \right\}^{1/2}$$

$$= \frac{c_x x^2 + c_y y^2 + \left(\frac{1}{a} - c_x \right) \left(\frac{c_y y^2}{1 + \left(1 - \frac{c_y y^2}{a} \right)^{1/2}} \right)^2}{1 + \left\{ 1 - c_x \left[c_x x^2 + c_y y^2 + \left(\frac{1}{a} - c_x \right) \left(\frac{c_y y^2}{1 + \left(1 - \frac{c_y y^2}{a} \right)^{1/2}} \right)^2 \right] \right\}^{1/2}} \quad (18.51)$$

where a is the major semi-axis of the generating ellipse. As it can be easily noticed, the axis of symmetry is at a distance c_x from the vertex.

It is interesting to notice that from this expression the reference astigmatic surface as well as several other surfaces can be obtained by setting the value of a as in Table 18.4.

Once more, as in the cases of the toroid and the ellipsoid, this surface can be rotated to have its axis of symmetry parallel to the x axis by interchanging x and y , as well as c_x and c_y .

18.4.5. Comparison Between Astigmatic Surfaces

All the astigmatic surfaces described here have in common the property that the cross section in the plane x - z are circles (if the axis of symmetry is parallel to the y

TABLE 18.4. Values of a in Eq. (18.51) to produce astigmatic surfaces with an axis of symmetry parallel to y axis.

Astigmatic surface	Value of a
Toroid	$1/c_y$
Ellipsoid or oblate spheroid	$1/c_x$
Reference astigmatic surface	$\frac{c_y}{c_x(2c_x - c_y)}$

axis). The Taylor expansion up to the fourth power of the astigmatic surfaces have in common the first three terms as follows:

$$z = \frac{1}{2}c_x x^2 + \frac{1}{2}c_y y^2 + \frac{1}{8}c_x^3 x^4 + z_4 \quad (18.52)$$

where z_4 represents the next two fourth order terms, and the coefficient of $x^2 y^2$ is a square ashtray term:

(a) For the sphero-cylindrical surface:

$$z_4 = \frac{1}{16}(3(c_x + c_y)c_x c_y - (c_x^3 + c_y^3))x^2 y^2 + \frac{1}{8}c_y^3 y^4 \quad (18.53)$$

(b) For the toroidal surface with axis of symmetry parallel to the y axis:

$$z_4 = \frac{1}{4}c_x^2 c_y x^2 y^2 + \frac{1}{8}c_y^3 y^4 \quad (18.54)$$

(c) For the ellipsoidal surface with axis of symmetry parallel to the y axis, passing through one of the semi-axes:

$$z_4 = \frac{1}{4}c_x^2 c_y x^2 y^2 + \frac{1}{8}c_x c_y^2 y^4 \quad (18.55)$$

(d) For the reference surface with axis of symmetry parallel to the y axis, passing through one of the focii, by approximating Eq. (18.51) with a Taylor expansion up to the fourth power and using the corresponding value of a in Table 18.4:

$$z_4 = \frac{1}{4}c_x^2 c_y x^2 y^2 + \frac{1}{8}c_y c_x (2c_x - c_y) y^4 \quad (18.56)$$

It must be pointed out that despite these formal differences for the astigmatic surfaces, in practice they might be almost identical if the two normal curvatures are similar.

18.5. OFF-AXIS CONICOIDS

We may think of a conicoid in terms of an ellipsoid, where the sphere, the paraboloid and the hyperboloid are particular cases where the only difference is the excentricity. The surface of any off-axis conicoid, as illustrated in Figure 18.12, can be represented by

$$z(x, y) = \frac{1}{2}c_x x^2 + \frac{1}{2}c_y y^2 + a_3 x^2 y + a_4 y^3 + a_5 x^4 + a_6 x^2 y^2 + a_7 y^4 \tag{18.57}$$

where c_x and c_y are the curvatures at the origin of coordinates in the off-axis conicoid and the coefficients a_3, a_4, a_5 and a_7 depend on the parameters defining the conicoid.

In general, for all off-axis conicoids we may see that

$$c_y c^2 = c_x^2 \tag{18.58}$$

as shown by Menchaca and Malacara (1984). Frequently, the curvature in the y direction is called the tangential curvature, and the curvature in the x direction is called the sagittal curvature.

If we define the ellipsoid parameters W and z_0 in terms of the off-axis angle θ and the semiaxes a and b by

$$W^2 = (a^2 - b^2) \sin \theta \cos \theta \tag{18.59}$$

and

$$z_0 = \sqrt{a^2 \cos^2 \theta + b^2 \sin^2 \theta} \tag{18.60}$$

as illustrated in Figure 18.12, we obtain the coefficients as

$$c_x = \frac{z_0}{b^2} \tag{18.61}$$

$$c_y = \frac{z_0^3}{a^2 b^2} \tag{18.62}$$

$$a_3 = -\frac{z_0^2 W^2}{2a^2 b^4} \tag{18.63}$$

$$a_4 = \frac{z_0^4 W^2}{2a^4 b^4} \tag{18.64}$$

$$a_5 = \frac{z_0(a^2 b^2 + W^4)}{8a^2 b^6} \tag{18.65}$$

$$a_6 = \frac{z_0^3}{4a^4 b^6} (a^2 b^2 + 3W^4) \tag{18.66}$$

$$a_7 = \frac{z_0^5}{8a^2 b^2 (a^2 b^2 + W^4)} \left(\frac{5W^8}{a^4 b^4} + \frac{6W^4}{a^2 b^2} + 1 \right) \tag{18.67}$$

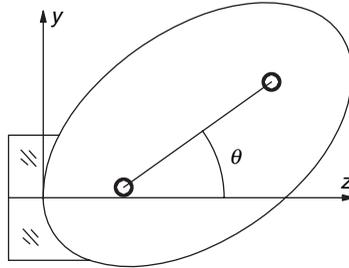


FIGURE 18.12. Off-axis ellipsoid.

These coefficients indicate the presence of coma, triangular astigmatism, as well as quadratic astigmatism (also known as ashtray astigmatism). These expressions can also be written in terms of the axial curvature c of the ellipsoid, that is, the curvature at the end of the semi-axis a , the conic constant K , and the off-axis angle θ by using

$$a = \frac{1}{c(K+1)} \quad (18.68)$$

$$b = \frac{1}{c\sqrt{K+1}} \quad (18.69)$$

Then, it easy to see that the sagittal and tangential curvatures become

$$c_x = c(1 + K \sin^2 \theta)^{3/2} \quad (18.70)$$

and

$$c_y = c(1 + K \sin^2 \theta)^{1/2} \quad (18.71)$$

respectively.

18.5.1. Off-Axis Paraboloids

As pointed out by Sasian (1997) an astigmatic surface, like a toroid, can sometimes replace an off-axis paraboloid like that in Figure 18.13, whose axis is tilted by an angle θ , and may be considered in a first approximation as a toroidal surface if the diameter is small compared with its radius of curvature.

From Eq. (18.71) by setting $K = -1$, we may see that the tangential curvature c_y is given by

$$c_y = c \cos^3 \theta \quad (18.72)$$

In an analogous manner, the sagittal curvature c_x is given by

$$c_x = c \cos \theta \quad (18.73)$$

The on-axis vertex curvature c of the paraboloid is related to the focal length by

$$c = \frac{1}{2F} \quad (18.74)$$

where F is the paraboloid focal length.

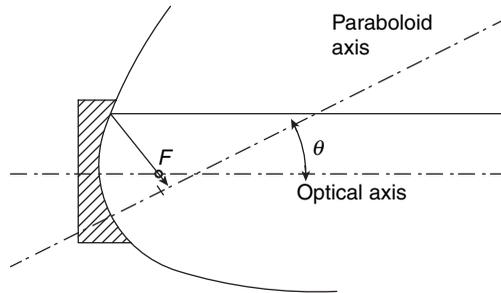


FIGURE 18.13. Off-axis paraboloid.

If the diameter of the off-axis paraboloid is not very small relative to its focal length, besides the astigmatic component, an additional coma like term has to be considered. With larger apertures, even a triangular astigmatism or quadratic terms appear (Malacara, 1990).

REFERENCES

- Malacara D., "Some Parameters and Characteristics of an Off-Axis Paraboloid," *Opt. Eng.*, **30**, 1277 (1990).
- Malacara D., and Z. Malacara, "Diamond Tool Generation of Toroidal Surfaces," *Appl. Opt.*, **10**, 975–977 (1971).
- Malacara-Doblado D., D. Malacara-Hernández, and J. García-Márquez, "Axially Astigmatic Surfaces: Different Types and Their Properties," *Opt. Eng.*, **35**, 3422–3426 (1996).
- McLeod J. H., "The Axicon: A New Type of Optical Element," *J. Opt. Soc. Am.*, **44**, 592–597 (1954).
- McLeod J. H., "Axicons and Their Uses," *J. Opt. Soc. Am.*, **50**, 166–169 (1960).
- Menchaca C. and D. Malacara, "Directional Curvature in a Conic Mirror," *Appl. Opt.*, **23**, 3258 (1984).
- Menchaca C. and D. Malacara, "Toroidal and Sphero-Cylindrical Surfaces," *Appl. Opt.*, **25**, 3008–3009 (1986).
- Mertz L., "Aspheric Potpourri," *Appl. Opt.*, **20**, 1127–1131 (1979b).
- Sasian J. M., "Double Curvature Surfaces in Mirror System Design," *Opt. Eng.* **36**, 183–188 (1997).
- Schulz G., "Aspheric Surfaces," in: E. Wolf (Ed.) *Progress in Optics*, vol. XXV, Chap. IV, North Holland, Amsterdam, 1988.
- Shannon R. R., "Aspheric Surfaces," in: R. Shannon and J. C. Wyant (Eds.), *Applied Optics and Optical Engineering*, vol. VIII, Academic Press, San Diego, 1980.

Appendix

Optical Testing Programs

Two computer programs for Windows, coded with National Instruments LabView, useful for optical testing are included with this book. These programs are:

1. A program to display and save data and images for several types of fringe patterns and images obtained by simulating some of the optical tests described in this book.
2. A program to analyze and design algorithms for measuring the phase in phase shifting interferometry.

A1. INTERFEROGRAM GENERATION

This program displays the test patterns on the computer screen and can also save them in the hard disc for from the following optical tests:

1. Wavefront plot (isometric).
2. Wavefront profiles (tangential and saggital).
3. Transverse aberrations ($TA_x(0, y)$, $TA_x(x, 0)$, $TA_y(0, y)$, $TA_y(x, 0)$).
4. Ronchi pattern.
5. Hartmann pattern.
6. Foucault pattern.
7. Twyman interferogram.
8. Lateral shear interferogram.
9. Radial shear interferogram.
10. Rotational shear interferogram.

The interferogram images have a resolution of 512×512 pixels, in the format *.bmp. The wavefront data can be specified in three different manners:

- (a) A polynomial represented by a sum of monomials.
- (b) A polynomial represented by a sum of primary aberration terms plus third and fifth order spherical aberration terms.
- (c) A polynomial represented by a linear combination of Zernike polynomials.

The coefficients in these three cases are those for a normalized pupil with a unit radius. The advantage of doing so is that the maximum deformation on the edge of a circular pupil is directly proportional to the value of the coefficient with its constant of proportionality being equal for all different terms, independently of its power.

Instead of the wavefront shape, we can define a reflecting surface shape, being tested at the vertex center of curvature. These are two types of surfaces.

- (a) An aspheric surface with rotational symmetry.
- (b) An a stigmatic surface, atoroid, an sphero-cylindrical surface, an ellipsoid or an astigmatic reference surface.

Once a type of data has been specified, the data being used is not modified until new data is entered. If desired, it can be seen in a screen called View Data.

If the data for a wavefront or surface is saved, it is stored in the folder:

C:\Program Files\Interferogram Generation\Interferograms Data

A2. PHASE SHIFTING ALGORITHMS

The data for this program is a set of coefficients W_1 in the numerator and W_2 in the denominator for the expression to find the phase in phase shifting interferometry, which is

$$\tan \phi = -\frac{\sum_{n=1}^N W_{1i} s(\alpha_n)}{\sum_{n=1}^N W_{2i} s(\alpha_n)}$$

where $s(\alpha_n)$ is the measured signal at the relative phase, measured with respect to the initial phase on the interferometer setting. The phase of the N sampling points is assumed to be equally spaced and the phase of the first sampled point is not necessarily zero.

The basic properties of the evaluated algorithm are calculated with the theory by Freischlad and Koliopoulos (1990). The exact phase error is calculated as described by Hernández-López and Malacara (2000) with the formula:

$$\tan \delta\phi = \frac{H_{01} + H_{11} \cos 2\phi + H_{21} \sin 2\phi}{H_{01} + H_{12} \cos 2\phi + H_{22} \sin 2\phi}$$

and the coefficients H_{ij} are calculated as described by the authors.

The first step when using this program should be either to enter new data or to open one of the data sets for an algorithm already stored in the folder:

C:\Program Files\Phase Shifting Algorithms\Algorithms data

When data for a new algorithm is saved, it is stored in the same folder.

A3. INSTALLATION

The installation is simple by just executing the program setup.exe. When installing these two programs please do not change the default installing directories, which are:

C:\Program Files\Interferogram Generation

for the program for the generation of interferograms and

C:\Program Files\Phase Shifting Algorithms

for the program to design phase shifting algorithms.

A4. SOME SUGGESTIONS FOR THEIR USE

The use of these programs is quite simple by just following the screens and menus. However, some hints may be useful: When modifying a data number while running any of these programs, if you do not want to modify the first significant digit, with the mouse place the vertical cursor line, at the right of the digit you want to modify.

These programs are provided as they are now. They may still contain a few bugs. If you find any, please contact the authors, so that they can be fixed in next versions.

REFERENCES

- Hernández-López and D. Malacara, "Exact Linear Detuning Error in Phase Shifting Algorithms," *Opt. Commun.*, **180**, 9–14 (2000).
- Freischlad and Koliopoulos, "Fourier Description of Digital Phase Measuring Interferometry," *J. Opt. Soc. Am. A.*, **7**, 542–551 (1990).

Index

- Aberration function, 499
- Aberration, spherical of aspherical surfaces, 839
- Aberration function
 - secondary, 504
 - Schwarzschild, 504
- Aberrations, 83, 259, 348, 408, 410, 416
 - astigmatism of a concave mirror, 841
 - balanced, 506, 508
 - coma of a concave mirror, 840
 - detection with Foucault test, 283
 - moiré pattern of, 766
 - primary Zernike, 521
 - rotationally symmetric pupil, 499
 - Seidel, 513
 - separation, 259
 - lateral, 504
 - variance, 507, 508, 535
- Absolute testing
 - of flats, 40
 - of spherical surfaces, 72, 665
- Accuracy in double-pass interferometers, 264
- Adaptive compensator, 471
- Airy pattern, 400
 - isophots, 404, 406
 - normalized intensity, 400
- Aldis spherometer, 817
- Angle block, 809
- Angle measurement, 808
 - interferometric, 812
 - in prisms, 813
- Angle vise, 809
- Annular polynomials, 525
- Armitage and Lohmann interferometer, 209
- Aspheric lens testing, 839
- Aspheric surface, 15, 102, 327, 435
 - astigmatism of, 841
 - caustic of, 837, 838
 - coma of, 840
 - definition, 832
 - spherical aberration of, 839
 - testing, 444, 452
- Aspheric wavefront, 435
 - testing with sub-Nyquist interferometry, 631
 - testing with two wavelengths, 488, 635
 - testing with wavefront stitching, 491, 637
 - testing methods, 437, 629, 631, 635, 637
- Astigmatic surfaces, 841
 - comparison, 847
 - testing, 846
- Astigmatism
 - axial, 504
 - detection, 841
 - detection with Foucault test, 287
- Atomic force microscopes, 678, 683
- Autocollimation, 443
- Autocollimation test for parabolic mirrors, 363
- Autocollimator, 810, 821
- Axial astigmatism, 504
- Axial coma, 504
- Axicon surface, 834
- Babinet compensator, 163
- Beam splitter, 48, 49, 51, 58
 - diffraction grating, 79, 80
 - birefringent, 104
 - non-polarizing, 55
 - polarizing, 53
 - required accuracy, 51
 - Saunders, 114
 - Wollaston prism, 106

- Beam walk-off, 269
- Bevel gauge, 816
- Brewster's angle, 48
- Brown interferometer, 192
- Burch compensator, 456
- Burch interferometer, 98

- Cartesian configuration, 451
- Caustic, 300, 837, 838
- Circle polynomials, 505
- Circular grid test, 347
- Cliromatic confocal microscopy, 693
- Coddington equations, 311
- Coded light projection, 780
- Coherence, 123
 - in radial shear interferometer, 191
 - requirements, 9, 20, 22, 56, 262
- Coma aberration detection, 284, 840
- Coma, axial, 504
- Commercial interferometers, 25, 735
- Common path interferometer, 97, 118
- Compensator, 454
 - adaptive, 471
 - Burch, 456
 - Couder, 456
 - Dall, 458
 - holographic, 474
 - Offner, 462
 - Offner reflection, 462
 - refractive, 461
 - reflecting, 466
 - Ross, 456
 - Shafer, 464
- Computer generated holograms, 478, 483, 485
- Concave mirror
 - astigmatism of, 841
 - coma of, 840
- Confocal microscopy, 668, 689, 693
- Confocal cavity, 270, 821
- Conic constant, 833
- Conic surface
 - aberration of the normals, 836
 - definition, 832
 - parameters for, 835
- Contact profilometers, 668, 670
- Couder compensator, 456
- Couder screen, 293
- Cube corner prism, 28, 66
- Curvature: measurement, 421

- Curvatures, local average, 421
- Cyclic radial shear interferometer, 194
- Cyclic interferometer, 140
- Cylindrical surfaces: testing, 453

- Dall compensator, 458
- Descartes ellipsoid, 450
- Descartes hyperboloid, 450
- Detectors, 613, 614
- Differential interference contrast microscope, 736
- Diffraction based interferometer, 158
- Diffraction ruling, 64
- Digital holography, 791, 794
- Digital holographic interferometry, 791
- Directional shearing interferometer, 166
- Displacement measurement, 781
- Divided circle, 808
- Double-focus interferometer, 112
- Double-focus system, 107
- Double-pass interferometer, 259
 - Fizeau, 262, 263, 269
 - Twyman-Green, 264
 - radial shear interferometer, 195
- Dyson system, 441
- Dyson interferometer, 112

- Electro-optic holography, 795
- Ellipsoidal surface
 - cross section, 845
 - definition, 835, 842
 - testing, 444, 448
- Encircled Energy, 402
- Equal chromatic order fringes, 220, 232
- Equivalent wavelength, 759
- ESPI (Electronic speckle pattern interferometry), 791, 794

- Fabry–Perot interferometer, 219, 236
- Film and plate thickness measurement, 729, 735
- First order parameters for an interferometer, 186
- Fitting the wavefront, 498
- Fizeau interferometer, 1, 229, 553
- Fizeau interferometer, double-pass, 262, 269
- Flat surface
 - and Ritchey-Common test, 310
 - testing, 4, 40, 237, 442

- Flatness measurement, 4, 40, 237
- Flats, 4, 7, 10, 11
 - absolute testing of, 40
 - liquid, 23
- Focal length measurement, 823
 - fiber optics, 827
 - Fourier transforms, 827
 - microlenses, 827
 - moiré, 825
 - nodal slide bench, 823
 - Talbot autoimages, 826
- Focimeters, 824
- Foucault test, 275
 - geometrical theory, 280
 - physical theory, 289
 - practical configurations, 280
 - with Couder screen, 293
 - with zonal test, 293
- Fourier transforms, 422
- Fresnel zone plate interferometer, 102
- Fringes
 - equal chromatic order (FECO), 220, 232
 - projection, 756, 769
 - stabilization, 77
 - Tolansky, 241
- Gardner-Bennett test, 393
- Gates reversal shear interferometer, 214
- Gaussian beam, 399, 409
- Geneva gauge, 820
- Geometric phase shifting, 738
- Geometrical theory
 - of Platzek-Gaviola test, 298
 - of wire test, 293
- Glass plate, 63
- Goniometer, 810
- Grazing incidence interferometer, 79, 270, 648
- Grazing incidence multipass interferometer, 271
- Group refractive index, 61
- Haidinger interferometer, 1
- Hariharan and Sen Interferometer, 192
- Hartmann test, 350
 - data reduction, 380
 - helical pattern screen in, 382
 - implementation, 362
 - ophthalmic lenses, 379
 - pattern for hyperboloidal mirror, 368
 - set up, 362
 - theory, 362
 - transverse aberrations, 363
 - with four holes screen, 376
 - with radial pattern screen, 380
 - with rectangular screen, 366
- Hartmann-Shack test, 383
 - concave and convex surfaces test, 389
 - convergent lenses, 388
 - crossed cylindrical test, 386
- Heteroscopic imaging, 689
- Hindle test, 445
- Holograms, 474
- Holographic compensator, 475
- Holographic test, 783
- Holographic interferometry
 - digital, 791
 - nondestructive testing, 784
 - static, 784
 - phase measurement, 783
 - time average, 787
- Holographic radial shear interferometer, 201
- Holography: electrooptic, 795
- Holography: two angle, 778
- Homogeneity testing, 27
- Hyperboloidal surface
 - definition, 835
 - Foucault test for, 278, 294
 - Hartmann test for, 362
 - Platzek-Gaviola test for, 298
 - testing, 445, 447, 451
- Inhomogeneity testing, 27
- Integrating bucket, phase shifting, 561
- Interferometers, 83, 638
 - lateral shear, 168
 - Twyman-Green, 82, 553
 - Armitage and Lohmann, 209
 - Brown, 192
 - Burch, 98
 - commercial, 647
 - common path, 97
 - compensation, 50
 - configurations, 3, 82, 168, 650
 - coupled, 244, 245
 - cyclic, 140
 - distance measuring, 649
 - double-pass, 259

Interferometers (*Continued*)

- double-pass Fizeau, 263
- double-pass radial shear, 195
- double-pass Twyman-Green, 264
- diffraction based, 158
- double focus, 112
- Dyson, 112
- Fabry-Perot, 219
- Fizeau, 1, 17, 229, 553
- Fresnel zone plate, 102
- fringe stabilization, 77
- Gates reversal shear, 214
- grazing incidence, 79, 648
- Haidinger, 1, 33
- Hariharan and Sen, 192
- holographic radial shear, 201
- Jamin, 137, 215
- Koster reversal shear, 213
- lateral shear, 108, 122
- Linnik or Smartt, 116, 302
- Mach-Zehnder, 553
- Murty, 150, 152
- Murty and Hagerott, 209
- multiple beam Fizeau, 224, 259
- multiple image, 80, 81, 556
- multiple-pass, 259, 266
- Mach-Zehnder, 77, 142, 146, 167, 553
- Michelson, 139, 143
- Newton, 1, 3
- oblique incidence, 78
- open path, 77
- plane parallel, 150, 152
- point diffraction, 116
- polarization based, 162
- phase shifting, 102, 547, 550, 740
- polarization, 669, 735
- radial shear, 185, 187
- reversal shear, 185, 211
- rotational shear, 185, 204
- Saunders prism, 114
- Saunders reversal shear, 214
- scatter plates, 98
- Shack-Fizeau, 34
- series, 244, 246
- Steel radial shear, 198, 200
- Som radial shear, 199
- thick lens radial shear, 202
- triangular-path, 79, 140
- Twyman-Green, 46
- unequal-path, 73
- wavelength scanning, 724
- wavefront-reversing, 115
- Williams, 47
- Interferogram analysis, 638
 - zero crossing, 638
 - synchronous detection, 639
 - heterodyne interferometry, 640
 - phase lock interferometry, 641
 - spatial synchronous, 642
 - Fourier methods, 642
 - computer processing for, 644
- Interferometric optical profilometers, 668, 695
 - grazing incidence, 270
 - multiple wavelength, 667
 - multiple beam, 219, 221
 - multiple pass, 266
 - Murty radial shear, 195, 199
 - Static holographic, 785
 - speckle, 791
 - two wavelength, 668, 702
 - white light, 667, 669, 711, 731
- Irradiance transport equation, 425
- Isophots in Airy pattern, 404, 406
- Jamin interferometer, 137, 215
- Knife-edge test, 275
- Köster reversal shear interferometer, 213
- Laplacian, 421
- Lateral shearing interferometer, 108, 122
- Lateral shearing interferometer
 - directional, 166
 - fringe patterns, 168
 - vectorial, 164
- Lateral aberrations, 504
- Lens testing, 69, 276
- Light sources, 9, 11, 23, 36, 123
- Light sources, 80, 74, 149, 207, 413, 650, 724
 - for star test, 413
 - infrared laser, 80
 - laser, 74, 149
 - size in radial shear, 207
 - wavelength tunable, 724
- Linnik interferometer, 116
- Lower test, 349
- Lyot test, 305

- Mach-Zehnder interferometer, 77, 142, 146, 167, 553
- Meinel tests, 449
- Michelson interferometer, 139, 143
- Microdisplays, 773
- Microscope
- atomic force, 678, 686
 - confocal, 689
 - differential interference contrast, 736
 - objectives, 415
 - scanning probe, 674
 - scanning tunneling, 676
 - testing, 71, 439
- Moiré, 756
- analysis of interferogram, 482
 - patterns of interferograms, 762
 - patterns of aberrations, 766
 - projection, 756
- Multiple beam interferometer, 24, 221
- fringe interval, 235
 - holographic, 247
 - with curved surfaces, 243
- Multiple beam interferometer: Fizeau, 224, 268
- Multiple image interferometer, 80, 81, 566
- Multiple wavelength interferometry, 667, 678, 679
- Multiple-pass interferometers, 259, 266
- Murty and Hagerott interferometer, 209
- Murty interferometer, 150, 152
- Murty radial shear interferometer, 195, 199
- Newton rings, 2, 3, 5
- Newton fringes, 2, 3, 5, 10
- Newton interferometer, 1, 3
- Nodal slide lens bench, 416, 823
- Non-null tests, 420, 439
- Nondestructive testing, 784
- Null Ronchi rulings, 328
- Null test configurations, 442
- Null test compensator, 454
- Oblate spheroid: definition, 835
- Oblate spheroid: testing, 472
- Oblique incidence interferometer, 78, 82
- Off-axis conicoids, 849
- Off-axis paraboloids, 850
- Offner compensator, 462
- Offner reflection compensator, 468
- Opaque surfaces, 17
- Open path interferometer, 77
- Optical flats, 4, 7, 10, 12
- Optical focus sensor, 668, 687
- Optical materials, 12
- Optical profilometers, 668, 685
- Optical ranging, 669, 741
- Optical surfaces: plane, 4, 7, 10, 12
- Optical surfaces
- aspherical, 15, 102
 - spherical, 13, 14, 30, 32
- Orthonormal coefficients, 535
- Orthonormal polynomials, 537
- Paraboloidal surface
- definition, 832, 835
 - Foucault test for, 278, 294
 - Hartmann test for, 362
 - off-axis paraboloids, 850
 - Platzeck-Gaviola test for, 298
 - testing, 443, 444
- Parallel plates, 26, 28, 35
- Phase conjugating interferometer, 81, 82
- Phase grating Ronchi test, 342
- Phase modulation test, 302
- Phase shifting, 547, 550
- geometric, 738
 - interferometer, 102, 740
 - methods to produce, 552
 - point diffraction interferometer, 117
 - quality functions for, 617
 - Ronchi test, 348
 - spatial, 560
 - temporal, 560
 - two wavelength, 635
- Phase shifting algorithms, 557, 568, 582
- 2+1, 580
 - averaging, 576
 - Carré, 574
 - characteristic polynomial, 589
 - four steps, 558
 - Fourier description, 586
 - Hariharan, 577
 - least squares, 571
 - N+1 bucket, 583
 - N+3 bucket, 584
 - methods to evaluate, 586
 - summary of, 591
 - three steps, 569

- Phase shifting error, 599, 600
 - detector non linearities, 602
 - source stability, 605
 - quantization errors, 606
 - vibration errors, 607
 - air turbulence, 610
 - extraneous fringes, 610
 - calibration, 596
 - optical, 611
- Phase unwrapping, 623
 - in one dimension, 623
 - in two dimensions, 625
 - temporal, 629
- Phase value photogrammetry, 780
- Phasogrammetry, 780
- Physical theory of Foucault test, 289
- Physical theory of wire test, 293
- Plane parallel interferometer, 150, 152
- Platzcek-Gaviola test, 298
- Point diffraction interferometer, 116, 302
- Polarization interferometers, 735
- Polarization based interferometer, 162, 669
- Polarization phase shifter, 557
- Polynomials, orthonormal, 537
- Porro prism, 67
- Primary aberrations, 83, 128, 166, 320, 501
- Prism, 64, 39, 812
 - cube corner, 28
 - Saunders, 114
 - Porro, 67
 - right angle, 816
 - testing right angle, 816
 - Wollaston, 106
- Profilers, 667
- Profilers. stylus, 670, 683
- Profilometers
 - contact, 668, 670
 - interferometric, 668, 695
 - optical, 668, 685
 - stylus, 668
- Projection moiré, 777
- Protactors, 808
- Pyramidal error in prisms, 813

- Radial shear interferometer, 185, 187
 - cyclic, 194
 - laser, 197
- Radial pattern screen for Hartmann test, 366

- Radius of curvature measurement, 154, 817
- Rayleigh criterion, 61
- Reflecting compensators, 466
- Refractive compensator, 461
- Reversal shear interferometer, 185, 211
- Reversing wavefront interferometers, 115
- Right angle prism, 34
- Rimmer and Wyant method, 135
- Ritchey test, 447
- Ritchey-Common test, 310
- Ronchi test, 317
 - circular grid, 347
 - null rulings, 328
 - patterns, 320
 - phase grating, 342, 348
 - physical theory, 337, 343, 344
 - sideband, 348
 - Talbot effect in, 341
- Ronchi-Hartmann test, 350
- Ross compensator, 456
- Rotational laser interferometer, 185, 204

- Saunders reversal shear interferometer, 214
- Saunders method, 134
- Saunders prism interferometer, 114
- Savart polariscope, 104, 108
- Scanning probe microscope, 668, 674
- Scanning tunneling microscopes, 676
- Scatter plates interferometer, 98
- Schwarzschild aberration function, 504
- Secondary aberration function, 504
- Seidel aberrations, 501
- Seidel sums, 503
- Separation of aberrations, 259
- Shack-Fizeau interferometer, 34
- Shadow moiré, 773
- Shafer compensator, 464
- Shearography, 801
- Sideband Ronchi test, 348
- Signal analysis, 728
- Silvertooth test, 447
- Smartt interferometer, 302
- Som radial shear interferometer, 199
- Source size in rotational shear interferometers, 211
- Spatial coherence requirements, 56
- Spatial phase shifting, 564

- Speckle
 - correlation fringes, 793
 - interferometry, 791
 - test, 783
- Spherical aberration: and Foucault test, 283
- Spherical aberration, 411, 417
- Spherical surfaces, 13, 30, 32
 - absolute testing, 72, 651
 - definition, 832
 - Foucault test for, 276
 - Gardner-Bennett test for, 394
 - Michelson test for, 393
 - wire test for, 293
- Sphero-cylindrical surface, 844
- Spheroid, definition, 833
- Spheroid: concave surface testing, 833
- Spherometers, 817
 - Abbe:, 819
 - Aldis, 817
 - bar, 820
 - dial, 820
 - ring, 819
 - Steinheil, 819
 - three leg, 818
 - precision, 819
 - optical, 821
- Square array screen, in Hartmann test, 376
- Star test
 - aspherical aberrations with, 419
 - astigmatism with, 419
 - aberrations field, 406
 - coma with, 419
 - distortion measurement with, 420
 - Star test, 398
 - light source for, 413
 - optical arrangement, 413
 - visual, 410
- Static holographic interferometry, 785
- Steel radial shear interferometer, 198, 200
- Structured light projection, 780
- Stylus profilometers, 667, 668, 670, 683
- Surface microtopography, 232, 234
- Surfaces: aspherical, 327
- Synthetic wavelength, 759
- Synthetic holograms, 477
- Temporal coherence requirement, 60, 74
- Test plates, 14, 817
 - holographic, 476
- Test, Speckle, 783
- Test, Holographic, 783
- Testing
 - aspherical surfaces, 444, 452, 491, 631
 - autocollimation, 443
 - cylindrical surfaces, 453, 844
 - flat surfaces, 4, 19, 40, 442
 - hyperboloidal surfaces, 445, 447, 451
 - oblate spheroid, 472
 - spherical surfaces, 72, 276, 393, 293, 651
 - telescope objectives, 442
 - with computer generated hologram, 478, 483, 485
 - paraboloidal surfaces, 443, 444
- Thick lens radial shear interferometer, 202
- Thickness measurement: film and plate, 729, 735
- Thin-film thickness measurements, 236
- Time-average holographic interferometry, 787
- Tolansky inequality, 228, 245
- Tolansky fringes, 241
- Toroidal surfaces, 842, 843
- Transverse aberrations: and wavefront deformations, 363
- Traveling microscope, 821
- Triangular path interferometer, 79, 140
- TV Holography, 791, 794
- Two-angle holography, 778
- Two-wavelength interferometry, 668, 703
- Two-wavelength testing, 488
- Twyman-Green, 553
 - aberrations compensation, 72
 - coherence requirements, 46
 - interferometer, 46, 62
 - interferograms, 82
 - phase conjugating, 81
 - unequal-path, 73
 - lens testing, 69
 - microscope objectives, 71
- Unequal-path interferometer, 73, 74
- Vectorial lateral shearing interferometer, 164
- Vertex power, 824

- Wave aberration, 501
- Wavefront, 126, 83, 91
 - deformations and transverse aberrations, 363
 - determination with Fourier transforms, 422
 - fitting, 498
- Wavefront retrieval, 134
 - with Southwell algorithm, 373
 - with trapezoidal integration, 370
 - with polynomial fitting, 368
 - with radial Shear, 189
 - with curvature measurements, 421
 - with two defocused images, 426, 430
 - with Ronchi Test, 331, 333, 335
 - with a defocused image, 429
- Wavefront: tilt and defocus removal, 368
- Wavefront: imaging at the pupil, 441, 442
- Wavefront: stitching, 491
- Wavefront-reversing interferometer, 115
- Wavelength, synthetic, 759
- Wavelength, equivalent, 759
- Wavelength scanning interferometer, 724
- Wedge measurement, 26, 28, 35
- White light interferometry, 667, 669, 711, 731
- Williams interferometer, 47
- Wire test, 293, 275
 - geometrical theory, 297
 - physical theory, 299
- Wollaston prism, 106, 111
- Wolter test, 307
- Zernike coefficients, 511, 538
- Zernike test, 118, 302
- Zernike polynomials, 498
 - annular, 525
 - circle, 505, 508
 - orthonormal, 506
 - primary aberrations, 521
- Zonal screen, 293
- Zone plate, 764