

UNIVERSITY OF LONDON

IMPERIAL COLLEGE OF SCIENCE, TECHNOLOGY
AND MEDICINE

DEPARTMENT OF PHYSICS
APPLIED OPTICS GROUP

**Absolute length measurement using
multiple-wavelength phase-stepping
interferometry**

Andrew John Lewis MA

Thesis submitted for the degree of
Doctor of Philosophy of the University of London
and for the Diploma of Membership of Imperial College

November 1993

(2002 Re-print edition)

ABSTRACT

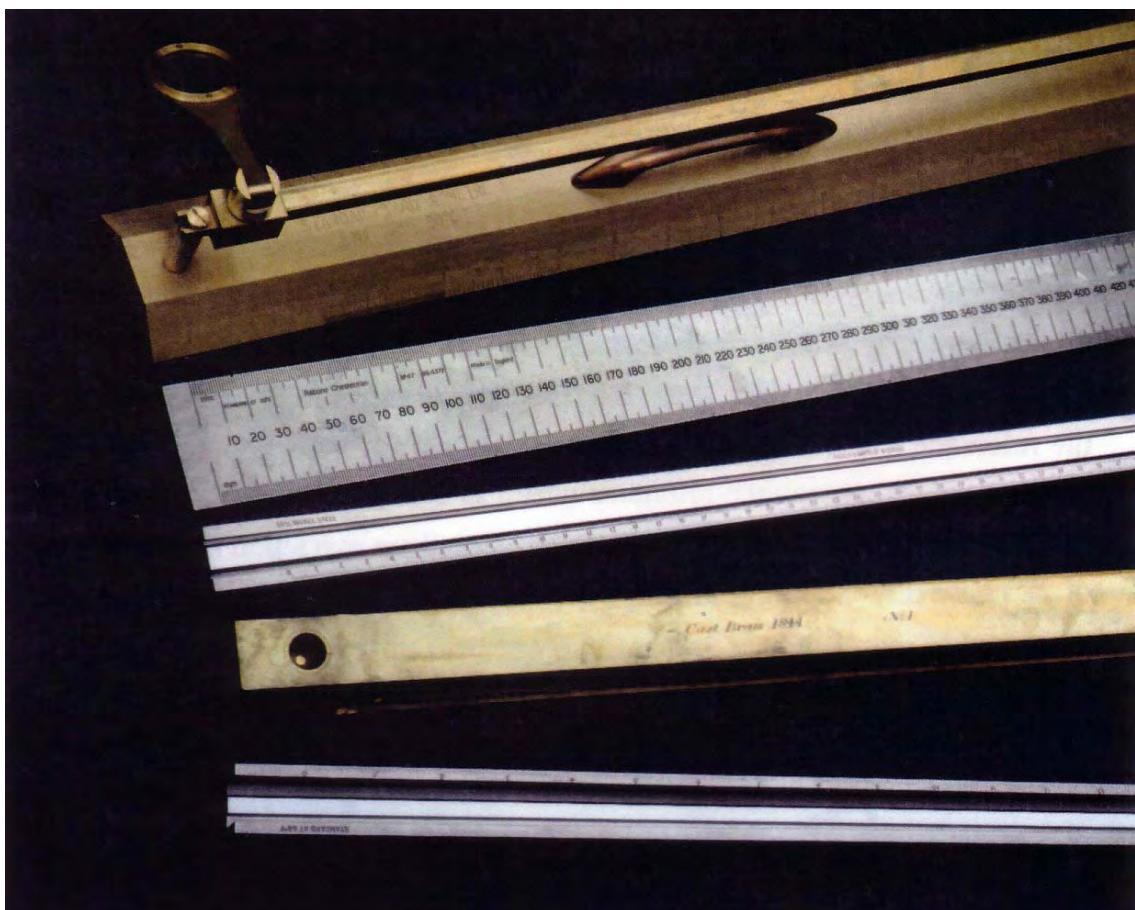
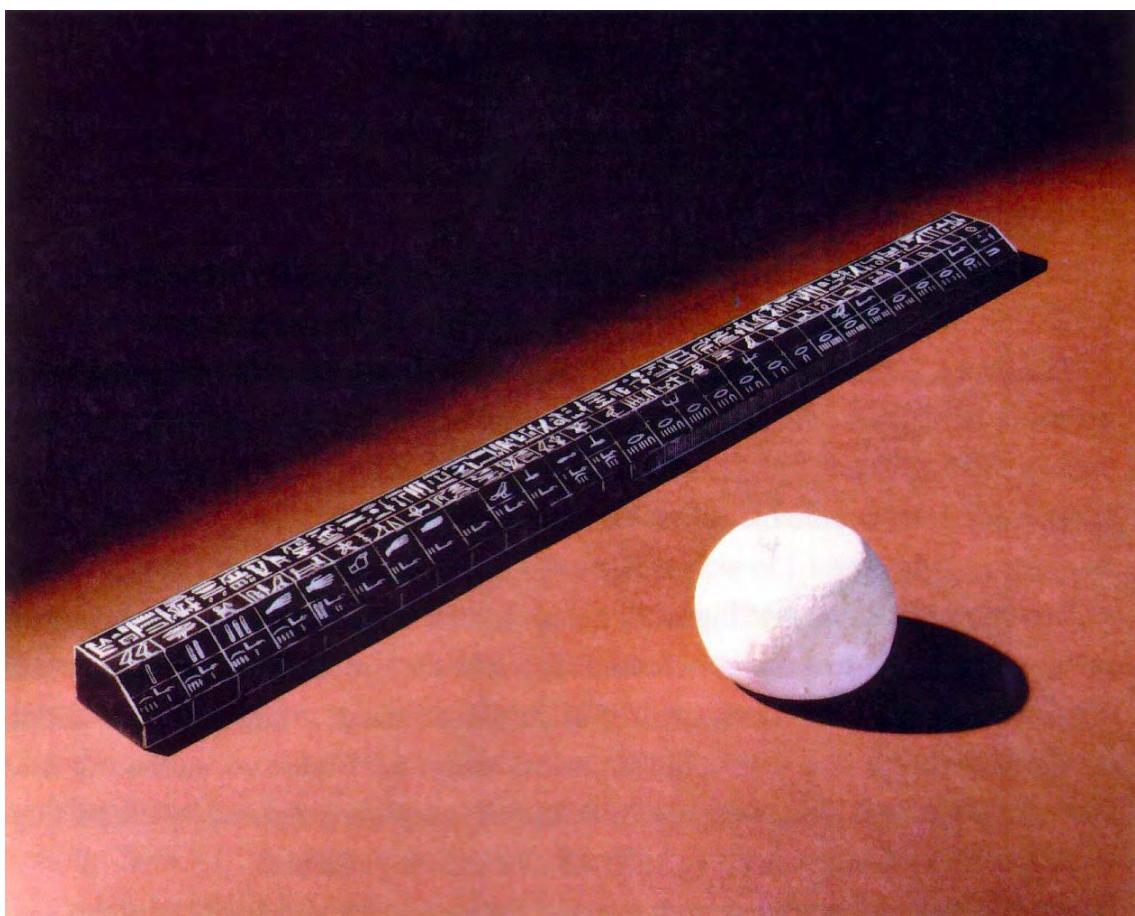
This thesis describes the design, construction, commissioning and testing of a new interferometer for the absolute measurement of length standards. The thesis starts with an introduction to the subject of length standards and length measurement by interferometry before considering various designs of interferometer. The design of the new interferometer is given in detail, including the operation of the lasers used as light sources. The alignment of the interferometer and the effects of incorrect alignment and collimation on the measured length are examined. A review of fringe analysis techniques is given, with an emphasis on phase-stepping algorithms. The 5-position algorithm used in the new interferometer is examined in more detail together with the phase-stepping actuator and its advantages over other devices. The data processing of images digitised in the interferometer is described, including the techniques developed for discontinuity removal and surface fitting. The measurement of the flatness and parallelism of the measuring faces of the length standards is described. The automated method of exact fractions is used to combine phase measurements at three wavelengths to allow accurate calculation of the length of the measured object, with a larger range than one or two-wavelength techniques. The techniques used for compensating for the refractive index of the air inside the interferometer chamber are summarised, with a comparison of calculated and directly-measured refractive index values, measured with a specially constructed refractometer. The thermal control of the interferometer is presented, including the use of the instrument to measure the thermal expansion coefficient of length standards, by measuring their lengths at different temperatures. Example results of length, flatness, parallelism, and thermal expansion are given for various sizes of length standard. A full uncertainty budget is calculated, allowing critical examination of the performance of the instrument. A chapter on conclusions is followed by several appendices.

PREFACE

This thesis describes the design, construction, commissioning and testing of a new multiple-wavelength phase-stepping interferometer for the measurement of length bars. This interferometer is called the NPL National Primary Length Bar Interferometer, or Primary Length Bar Interferometer (PLBI) for short. The instrument was designed and built at the National Physical Laboratory (NPL), Teddington. NPL was founded in 1900 and is the focus for the UK's National Measurement System.

The PLBI has been developed as a link between the realisation of the metre as a wavelength of a frequency-stabilised laser and the use of secondary length standards known as length bars, which are used as transfer standards to calibrate other standards or to verify the performance of instruments such as co-ordinate measuring machines (CMMs). The operation of the PLBI is described in a paper published by the author, which is reproduced in Appendix F (p 317). An overview is given here.

The PLBI is an interferometer of a modified Twyman-Green design, which measures the length of a bar in terms of the wavelength of the light emitted by calibrated frequency-stabilised lasers. The length, L , is calculated from the equation $L = (n + f)\lambda / 2$ where n is an integer (fringe order) and f is a fraction. In the interferometer, f is measured as a fringe fraction but n is initially unknown. The PLBI uses 3 wavelengths and measures a fringe fraction for each wavelength. The method of exact fractions is used to combine these 3 fractions with an estimate for L (within $\pm 9 \mu\text{m}$) to calculate L more accurately. In order to measure the fringe fractions with sufficient accuracy, Phase Stepping Interferometry is used to measure the fractions in terms of the phase difference between the reference and measurement beams. A 5-position phase-stepping algorithm is used which involves moving the reference mirror in 5 steps of size $\lambda/8$ for each wavelength and digitising the image at each step. Point-wise phase extraction is followed by removal of 2π phase discontinuities at fringe boundaries. After tilt removal, fringe fractions are measured as the difference in phase between pixels at the centre of the bar and those corresponding to the platen surface, in the 3 phase maps.



Egyptian Cubit (& mass standard) and more modern line standards of length

To my parents ...

CONTENTS

ABSTRACT	2
PREFACE	3
CONTENTS.....	7
LIST OF FIGURES.....	11
LIST OF TABLES	14
CHAPTER 1 INTRODUCTION	15
1.1 BACKGROUND.....	15
1.2 THE MEASUREMENT OF LENGTH	16
1.2.1 What is length?.....	16
1.2.2 The International System of units (SI)	16
1.2.3 The definitions of the SI units	17
1.3 HISTORICAL UNITS OF LENGTH	19
1.3.1 Timetable of events	19
1.3.2 The first definition of the metre	20
1.3.3 The 1927 definition of the metre.....	21
1.3.4 The 1960 definition of the metre	22
1.3.5 The present (1983) definition of the metre	24
1.3.6 Limitations of the present realisation of the metre	26
1.3.7 Future realisations of the metre.....	27
1.4 SECONDARY LENGTH STANDARDS.....	29
1.4.1 Modern secondary length standards	29
1.4.2 Gauge blocks and length bars.....	29
1.4.3 Gauge blocks.....	30
1.4.4 Length Bars	31
1.4.5 Definitions and specifications for Reference Grade length bars	32
1.4.6 Calibration of End Standards of length	33
1.4.7 The rationale behind the development of the new interferometer	34
1.4.8 Traceability of length bar length measurements	35
1.5 CONTENTS OF THE THESIS	38
CHAPTER 2 REVIEW OF LENGTH MEASURING INTERFEROMETERS	41
2.1 THEORY OF INTERFEROMETRIC LENGTH MEASUREMENT.....	41
2.2 BASIC INTERFEROMETER TYPES	42
2.3 REVIEW OF SMALL FIELD, DYNAMIC INTERFEROMETERS	42
2.3.1 Fringe counting interferometry	42
2.3.2 Heterodyne fringe counting interferometry.....	45
2.3.3 Two-wavelength fringe counting interferometry	46
2.3.4 Other fringe counting systems.....	46
2.3.5 The effect of laser beam diffraction on measured length.....	47
2.4 REVIEW OF LARGE-FIELD INTERFEROMETER DESIGNS.....	48
2.4.1 The Fizeau interferometer	48
2.4.2 The Fabry-Perot interferometer	50
2.4.3 The Michelson interferometer.....	50
2.4.4 The Twyman-Green interferometer	51
2.4.5 Other designs of length measuring interferometer	52
2.5 PRIMARY LENGTH BAR INTERFEROMETER - BASIC INTERFEROMETER TYPE	53
2.5.1 The technique of 'wringing'.....	54
CHAPTER 3 THE DESIGN OF THE INTERFEROMETER.....	57
3.1 OPTICAL DESIGN OF THE INTERFEROMETER	57
3.1.1 Description of the interferometer	57
3.1.2 The interferometer chamber.....	61
3.2 DETAILS OF INTERFEROMETER COMPONENTS	64
3.2.1 Design of lasers used with the interferometer.....	64
3.2.1.1 Helium-neon laser theory	64
3.2.1.2 Single mode laser wavelength stabilisation schemes	67
3.2.1.3 Laser frequency-stabilisation using saturated absorption	68
3.2.1.4 Zeeman-Stabilised 633 nm Lasers.....	69

3.2.1.5 Calibration of the Zeeman-stabilised 633 nm laser	72
3.2.1.6 Calibration of 612 nm and 543 nm lasers.....	73
3.2.2 The optical fibre illumination delivery system	73
3.2.3 Reference mirror assembly	76
3.2.4 Design of the imaging optics	78
3.2.5 Design of length bar supports.....	80
3.2.6 Length bar support carriage design.....	82
3.2.7 Design of the adjustable mirror in the measurement arm	85
3.2.8 External equipment rack and computing equipment	85
3.3 DOUBLE-ENDED INTERFEROMETRY	88
3.3.1 The wringing film thickness	88
3.3.2 Surface roughness.....	88
3.3.3 Phase change on reflection	89
3.3.4 Effect of surface form errors of measuring faces	91
3.4 OTHER DOUBLE-ENDED DESIGNS	91
3.5 DOUBLE-ENDED INTERFEROMETRY IN THE PRIMARY LENGTH BAR INTERFEROMETER	93
3.5.1 Analysis of new double-ended interferometer	95
3.5.2 Double-ended phase-stepping	96
CHAPTER 4 ALIGNMENT, COHERENCE AND OPTICAL TESTING.....	99
4.1 ALIGNMENT OF THE INTERFEROMETER.....	99
4.1.1 Approximate alignment of interferometer	99
4.1.1.1 Laser beam launching into fibres.....	99
4.1.1.2 Component positioning in interferometer	100
4.1.1.3 Fibre positioning in collimator	100
4.1.1.4 Reference mirror alignment	100
4.1.1.5 Measurement beam alignment	101
4.1.1.6 Alignment of length bars with measurement beam	101
4.1.1.7 Alignment for double-ended interferometry	102
4.1.2 Accurate alignment of the interferometer	103
4.1.2.1 Cosine error due to measurement beam mis-alignment	104
4.1.2.2 Alignment of the three interferometer axes	105
4.1.2.3 Two-fibre autocollimation technique	105
4.1.3 Obliquity effect due to position and size of light source	108
4.1.3.1 Obliquity effect due to source size - full derivation	109
4.1.4 Collimation check using a shearing plate interferometer	111
4.1.5 Tilt in the measurement beam	112
4.1.5.1 Prismatic dispersion at the beamsplitter.....	114
4.1.5.2 Methods for compensation of tilt.....	114
4.1.6 Chromatic aberration - tolerance on collimator focal position	115
4.1.7 Optical component quality and spherical aberration	116
4.1.7.1 Quality of optical components.....	116
4.1.7.2 Effect of spherical aberration in collimator.....	116
4.1.8 Effect of squareness of length bar on measured length.....	118
4.2 COHERENCE IN THE INTERFEROMETER	119
4.2.1 Temporal coherence.....	120
4.2.2 Spatial coherence	121
4.2.2.1 An approximate estimate of the spatial coherence.....	121
4.2.2.2 Detailed estimate of the spatial coherence.....	122
CHAPTER 5 FRINGE ANALYSIS & PHASE-STEPPING INTERFEROMETRY	135
5.1 ANALYSIS OF INTERFERENCE FRINGES	135
5.1.1 Introduction to interference fringe analysis.....	135
5.1.2 Fringe skeletonisation methods.....	137
5.1.3 Fourier transform methods	137
5.1.4 Temporal heterodyning methods.....	138
5.1.5 Spatial heterodyning methods	138
5.1.6 Phase locking methods.....	139
5.1.7 Summary of phase measurement methods	139
5.2 PHASE-STEPPING INTERFEROMETRY (PSI).....	140
5.2.1 History of PSI.....	140
5.2.2 Basic theory of PSI	141
5.2.3 Derivation of generic PSI equations	141
5.2.4 Typical applications of PSI	144

5.2.5 Phase variation methods for PSI	144
5.3 PHASE STEPPING TECHNIQUES	144
5.3.1 Basic phase-stepping techniques	144
5.3.2 Phase-shifting interferometry	147
5.3.3 Four quadrant arctangent routine.....	147
5.3.4 Two position phase-stepping technique.....	148
5.3.5 Three position phase-stepping technique	149
5.3.6 Four position phase-stepping technique	150
5.3.7 Errors for three and four position techniques.....	150
5.3.7.1 Error due to phase stepper error	150
5.3.7.2 Error due to detector response.....	153
5.3.7.3 Error due to multiply-reflected beams	154
5.3.7.4 Error due to quantisation noise during digitisation	155
5.4 AN ERROR-COMPENSATING FIVE POSITION TECHNIQUE	156
5.5 IMPLEMENTATION OF THE FIVE POSITION TECHNIQUE IN THE PRIMARY INTERFEROMETER	160
CHAPTER 6 DATA PROCESSING.....	165
6.1 OVERVIEW OF THE DATA PROCESSING.....	165
6.2 COMPUTING SYSTEM.....	167
6.3 IMAGE PROCESSING	168
6.3.1 Interferogram digitisation.....	168
6.3.2 Phase extraction.....	170
6.3.3 Discontinuity removal	171
6.4 MULTIPLE-WAVELENGTH INTERFEROMETRY.....	175
6.4.1 Multiple wavelength analysis.....	175
6.4.2 Limit to multiple-wavelength technique due to source instability	179
6.4.3 Multiple wavelength algorithm - method of exact fractions	180
6.5 FLATNESS AND PARALLELISM MEASUREMENTS	184
6.5.1 Measurement of parallelism (variation)	184
6.5.2 Measurement of flatness	186
6.5.3 Example measurements.....	187
6.6 COMPUTER PROGRAM	189
6.7 DOUBLE-ENDED ANALYSIS	192
CHAPTER 7 REFRACTOMETRY	197
7.1 REFRACTIVITY OF AIR	197
7.1.1 Dispersion factor.....	197
7.1.2 Density factor.....	198
7.2 EDLÉN'S EQUATIONS FOR THE REFRACTIVITY OF AIR.....	199
7.2.1 Refractivity of standard air.....	199
7.2.2 Corrections for temperature, pressure, water vapour and CO ₂	199
7.3 EFFECTS OF PRESSURE, TEMPERATURE, HUMIDITY AND CO ₂ ON REFRACTIVITY.....	201
7.3.1 Pressure measurement	201
7.3.2 Temperature measurement.....	202
7.3.3 Humidity measurement	203
7.3.4 Carbon dioxide measurement	204
7.3.5 Air parameter measurement order	204
7.4 OTHER WAVELENGTH COMPENSATION TECHNIQUES	205
7.4.1 Two-wavelength compensation.....	205
7.4.2 Gas refractometry	205
7.4.2.1 Gas refractometer design	206
7.4.2.2 Refractometer operation.....	207
7.4.2.3 Corrections and errors.....	207
7.4.2.4 Comparison between Edlén and refractometer.....	209
7.5 ADDENDUM.....	212
CHAPTER 8 THERMAL EXPANSION.....	215
8.1 THERMAL REQUIREMENTS	215
8.2 TEMPERATURE CONTROL SYSTEM.....	217
8.3 TEMPERATURE MEASUREMENT SYSTEM	219
8.3.1 PRTs and resistance bridge details	219
8.3.2 Calibration of PRTs	220
8.3.3 Temperature measurements using ITS-90	221

8.4 STABILITY OF TEMPERATURES INSIDE CHAMBER	223
8.4.1 Measurements at 20 °C	223
8.4.2 Heating from 20 °C to 30 °C	224
8.5 CALCULATION OF THERMAL EXPANSION COEFFICIENTS	226
8.6 ERRORS IN α AND β	227
8.6.1 Error propagation method - calculation of error in α and β	227
8.6.2 Least-squares fit to data with errors in both coordinates	230
8.7 EXAMPLE OF THERMAL EXPANSION MEASUREMENT	231
CHAPTER 9 PERFORMANCE OF THE INTERFEROMETER	235
9.1 ASSESSMENT OF THE INTERFEROMETER	235
9.2 FRINGE FRACTION MEASUREMENTS	236
9.3 ZERO-LENGTH MEASUREMENTS	237
9.4 CENTRAL LENGTH MEASUREMENTS	238
9.5 FLATNESS & PARALLELISM MEASUREMENTS	240
9.6 THERMAL EXPANSION MEASUREMENTS	242
9.7 GAUGE BLOCK MEASUREMENTS	245
9.8 DOUBLE-ENDED MEASUREMENTS	245
9.9 CONCLUSIONS ON PERFORMANCE	247
CHAPTER 10 UNCERTAINTY OF MEASUREMENTS	249
10.1 THE NATURE OF ERRORS	249
10.1.1 The 'orthodox' theory of errors	249
10.1.2 Combination of errors	250
10.1.3 Random errors	251
10.1.4 Systematic errors	252
10.2 BIPM RECOMMENDATIONS ON ERROR ASSESSMENT	252
10.3 COMPARISON OF 3 THEORIES OF ERROR AND RECOMMENDATIONS	253
10.4 SOURCES OF UNCERTAINTY	256
10.4.1 Air pressure measurement	257
10.4.2 Air temperature measurement	258
10.4.3 Air humidity measurement	259
10.4.4 Air CO ₂ measurement & Edlén's equations	260
10.4.5 Laser wavelength	261
10.4.6 Mechanical - optical effects	262
10.4.7 Bar expansivity at 20 °C	263
10.5 SUMMATION OF UNCERTAINTIES	264
10.6 POSSIBLE STEPS TO IMPROVE THE ACCURACY	266
10.7 COMBINED UNCERTAINTY BUDGETS OF INSTRUMENTS	266
CHAPTER 11 CONCLUSIONS	269
11.1 CONCLUSIONS	269
11.2 PROPOSALS FOR IMPROVEMENTS	271
11.3 ACKNOWLEDGEMENTS	274
APPENDIX A OPTO-MECHANICAL EQUIPMENT LIST	277
APPENDIX B QUALITY OF OPTICAL SURFACES	281
APPENDIX C FLEXING OF LENGTH BARS	283
C.1 FLEXING OF A LENGTH BAR DUE TO ITS OWN WEIGHT	283
C.2 DERIVATION OF POSITIONS OF AIRY POINTS	284
C.3 COMPENSATION FOR MASS OF WRUNG FLAT	287
C.4 EFFECT OF FLEXURE OF BARS ON THEIR LENGTH	293
APPENDIX D PRISMATIC COMPRESSION OF LENGTH BARS	295
APPENDIX E CONNECTORS & CONNECTIONS	299
APPENDIX F PUBLICATIONS BY THE AUTHOR	305
END NOTE	323

LIST OF FIGURES

Figure 1.1 - Cross-section of International Prototype Metre designed by Tresca	21
Figure 1.2 - Frequency chain for methane laser frequency determination.....	24
Figure 1.3 - Trends in accuracy of determination of length and speed of light	28
Figure 1.4 - A set of gauge blocks, 2 long series gauge blocks and 2 length bars	30
Figure 1.5 - Gauge blocks being wrung to a platen.....	31
Figure 1.6 - NPL Length Bar Machine	34
Figure 1.7 - Traceability of length measurements	35
Figure 1.8 - Iodine-stabilised He-Ne Primary laser	36
Figure 1.9 - NPL design Zeeman stabilised laser	37
Figure 2.1 - Schematic diagram of a fringe counting interferometer	43
Figure 2.2 - Use of fringe counting to measure the length of an object.....	44
Figure 2.3 - Schematic diagram of a heterodyne fringe counter.....	45
Figure 2.4 - Laser beam waist - alteration of effective propagation speed.....	47
Figure 2.5 - Standard Fizeau interferometer.....	49
Figure 2.6 - Modification to Fizeau interferometer	49
Figure 2.7 - Normalised intensity profiles of fringes in a Fizeau interferometer.....	50
Figure 2.8 - Fabry-Perot interferometer	50
Figure 2.9 - Michelson interferometer.....	51
Figure 2.10 - Twyman-Green interferometer	51
Figure 2.11 - Coherence depth doubling by positioning of reference mirror	52
Figure 2.12 - Kösters-Zeiss interference comparator	53
Figure 2.13 - Twyman-Green interferometer for the measurement of length bars	53
Figure 2.14 - Wringing a platen onto the end of a length bar	54
Figure 3.1 - Schematic design of interferometer optics.....	58
Figure 3.2 - Diagram showing the opto-mechanical components of the interferometer	59
Figure 3.3 - Perspective view of optical components and directions of beams	60
Figure 3.4 - Lid of chamber	63
Figure 3.5 - Aluminium 'collar' used as side walls of chamber.....	63
Figure 3.6 - Optical table used as base of interferometer chamber	64
Figure 3.7 - Energy levels in the He-Ne gas laser for 632.8 nm radiation	65
Figure 3.8 - Schematic diagram of an iodine-stabilised He-Ne laser.....	69
Figure 3.9 - Magnetic splitting of neon	70
Figure 3.10 - Refractivities of two Zeeman modes in 632.8 nm laser mode	70
Figure 3.11 - Calibration scheme for Zeeman stabilised laser	72
Figure 3.12 - Synchronisation diagram for red Zeeman stabilised laser	73
Figure 3.13 - Focusing of laser beam into fibre core	74
Figure 3.14 - Coupling efficiency of fibre launch.....	74
Figure 3.15 - Three source coupling using 2:2 couplers	75
Figure 3.16 - Three source coupling and return spot detection using 2:1 couplers	75
Figure 3.17 - Reference mirror assembly	77
Figure 3.18 - Exaggerated tilting of reference mirror when translating horizontally.....	78
Figure 3.19 - The imaging optics of the NPLBI	79
Figure 3.20 - Comparison of end face sizes of length bars & long series gauge blocks	79
Figure 3.21 - Video prints showing different diffraction effects at the edge of bars	80
Figure 3.22 - Length bar supports with integral PRTs, end view	81
Figure 3.23 - Length bar supports with integral PRTs, side view.....	81
Figure 3.24 - Length bar support carriage.....	83

Figure 3.25 - View of carriage, supports, length bars and optics inside chamber.....	84
Figure 3.26 - Flexure system for tilting mirror.....	85
Figure 3.27 - Schematic front view of equipment rack.....	86
Figure 3.28 - Schematic diagram of interferometer & external equipment	87
Figure 3.29 - Double-ended interferometer of Dorenwendt	91
Figure 3.30 - Analysis of triangular interferometer of Dorenwendt	92
Figure 3.31 - Additional roof-mirror optics for double-ended interferometry.....	94
Figure 3.32 - Double ended image	95
Figure 3.33 - Analysis of new double-ended interferometry.....	95
Figure 4.1 - Aligning the measurement beam with the length bar	101
Figure 4.2 - Incorrect adjustment of roof mirror orthogonality	103
Figure 4.3 - Correct adjustment of roof mirror orthogonality	103
Figure 4.4 - Cosine error of measurement beam	104
Figure 4.5 - Three fibre system	105
Figure 4.6 - Autocollimation arrangement	106
Figure 4.7 - Detected intensity during (a) radial and (b) axial positioning of the fibre	107
Figure 4.8 - Obliquity effect due to a circular aperture on axis of interferometer	109
Figure 4.9 - Convergence of un-collimated wavefront	112
Figure 4.10 - Difference in tilt between wavelengths 633 nm, 612 nm, 543 nm	113
Figure 4.11 - Chromatic dispersion - effect on focal length of collimator.....	115
Figure 4.12 - Spherically aberrated wavefront	116
Figure 4.13 - Interference between two spherically aberrated beams.....	117
Figure 4.14 - Non-square, singularly non-parallel length bar	118
Figure 4.15 - Coherence of an extended source with slit and screen	122
Figure 4.16 - Source sphere centred on image plane origin.....	123
Figure 4.17 - Coherence between source and image planes separated by R.....	125
Figure 4.18 - Coherence from a uniform circular source centred at origin.....	126
Figure 4.19 - Variation of coherence for the Primary Length Bar Interferometer	129
Figure 4.20 - Section through figure 4.19 showing detail	129
Figure 4.21 - Region of figure 4.19 about coherence limit of 0.88.....	130
Figure 4.22 - Double-ended interferogram showing different fringe contrasts.....	131
Figure 4.23 - Double-ended images for three wavelengths	132
Figure 5.1 - Example Twyman-Green interferometer for optical testing	136
Figure 5.2 - Idealised interferometer for testing surfaces.....	142
Figure 5.3 - Errors in general 3, 4 and 5-position techniques for phase step error.....	152
Figure 5.4 - Variation of phase step factor as a is varied.....	157
Figure 6.1 - Flow diagram of data processing	166
Figure 6.2 - Organisation of framestore memory	168
Figure 6.3 - Simulated intensity arrays for the first 4 digitised images	169
Figure 6.4 - Typical digitisation histogram	170
Figure 6.5 - Typical digitisation histogram but with more dark fringes in the image	170
Figure 6.6 - Simulated phase map	171
Figure 6.7(a) - Discontinuity removal 1st pass across top line of image	172
Figure 6.7(b) - Discontinuity removal - comparison of phase of top line with next	172
Figure 6.7(c) - Discontinuity removal 1st pass - continuation until middle of image	173
Figure 6.8 - Discontinuity removal 2nd pass from left to right	173
Figure 6.9 - Simulated phase map after discontinuity removal	174
Figure 6.10 - Phase maps after discontinuity and tilt removal,	175
Figure 6.11 - Coincidences for two wavelengths 633 nm & 543 nm	177
Figure 6.12 - Coincidences for three wavelengths 633 nm, 543 nm, 612 nm.....	178
Figure 6.13 - Measurement of deviation from parallelism of measurement faces	185

Figure 6.14 - Measurement of deviation from flatness of measurement face	186
Figure 6.15 - Example showing errors in the phase map	187
Figure 6.16 - Example measurements of flatness and parallelism (variation).....	188
Figure 6.17 - Program flowchart (simplified).....	189
Figure 6.18 - Example printout from interferometer.....	192
Figure 7.1 - Dispersion of standard air over the region 350 nm - 700 nm.....	198
Figure 7.2 - Gas refractometer schematic	206
Figure 7.3 - Comparison of Edlén and refractometer	210
Figure 7.4 - Comparison of Edlén and refractometer	210
Figure 7.5 - Periodic error due to drift	211
Figure 8.1 - Spiral of pipework on lid and baseplate.....	217
Figure 8.2 - Heating, cooling and insulation of interferometer	218
Figure 8.3 - Fringe distortion due to opening of chamber at raised temperature	219
Figure 8.4 - Triple point of water cell used for temperature calibrations.....	221
Figure 8.5 - Stability of air temperature inside chamber at 20 °C	223
Figure 8.6 - Stability of bar and air temperatures at 20 °C	224
Figure 8.7 - Comparison of heating rates using PRT on baseplate or in water bath	225
Figure 8.8 - Temperature measurements of air temperature gradients	226
Figure 8.9 - Least squares quadratic fit to thermal expansion data for 1000 mm bar	232
Figure 9.1 - Comparison of length bar measurements from three instruments.....	239
Figure 9.2 - Photograph of screen showing results for a 1000 mm length bar	241
Figure 9.3 - Example measurement of a gauge block	245
Figure 9.4 - Double-ended image, stored in framestore during measurement.....	246
Figure 9.5 - Three phase maps obtained during a double-ended measurement.....	246
Figure 10.1 - Four classes of experimental error	254
Figure 10.2 - Plot of total uncertainty in length measurement over range 0.1 - 1.5 m	265
Figure C.1 - Bar supported at two points	284
Figure C.2 - Effect of support point position, a, on change in length, dL, of bar	286
Figure C.3 - Bar supported at new support points with flat attached to one face	287
Figure C.4 - Variation in vertical position and gradient (dashed line) of neutral plane	294
Figure D.1 - Compression on cross section of bar	295
Figure D.2 - Contraction of a steel length bar, standing vertically	298
Figure E.1 - Electronics rack, left side view of connectors	299
Figure E.2 - Electronics rack, right side view of connectors.....	300
Figure E.3 - Pump control using solid state relay	300
Figure E.4 - Motor power connections from motor to PSU	301
Figure E.5 - Motor direction control, from joystick microswitches to PSU board	301
Figure E.6 - Motor inhibit connections, microswitches on stage to PSU board	302
Figure E.7 - Motor control from joystick switches to motor.....	302
Figure E.8 - Low voltage PZT power supply connections.....	303
Figure E.9 - High voltage PZT power supply connections	303
Figure E.10 - Gas sample flow connections.....	304

LIST OF TABLES

Table 1.1 - The 7 base units of the SI system	17
Table 1.2 - Realisations of the SI units at NPL	18
Table 1.3 - Recommended wavelengths for realisation of the metre	26
Table 1.4 - Tolerances: parallelism, flatness & length for reference bars to BS 5317	33
Table 2.1 - Typical fringe interpolation errors (FT Technologies FT612AS)	44
Table 5.1 - Summary of phase measurement techniques	140
Table 5.2 - Four-quadrant lookup table	148
Table 5.3 - Non-linearity effects present for R-step algorithm	153
Table 5.4 - Phase measurement error (radians) due to digitisation quantisation noise.....	156
Table 7.1 - Typical variations in air temperature, pressure, humidity and CO ₂	201
Table 7.2 - Effect of air temperature, pressure, humidity & CO ₂ content on refractivity.....	201
Table 7.3 - Errors in pressure reading due to change from 0% to 70%RH.....	202
Table 8.1 - Constants used in ITS-90 reference equation	222
Table 8.2 - Calibration data for interferometer PRTs.....	222
Table 8.3 - Error in measured value of a (10 ⁻⁶ K ⁻¹).....	230
Table 8.4 - Measured thermal expansion data for a 1000 mm length bar.....	231
Table 9.1 - Fringe fraction sample results	237
Table 9.2 - Results of zero-length measurements	237
Table 9.3 - Comparison results of length bars measured in three instruments	238

CHAPTER 1

INTRODUCTION

“If God had wanted us to use the metric system, he would have given us ten fingers and ten toes”
Anon.

1.1 BACKGROUND

The measurement of length and the provision of length standards are of fundamental importance to any technologically developed society. The ability to measure length to a required accuracy and demonstrate that the measurement has been performed in terms of universally recognised units underlies much of the world’s trade. The world has seen many standards of length in use throughout history [1] ranging from the simple use of thumbs and feet, to the more advanced definitions relating to the wavelength and speed of light. One common theme is the continuing refinement of standards, leading to more specific definitions and more accurate methods of realising them.

Perhaps the most significant step forward in the field of length standards was made by Michelson at the end of the nineteenth century, when he measured the wavelength of the red light from cadmium in terms of the International Prototype Metre bar [2]. This established the techniques of interferometric comparison whereby material lengths are measured in terms of known wavelengths of light. The work of Michelson and his contemporaries stimulated international development of wavelength standards which resulted in the adoption in 1960 of a new definition of the metre, based on the wavelength of light from a krypton lamp [3]. The development of the laser in the 1960s produced a new wavelength source which, with careful control, could surpass the stability and accuracy of the krypton wavelength standard. This prompted a new definition of the metre [4] in 1983 which, although it abandoned the direct concept of wavelength in the definition, recommended the use of laser wavelengths in the realisation of the metre.

At the time of writing, this definition still stands and the realisation of the metre at the National Physical Laboratory (NPL) as a wavelength standard using iodine-stabilised helium-neon lasers serves as the UK’s primary length standard.

1.2 THE MEASUREMENT OF LENGTH

1.2.1 What is length?

Whilst it is quite difficult to say exactly what is meant by ‘length’, it is relatively easy to actually measure length, especially the length of material objects. Measurements of length range from the dimensions of atoms and their constituents to the size of the visible universe. Units used to describe this wide range of lengths include parsecs, light years, solar diameters, light seconds, kilometres, metres, yards, feet, inches, millimetres, microns, thousandths of an inch, micrometres, nanometres, angstroms and picometres. Each unit or sub-multiple of a unit must be connected to each other unit if the measurements at different scales are to be related. This necessitates a standardised system of units, with consistent definitions and interrelationships. The most common such system in use today is the Système International d’Unités (SI).

1.2.2 The International System of units (SI)

In the second half of the nineteenth century the inch, yard and foot were the most common units in use in Britain, but the centimetre, gram and second were also in use. These units, called the “CGS electromagnetic system” were coherent, *i.e.* there were no numerical factors other than unity used in the definitions of the derived units. There was another set of units, the “CGS electrostatic system” which was used for measurements of charge, potential and capacitance. The problem with this latter system of units was that the sizes of the units were inconvenient. In 1881 an international agreement defined new units: the volt as 10^8 CGS potential units, the ohm as 10^9 CGS resistance units and the ampere as 0.1 CGS units.

These new units were mutually coherent but were not coherent with the magnetic or mechanical units. This led Giorgi in 1902 to propose a new set of units based on the metre, kilogram, second and ampere. This allowed the magnetic field strength to be expressed in amperes per metre thus removing a factor of π from most electromagnetic formulae involving rectilinear geometry, and transferring it to formulae using cylindrical or spherical geometry.

In 1948 the 9th Conférence Générales des Poids et Mesures (CGPM) adopted these mechanical units. The 11th CGPM later added to these the candela and kelvin and the supplementary units the radian and steradian. In 1971 the 14th CGPM added the mole as the amount of substance, completing the set of 7 base units.

Quantity	Unit	Symbol
Time	second	s
Length	metre	m
Mass	kilogram	kg
Electric current	ampere	A
Thermodynamic temperature	kelvin	K
Luminous intensity	candela	Cd
Amount of substance	mole	mol

Table 1.1 - The 7 base units of the SI system

1.2.3 The definitions of the SI units

The second is the duration of 9 192 631 770 periods of the radiation corresponding to the transition between the two hyperfine levels of the ground state of the caesium-133 atom.

The metre is the length of the path travelled by light in vacuum during a time interval of 1/299 792 458 of a second.

The kilogram is the unit of mass; it is equal to the mass of the international prototype of the kilogram.

The ampere is that constant current which, if maintained in two straight parallel conductors of infinite length, of negligible circular cross-section, and placed 1 metre apart in vacuum, would produce between these conductors a force equal to 2×10^{-7} newton per metre of length.

The kelvin, unit of thermodynamic temperature, is the fraction 1/273.16 of the thermodynamic temperature of the triple point of water.

The candela is the luminous intensity, in a given direction, of a source that emits monochromatic radiation of frequency 540×10^{12} hertz and that has a radiant intensity in that direction of (1/683) watt per steradian.

The mole is the amount of substance of a system which contains as many elementary entities as there are atoms in 0.012 kilogram of carbon 12.

The *supplementary* units are defined thus:

The radian is the plane angle between two radii of a circle which cut off on the circumference an arc equal in length to the radius.

The steradian is the solid angle which, having its vertex in the centre of a sphere, cuts off an area of the surface of the sphere equal to that of a square with sides of length equal to the radius of the sphere.

From these basic units, one can derive other units to describe any quantity which can be measured, *e.g.* velocity as the rate of change of length in unit time (m s^{-1}). For a brief history of the units, see for example Kaye & Laby [5] or The International System of Units [6].

Some of these base units may be considered less fundamental than the others as they contain references to other units in their definitions. Thus the metre relies on the definition of the second, the ampere on the metre. The definition of the kilogram is also somewhat unusual in that it is the only unit which is derived from a physical object. It is also unfortunate that the unit is the *kilogram*, rather than the gram, and thus contains one of the recommended prefixes used to denote multiples and sub-multiples of units.

Unit	Realisation	Accuracy of realisation
second	Caesium beam clock	$1 \text{ in } 10^{13}$
metre	Wavelength of laser	$2.5 \text{ in } 10^{11}$
kilogram	British copy no. 18	$1 \text{ in } 10^9$
ampere	Via the Watt	$8 \text{ in } 10^6$
kelvin	Water triple point cells	$1 \text{ in } 10^4$
mole	Directly from definition	(Avogadro const: $6 \text{ in } 10^7$)
candela	Cryogenic radiometer	$1 \text{ in } 10^3$

Table 1.2 - Realisations of the SI units at NPL

For most measurements, only 5 of these base units are required: length, time, mass, amount of substance and electrical current. If required, temperature can be defined in terms of energy and hence in terms of mass, length and time, without requiring its own separate unit. The unit for radiated intensity can be expressed similarly as watts per square metre per steradian.

1.3 HISTORICAL UNITS OF LENGTH

1.3.1 Timetable of events

- 3000 BC Egyptian & Mesopotamian cubit in common use
- 12 BC Lower Germany recognises 'northern cubit' (just over 2 feet in length)
- 410 Anglo-Saxon foot used throughout Britain
- 1305 Edward I decreed that "3 dry round grains of barley makes 1 in, 12 inches equals 1 foot, 3 feet equal 1 ulna"
- 1497 Size of ulna (yard) of Henry VII same as modern to within 0.04 inch
- 1588 Elizabeth I yard same as modern yard to within 0.01 inch
- 1742 Two brass bars constructed, 42 in x 0.5 in x 0.25 in - yard engraved on surface
- 1758 Royal Commission start work on new bronze standard, 1 in square, gold inserts
- 1760 New bronze yard standard completed
- 1790 Talleyrand (Bishop of Autun) proposes new system of lengths for France
- 1791 Academy of France Commission set up to consider a new decimal scale of units
- 1791 Commission chose quadrant of meridian of Earth, terrestrial pole to equator, as basis for length standard, rejecting use of a pendulum beating the second
- 1792 Delambre and Méchain commence measurements along meridian
- 1793 National Convention (France) adopts 1 m = 443.44 lignes of Toise de Pérou
- 1795 Basic law passed adopting metric system passed by Convention
- 1799 Mètre des Archives, (platinum), 25.3 x 4 mm
constructed and adjusted by Janetti to be 1/10 000 000 of Earth quadrant,
- 1824 Royal Commission bar (yard) adopted as new primary standard for UK
- 1829 Babinet suggests use of wavelengths as length standards
- 1834 Fire in Houses of Parliament destroys yard standard
- 1837 France rescinded all weights and measures other than the metric system
- 1843 Sheepshanks & Bailey of British Royal Commission work on new bronze yard standard
- 1855 Work completed on 1 inch square bronze yard standard replacement
- 1864 Metric system sanctioned in UK by Act of Parliament
- 1872 30 new prototype metres construction started (X cross-section) 90% Pt, 10% Ir, based on design by Tresca
- 1875 Convention du Mètre signed
- 1878 Weights and Measures Act - metric system fully legalised in UK
- 1884 Britain signs Convention du Mètre
- 1889 Work on 30 prototype metres completed, deposited at BIPM
- 1889 Definition of metre in terms of Prototype Metre
- 1893 Michelson & Benoit compare Cd red line to Prototype Metre
-

1893	American yard linked to metre
1899	Metre Copy no. 16 measured as 0.999 999 400 m
1899	Bill and Order in Parliament, legalising metre copy no. 16 as UK metre
1922	Comparison of metre and yard (inch) : 1 m = 39.370 147 inches
1927	Re-definition of the metre
1956	Engraved lines on copy no. 16 re-ruled to allow measurements at 0 °C and 20 °C
1959	New value for inch : 1 inch = 25.4 mm exactly, 1 m = 39.370 079 inches
1960	Re-definition of metre in terms of krypton-86 wavelength ($\pm 4 \times 10^{-9}$)
1963	Weights and Measures Act legalises new definition of inch as 2.54 cm
1975	Speed of light fixed at 299 792 458 m s ⁻¹
1983	Definition of metre as distance travelled by light in 1/299 792 458 s (absolute)
1983	Accuracy of second $\pm 1 \times 10^{-13}$, accuracy of metre realisation as wavelength : $\pm 2 \times 10^{10}$ to $\pm 2 \times 10^9$ for lasers stabilised to saturated absorption, 10^8 for others

1.3.2 The first definition of the metre

The first Conférence Générales des Poids et Mesures [4] (CGPM) in 1889 stated of the Prototype Metre, that:

<<Ce prototype représentera désormais à la température de la glace fondante, l'unité métrique de longeur.>>

“This prototype, at the temperature of melting ice, will henceforth represent the unit of length.”

This was a rather informal definition as it did not include certain details concerning how the bar was to be supported.

(Note that the official language of the SI system of units is French: the above wording in English is only an approved translation, NOT the definition. The language difference can sometimes be subtle, but important. For instance in a later definition, use is made of the word “vide” or “vacuum”. In English, the phrase “free space” is more often used to indicate the absence of cosmological matter such as black holes and virtual particles, however there is no distinction in French between “vacuum” and “free space” [7]).

1.3.3 The 1927 definition of the metre

In 1927 the 7th CGPM adopted a new definition [3] of the metre:

<<L'unité de longueur est le mètre, défini par la distance, à 0°, des axes de deux traits médians tracés sur la barre de platine iridié déposée au Bureau International des Poids et Mesures et déclarée Prototype du mètre par la première Conférence Générale des Poids et Mesures, cette règle étant soumise à la pression atmosphérique normale et supporté par deux rouleaux d'au moins un centimètre de diamètre, situés symétriquement dans un même plan horizontal et à la distance de 571 mm l'un de l'autre.>>

“The unit of length is the metre, defined as the distance at 0° between the two lines engraved in the platinum iridium bar, deposited in the BIPM, and declared as the Prototype metre by the 1st CGPM, this standard supported at normal atmospheric pressure on two rollers, of less than 1 cm diameter, situated symmetrically in a horizontal plane, at a distance of 571 mm from each other.”

Thus the first formal definition of the metre was based on a single material object (as the kilogram is today). This bar had an ‘X’-shaped cross section designed by Tresca [3] to have maximum rigidity with minimum use of material, with the 2 lines defining the metre engraved on the neutral plane or surface of the bar, *i.e.* the surface which experiences no net compression or expansion when the bar is supported on its designated rollers. The positions of the rollers corresponded to the “Bessel points” of the bar, positioned 571 mm apart, either side of the centre. When supported at these points, the length of the bar is unchanged from its free state.

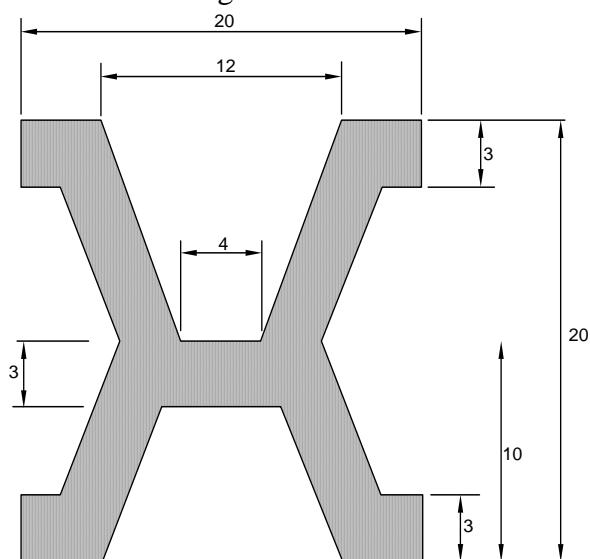


Figure 1.1 - Cross-section of International Prototype Metre (and copies), based on a design by Tresca, manufactured between 1882 and 1889 (dimensions in mm)

Although it was discovered that the original Mètre des Archives (an End Standard) was short by 0.23 mm of the size it was meant to be, namely one ten-millionth of the Earth's quadrant (realised on a meridian connecting Dunkirk and Barcelona), the new Prototype Metre was made to be the same length as its predecessor to avoid any change to the physical size of the metre which was in use.

1.3.4 The 1960 definition of the metre

In 1893 Michelson and Benoit working at the Bureau International des Poids et Mesures used an interferometer to measure the wavelength of the cadmium red line in terms of the metre. Following the definition of the metre in 1927 progress was made in the determination of the wavelengths of emission lines and also in the purification of isotopes of elements, notably krypton 86, cadmium 114, and mercury 198. This allowed the production of wavelengths from emission lines with suitably narrow spectral lines, the stability and reproducibility of which could surpass the values obtainable from the Prototype Metre. Thus in 1960 the CGPM adopted a new definition of the metre [3] based on wavelengths:

<<

1. *Le mètre est la longueur égale à 1 650 763.73 longueurs d'onde dans le vide de la radiation correspondant à la transition entre les niveaux 2p₁₀ et 5d₅ de l'atome de krypton-86.*
2. *La Définition du Mètre en vigueur depuis 1889, fondée sur le Prototype International en platine iridié, est abrogée.*
3. *Le Prototype International du Mètre sanctionné par la Première Conférence Générale des Poids et Mesures de 1889 sera conservé au Bureau International des Poids et Mesures dans les mêmes conditions que celles qui ont été fixées en 1889.>>*

which translates as

“

1. The metre is the length equal to 1 650 763.73 wavelengths in vacuum of the radiation corresponding to the transition between the energy levels 2p₁₀ and 5d₅ of an atom of krypton 86.
2. The definition of the metre in use since 1889, based on the International Prototype of platinum iridium, is abrogated.
3. The International Prototype Metre sanctioned by the first CGPM of 1889 will be conserved at the BIPM in the same conditions that it was placed in 1889.”

The source of the standard radiation recommended by the 1960 Committee was the Engelhard lamp, operated according to the following instructions [3]:

<< Conformément au paragraphe 1 de la Résolution 2 adoptée par la Onzième Conférence Générales des Poids et Mesures (octobre 1960), le Comité International des Poids et Mesures recommande que la radiation du krypton 86 adoptée comme étalon fondamental de longueur soit réalisée au moyen d'une lampe à décharge à cathode chaude contenant du krypton 86 d'une pureté non inférieure à 99 pour cent, en quantité suffisante pour assurer la présence de krypton solide à la température de 64 °K, cette lampe étant munie d'un capillaire ayant les caractéristiques suivantes: diamètre intérieur 2 à 4 millimètres, épaisseur des parois 1 millimètre environ.

On estime que la longueur d'onde de la radiation émise par la colonne positive est égale, à 1 cent-millionième (10^{-8}) près, à la longueur d'onde correspondant à la transition entre les niveaux non perturbés, lorsque les conditions suivantes sont satisfaites:

1. *le capillaire est observé en bout de façon que les rayons lumineux utilisés cheminent du côté cathodique vers le côté anodique;*
2. *la partie inférieure de la lampe, y compris le capillaire, est immergée dans un bain réfrigérant maintenu à la température du point triple de l'azote, à 1 degré près;*
3. *la densité du courant dans le capillaire est $0,3 \pm 0,1$ ampère par centimètre carré. >>*

“Conforming to paragraph 1 of resolution 2 of the 11th CGPM (October 1960), the CIPM recommends that the radiation of krypton 86 adopted as the fundamental standard of length should be realised by means of a hot cathode discharge tube containing krypton 86 of a purity not less than 99 per cent in a sufficient quantity to ensure the presence of solid krypton at a temperature of 64 K; the lamp having a capillary with the following characteristics: internal diameter 2 to 4 mm, wall thickness about 1 mm.

It is believed that the wavelength of the radiation emitted by the positive volume is equal to within about one hundred millionth (10^{-8}) of the wavelength corresponding to the transition between the unperturbed levels, provided the following conditions are satisfied:

1. the capillary is observed end-on in such a way that the rays used travel from its cathodic to its anodic end
2. the lower part of the lamp, including the capillary, is immersed in a cooling bath maintained at the temperature of the triple point of nitrogen within about 1 degree
3. the current density in the capillary is 0.3 ± 0.1 amperes per square centimetre.”

The number of wavelengths in the definition was chosen to correspond as closely as possible with the previous definition. By removing the dependence on a single material object, the new definition allowed greater access for metrologists to a more accurate length standard. However the variation in the wavelength due to impurities in the krypton and other constructional and operational details required extra “instructions for use” as recommended by the CIPM.

1.3.5 The present (1983) definition of the metre

The first simultaneous measurements of both the frequency and wavelength of light were performed in 1972 leading to a more accurate value for the speed of light [8] (see figure 1.2).

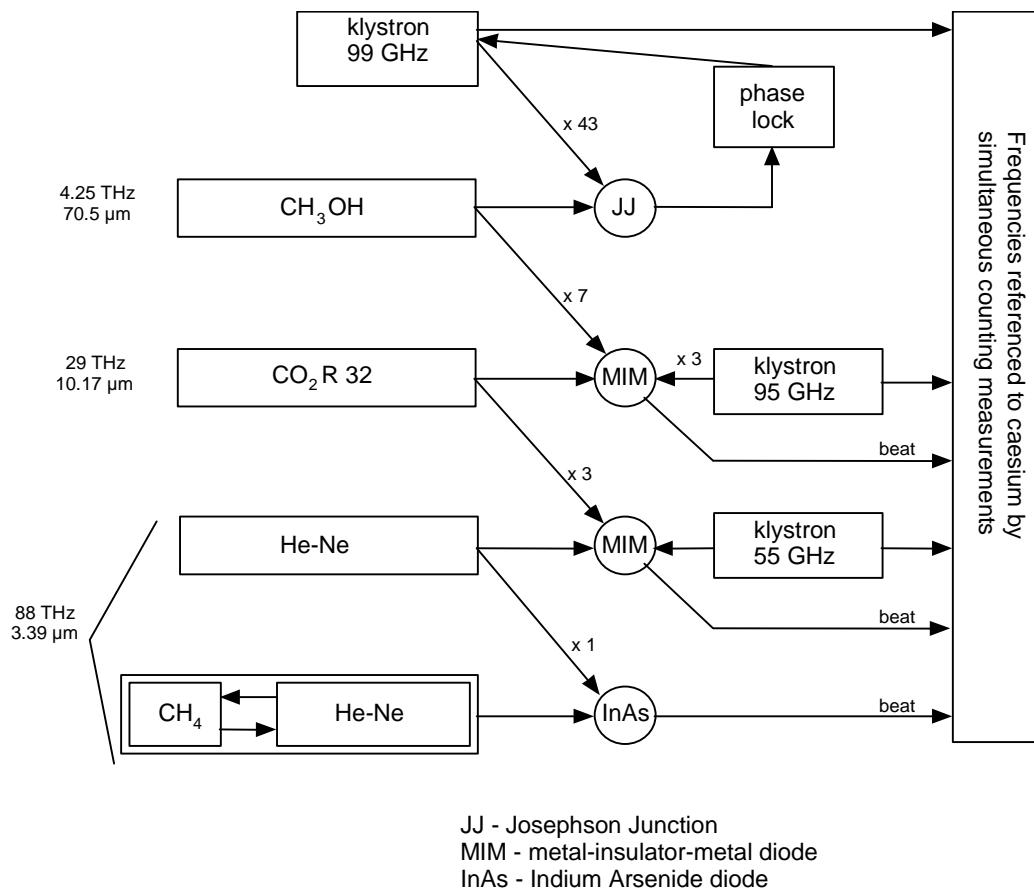


Figure 1.2 - Frequency chain for methane laser frequency determination

The uncertainty in this value was estimated to be 1.2 m s^{-1} . This value was used to link the frequencies and wavelengths of other lasers. The weak link in the chain was found to be the realisation of the metre using the krypton 86 lamp, which had an uncertainty similar to that of the measurement of the speed of light. Also, by the early 1970s the

frequencies of lasers could be compared with greater accuracy than the results could be expressed in absolute units. It seemed that a more accurate definition of the speed of light was required. To prevent discrepancies between different experiments, the Comité Consultatif pour la Définition du Mètre (CCDM) recommended a value for the speed of light which was fixed in 1975 by the 15th CGPM at $299\,792\,458 \text{ m s}^{-1}$. This implicitly contained a new definition of the metre which at the level of uncertainty within 4 parts in 10^9 could cause discrepancies with previously determined wavelengths, with the danger that a laser wavelength scale might have become established, separate from the SI system.

To avoid these problems, in 1983 the CCDM proposed a new definition for the metre based on the speed of light. This was then adopted by the Comité Consultatif des Unités (CCU), the Comité International des Poids et Mesures (CIPM) and the 17th Conférence Générale des Poids et Mesures (CGPM) in October 1983. The definition was chosen to be both intelligible enough to be understood by physics students and yet be precise enough to allow metrologists working at the frontiers of measurement to use it as a working definition. Thus the definition was kept as simple as possible, with an additional recommendation of how to use it in practice [4]:

<<Le mètre est la longueur du trajet parcouru dans le vide par la lumière pendant une durée de $1/299,792,458$ de seconde.>>

<<la définition du mètre en vigueur depuis 1960, fondée sur la transition entre les niveaux $2p_{10}$ et $5d_5$ de l'atome de krypton 86, soit abrogée.>>

“The metre is the length of the path travelled by light in vacuum during a time interval of $1/299\,792\,458$ of a second.”

“The definition of the metre in use since 1960, based on the transition between the two lines $2p_{10}$ and $5d_5$ of the krypton atom, is abrogated.”

RECOMMENDATION 1(CI-1983)

The Comité International des Poids et Mesures

recommends

- that the metre be realised by one of the following methods:

- (a) by means of the length l of the path travelled in vacuum by a plane electromagnetic wave in a time t ; this length is obtained from the measured time

t , using the relation $l = c.t$ and the value of the speed of light in a vacuum $c = 299\ 792\ 458\ \text{m s}^{-1}$;

- (b) by means of the wavelength in vacuum λ of a plane electromagnetic wave of frequency f ; this wavelength is obtained from the measured frequency f , using the relation $\lambda = c/f$ and the value of the speed of light in vacuum $c = 299\ 792\ 458\ \text{m s}^{-1}$;
- (c) by means of one the radiations from the list below, whose stated wavelength in vacuum, or whose stated frequency, can be used, provided that the given specifications and good practice are followed;

Laser	Absorber	Trans ⁿ , line, comp ^t	f/MHz	λ/nm
He-Ne	CH ₄	n ₃ ,P(7),F ₂	88 376 181.608	3392.231 397
He-Ne	¹²⁷ I ₂	17-1,P(62),O	520 206 808.51	576.294 760 27
He-Ne	¹²⁷ I ₂	11-5,R(127),i	473 612 214.8	632.991 398 1
He-Ne	¹²⁷ I ₂	9-2,R(47),O	489 880 355.1	611.970 769 8
Ar ⁺	¹²⁷ I ₂	43-0,P(13)a ₃	582 490 603.6	514.673 466 2

Table 1.3 - Recommended wavelengths for realisation of the metre

1.3.6 Limitations of the present realisation of the metre

The first limitation of the current realisation of the metre lies in its dependence on the speed of light. The value $299\ 792\ 458\ \text{m s}^{-1}$ is the latest and most accurate result with an uncertainty of $\pm 1.2\ \text{m s}^{-1}$ and is based on measurements of the frequency and wavelength of a helium-neon laser radiation, stabilised to an infra-red transition in methane at about $3.39\ \mu\text{m}$, with later confirmation at a wavelength of $9.3\ \mu\text{m}$ using a stabilised carbon dioxide laser [9]. Although the definition is fixed, the realisation in absolute terms contains an uncertainty at the level of 4 parts in 10^9 due to the uncertainty in the speed of light. This has been recognised by the CGPM whose recommendation is that any changes in the measured value of the speed of light will be ascribed to discrepancies between the maintained metre and the metre of the SI definition.

The second source of uncertainty lies in the uncertainty in the realisation of the second through the use of the caesium beam clock, which is 1 in 10^{13} . This directly affects the

uncertainty in the frequency of the iodine-stabilised laser, which in turn affects the wavelength.

It is assumed that during any realisation of the metre, the experimenter will seek to take account of any effects of relativity or other influences which would affect either the duration of the second or the speed of light.

1.3.7 Future realisations of the metre

The first definition of the metre in terms of the metre bar stood for 130 years until it was replaced with the second definition in terms of wavelengths. This in turn was replaced after only 23 years with the current definition in terms of light path. How long will the current definition stand? Looking at the increasing accuracy of length measurements and the corresponding reduction in the uncertainty of the realisation of the metre [7], as shown in figure 1.3, it appears that a new realisation will probably be required early in the next century. In particular, the demands of nanometrology [10,11] and the increasing accuracy of commercial interferometers which are now accurate to between 10^{-8} and 10^{-9} make a more accurate realisation a necessity in 10 to 20 years time.

According to Petley [12], the whole SI system is dynamic. During the last two decades the accuracy with which the fundamental constants can be measured has caught up with the accuracy to which the definitions of many of the base units can be realised. In many cases fundamental constants are now either part of the definition of a base unit or used to maintain a reproducible secondary unit, *e.g.* the Josephson effect as a voltage standard. In principle it should thus be possible for users to realise their own base units in terms of these fundamental constants, however this is not yet the case: commercial atomic clocks and helium-neon lasers still require calibration against primary standards.

The ability to measure physical quantities has improved over the last decade, sometimes at the rate of a tenfold increase in accuracy per decade. This progress has actually been made in discrete steps. There is a time lag of approximately 10 years between the advancement of a new accuracy of calibration and the routine achievement of the same accuracy in commercial products. Similarly, work in national standards laboratories is at an accuracy above that of the commercial sector, but below the most accurate possible. As the accuracy of commercial instruments approaches that of calibrations offered by national standards laboratories and the accuracy of those calibrations

approaches that of the national standard, the standard is replaced by a newer one of greater accuracy, and the cycle starts again.

Note also that it is a new realisation of the metre that will be required, not a new definition. The current definition is sufficiently open-ended, being based on a fundamental constant, to allow future methods of realising the metre to be incorporated, without the need to change the definition. In particular new sources will be added to the list of recommended radiations and their uncertainties will be reduced as new frequency measurements are made. Thus the realisation of the unit of length will be kept up to date, delaying the time when a new definition will be needed. This does however lead to an extensive list of recommended wavelengths and some rationalisation may be required as older sources become less commonly used.

Note also that the definition does not mean that the speed of light can never change; rather that, if it does, then the size of the metre will change in accordance so that the numerical value 299 792 458 is preserved.

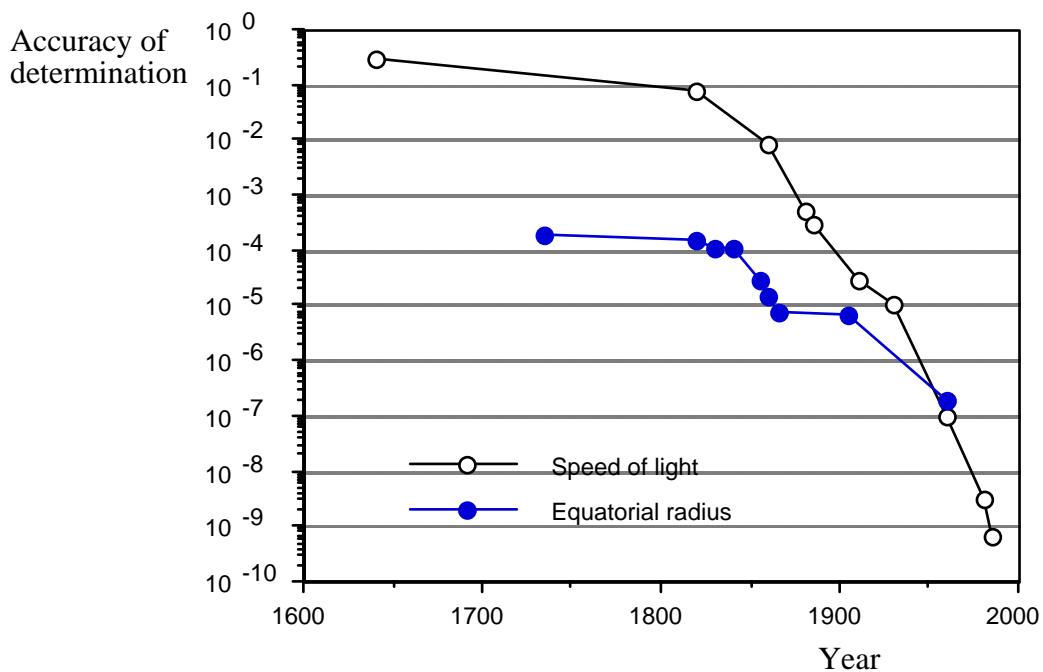


Figure 1.3 - Trends in accuracy of determination of length and speed of light

Future realisations of the metre may include some form of ion-trapping arrangement in which individual atoms or ions are cooled by lasers and trapped in fields, where they undergo transitions between energy levels [13]. It is thought that these “single oscillators” will exhibit exceptional frequency stability and purity, of the order of 10^{-18} .

Other proposals for slow-atom frequency standards include laser cooled atomic fountains [14] with an estimated accuracy of 10^{-16} .

It will be necessary to link the frequency of these transitions to the caesium clock, and then establish a new trapped ion standard as the frequency standard. This will allow accurate beat frequency comparison with stable laser sources, which can then serve as length standards. Thus it may require a change of viewpoint from the current notion of measuring length against wavelengths, to timing the flight of the light, using the laser as an accurate frequency standard or clock.

1.4 SECONDARY LENGTH STANDARDS

1.4.1 Modern secondary length standards

Although the unit of length is the metre, and the realisation of the metre is via the wavelength or frequency of a frequency-stabilised laser, it is inappropriate to use this primary standard for everyday measurements and so use is made of secondary length standards for less demanding measurements. There are two types of length standard in use: line standards and end standards. Examples of line standards include survey tapes, photomasks, the Prototype Metre and of course rulers. End standards include gauge blocks, length bars, the Mètre des Archives, Hoke gauges and combination bars.

The measurement of line standards usually requires an instrument which works on the same principles as, or is actually a travelling microscope which traverses the distance between two or more lines marked on the standard. Line standards range from the dimensions of micro-lithographic standards of the $1\text{ }\mu\text{m}$ size, up to 50 m survey tapes.

End standards usually take the form of bars of durable material and have flat, polished end faces, the separation between them defining the length of the standard. Calibrations of end standards are made using mechanical probes or interferometrically. End standards are widely used in industry for calibrating verniers, micrometers and for verifying the performance of Co-ordinate Measuring Machines (CMMs) [15]. They are more suitable for engineering measurements than line standards as they represent a “mechanical” length which can be physically probed.

1.4.2 Gauge blocks and length bars

The two most common forms of end standard in use throughout Europe are *gauge blocks* and *length bars*. These are material standards and take the form of rectangular or

circular bars made to various nominal lengths according to certain standards [16,17,18,19,20,21].



Figure 1.4 - A set of gauge blocks, 2 long series gauge blocks and 2 length bars (foreground)

1.4.3 Gauge blocks

A gauge block is a block of rectangular section, made of durable material, with one pair of plane, mutually parallel measuring faces. The length of a gauge block at a particular point on the measuring face is the perpendicular distance between this point and a rigid plane surface of the same material and surface texture upon which the other measuring face has been wrung. ‘Wringing’ is a technique by which very flat, lapped surfaces can be made to adhere to one another by molecular attraction (see § 2.5.1).

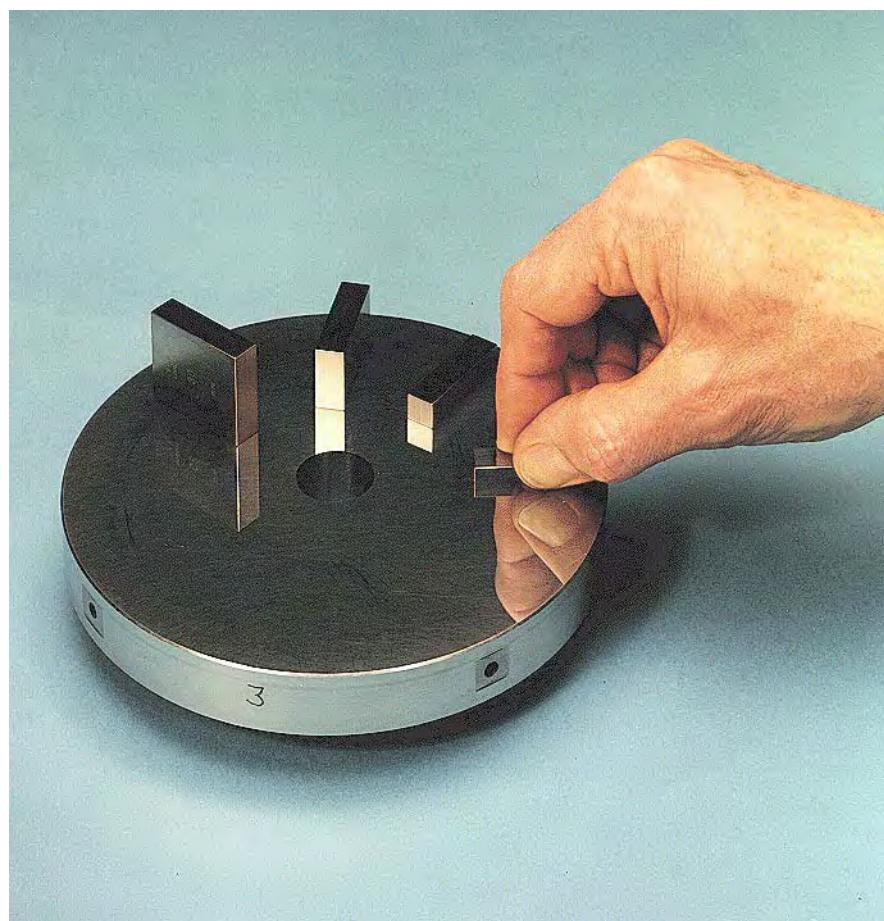


Figure 1.5 - Gauge blocks being wrung to a platen

The measured length of a gauge block is corrected to the reference temperature of 20 °C and standard air pressure 101 325 kPa (1 013.25 mbar). The lengths of gauge blocks up to and including 100 mm refer to the length of the gauge block in the vertical position, *i.e.* with the measuring faces horizontal. The lengths of gauge blocks over 100 mm refer to the length of the gauge block in the horizontal position, the block being supported on one of the smaller side faces without additional stress by two suitable supports, each at a distance of 0.211 times the nominal length from the ends. This is to account for prismatic compression of the gauge under its own weight, when standing vertically and ‘sagging’ when supported horizontally (see Appendices C & D).

1.4.4 Length Bars

Length bars are end standards of cylindrical cross-section, 22 mm in diameter, having flat, parallel end faces finished by lapping. They are made from high quality tool steel, free from non-metallic inclusions. The length of a bar is defined with the bar mounted horizontally (and referred to the standard reference temperature of 20 °C) as *the distance from the centre of one of its faces to a flat surface in wringing contact with the opposite face, measured normal to the surface*.

25 mm bars are hardened throughout their length. Bars over 25 mm up to and including 125 mm are hardened either throughout their length or at the ends only for a distance of not less than 4 mm. Longer bars are hardened at the ends only for a distance of about 6 mm and not less than 4 mm from each end.

1.4.5 Definitions and specifications for Reference Grade length bars

Additionally, the British Standard (BS 5317) places detailed specifications on length bars, depending on the grade of manufacture, as follows.

Deviation from flatness

The minimum distance between two parallel planes which just envelop the measuring face. The maximum permissible values are given in table 1.4.

Deviation from parallelism

The difference between the maximum and minimum lengths at any points on the measuring faces measured perpendicular to the surface to which one face is wrung. The maximum permissible values are given in table 1.4.

Deviation from squareness

The minimum distance between two parallel planes normal to the axis of the bar which just envelop the measuring face under consideration.

Diameter

The diameter of each bar shall be uniform within 15 μm for bars up to 300 mm in length, 25 μm for bars over 300 mm up to and including 600 mm in length, and within 50 μm for bars longer than 600 mm.

Straightness

The body shall be straight within 10 μm per 100 mm of length.

Squareness

The end faces of all grades of bars shall be square with the axis of the bar to within 1.2 μm over the diameter of the face for bars up to and including 400 mm in length and to within 2.5 μm for bars over 400 mm in length.

Nominal length mm	Tolerances on accuracy of faces		Tolerance on length at 20 °C ± µm
	Flatness ± µm	Parallelism ± µm	
up to 25	0.08	0.08	0.08
50	0.08	0.10	0.12
75	0.10	0.16	0.15
100	0.10	0.16	0.20
125	0.10	0.20	0.25
150	0.10	0.20	0.30
175	0.15	0.20	0.30
200	0.15	0.20	0.35
225	0.15	0.20	0.40
250	0.15	0.30	0.40
275	0.15	0.30	0.45
300	0.15	0.30	0.50
400	0.15	0.30	0.65
500	0.15	0.30	0.80
600	0.15	0.30	0.95
700	0.15	0.30	1.10
800	0.15	0.30	1.25
900	0.15	0.30	1.40
1000	0.15	0.30	1.55
1200	0.15	0.30	1.85

Table 1.4 - Tolerances on parallelism, flatness and length for reference bars according to BS 5317:1976

1.4.6 Calibration of End Standards of length

Gauge blocks and length bars of the highest grade of accuracy are calibrated at the National Physical Laboratory (NPL). These calibrations are traceable to the definition of the metre through the use of stabilised laser wavelengths. For gauge block calibrations, this traceability is provided directly by the NPL automatic Gauge Block Interferometer [22].

This instrument uses calibrated frequency-stabilised lasers to relate the lengths of gauge blocks to the realisation of the metre in terms of the wavelength emitted by an iodine-stabilised laser. The instrument is computer controlled and performs corrections to take account of the variations in the refractive index of the air, and the temperature of the gauge blocks.

For length bar calibrations, use is made of the NPL Length Bar Machine. This instrument functions as a comparator in which the length of a test bar is compared to the length of a short (3 in) standard length bar which has been measured in the Gauge Block Interferometer (GBI). The instrument uses a mechanical probing system with the separation of the two probes monitored using a commercial fringe counting interferometer. Corrections are made for the variations in refractive index of the air inside the interferometer and for the thermal expansion of the length bars.

This instrument simply provides a measurement of the central length of the bar and relies on the calibrated standard for its traceability. The instrument can experience problems with alignment which can limit its accuracy.

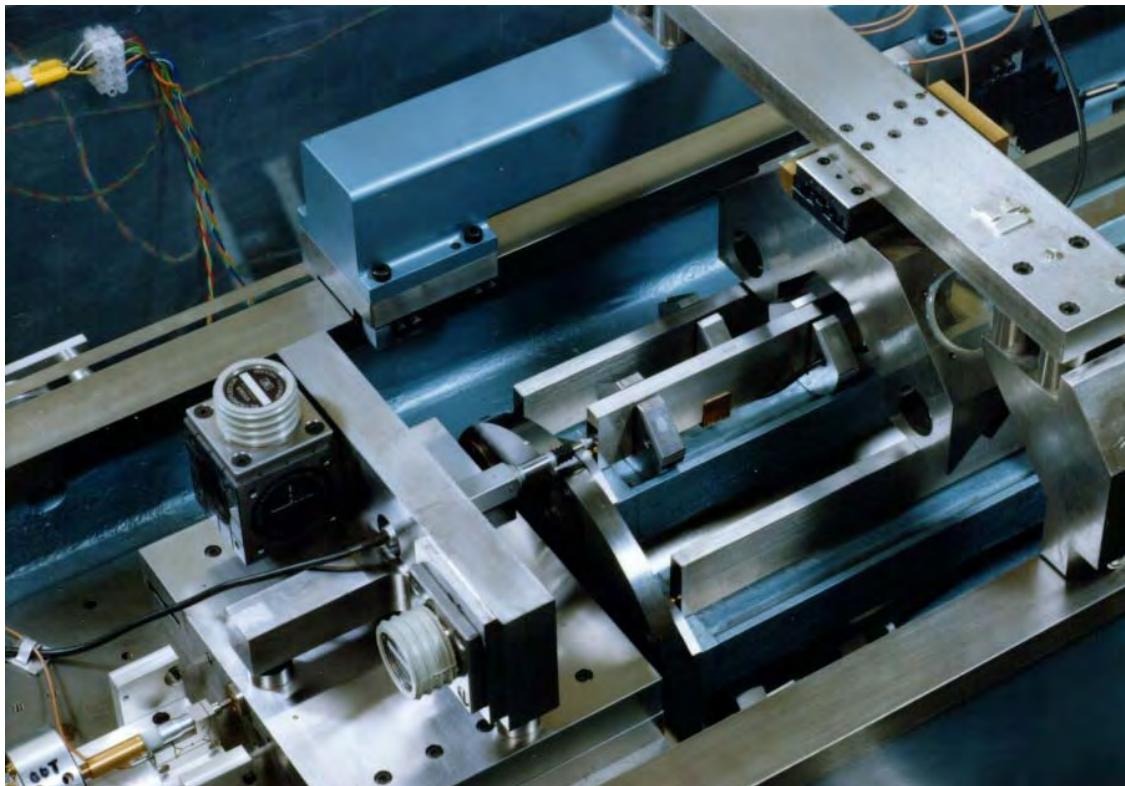


Figure 1.6 - NPL Length Bar Machine: probes contacting a long series gauge block

1.4.7 The rationale behind the development of the new interferometer

Emerging customer requirements showed a need for increased accuracy in length bar calibrations, especially for verifying the performance of co-ordinate measuring machines (a £20 M per annum CMM manufacturing market for the UK in 1986 [23]). Also there was a need to provide accurate longer standards for the NPL Length Bar Machine (over 300 mm) with measured values of flatness, parallelism and thermal expansion coefficient, and also an independent technique for verifying the performance of the LBM.

It was decided that a new instrument should be constructed to overcome as many of the limitations of the Length Bar Machine as possible and to measure additional parameters of length bars including: thermal expansion coefficient, flatness of measurement faces, parallelism of faces, as well as providing a more accurate measurement of the length of the bars. It was decided that this new interferometer should be capable of measuring both length bars and long series gauge blocks (gauge blocks over 100 mm) between 100 mm and 1500 mm in length. The interferometer could be used either as a calibration instrument, or to provide traceable standards for use in the Length Bar Machine. This

new interferometer, the National Primary Length Bar Interferometer (NPLBI), is the subject of the research presented in this thesis. To put the importance of this work into context, the next section deals with the traceability chain of length measurements through the use of calibrated end standards in the UK.

1.4.8 Traceability of length bar length measurements

When length bars are measured in terms of the primary standard of length, *i.e.* the wavelength emitted by the iodine-stabilised laser, they can then be used to calibrate the lengths of other standards through comparison or can be used to verify the performance of length measuring instruments. This hierarchical system of standards represents the *traceability* of length measurements - in theory any measurement of length can be traced to the definition of the metre. By nature of the hierarchical structure and the loss of accuracy at each comparison stage, it is obvious that the higher in the pyramid, the more accurate must be the standard, with the primary standard being the most accurate (see figure 1.7).

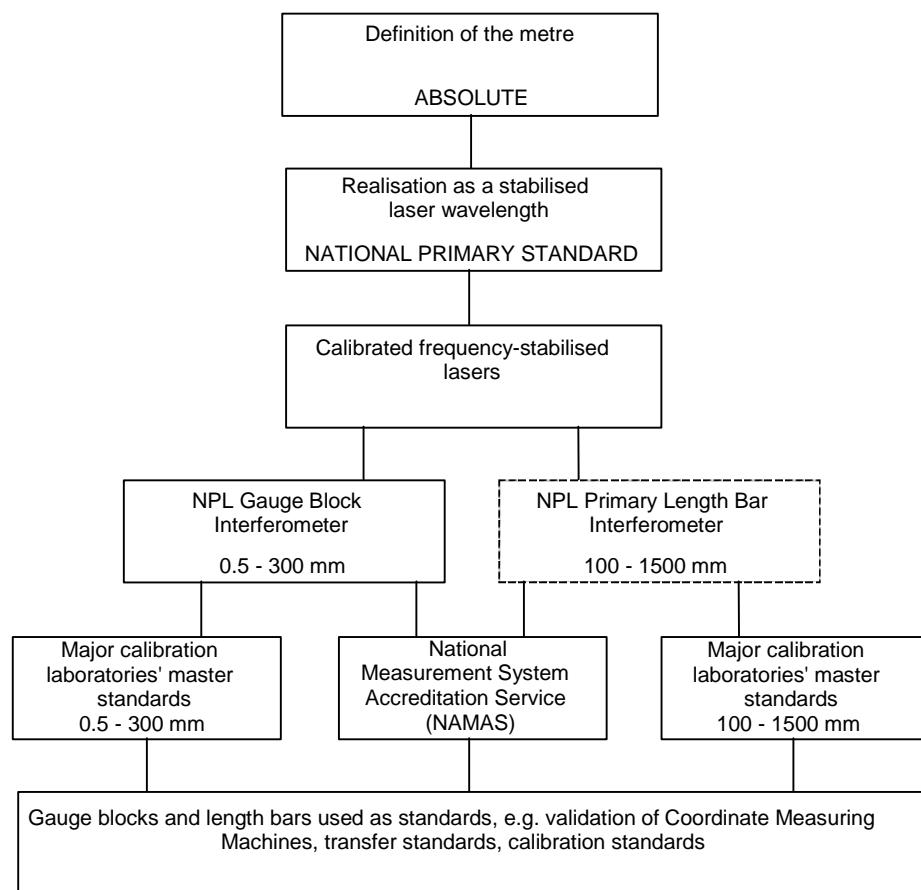


Figure 1.7 - Traceability of length measurements

At the head of the traceability chain is the definition of the metre. This is an absolute standard and is part of the SI system of units (see earlier). The metre is realised at NPL as a wavelength of a helium-neon laser stabilised to a saturated absorption in iodine at 632.991 398 1 nm. Details of the operation of this laser are given in § 3.2.1.3. A recent intercomparison of iodine-stabilised lasers has shown agreement at the level of 2×10^{-11} between lasers and has resulted in a new uncertainty being adopted for the UK realisation of the metre of $\pm 2.5 \times 10^{-11}$.



Figure 1.8 - Iodine-stabilised He-Ne Primary laser

As the Primary laser is the UK's Primary length standard, it follows that it is not used for routine calibration of end standards via interferometry. Instead, commercial stabilised lasers (based on an NPL design) are used after having their wavelengths calibrated against the Primary laser. These lasers are used in both the existing Gauge Block Interferometer and the National Primary Length Bar Interferometer.

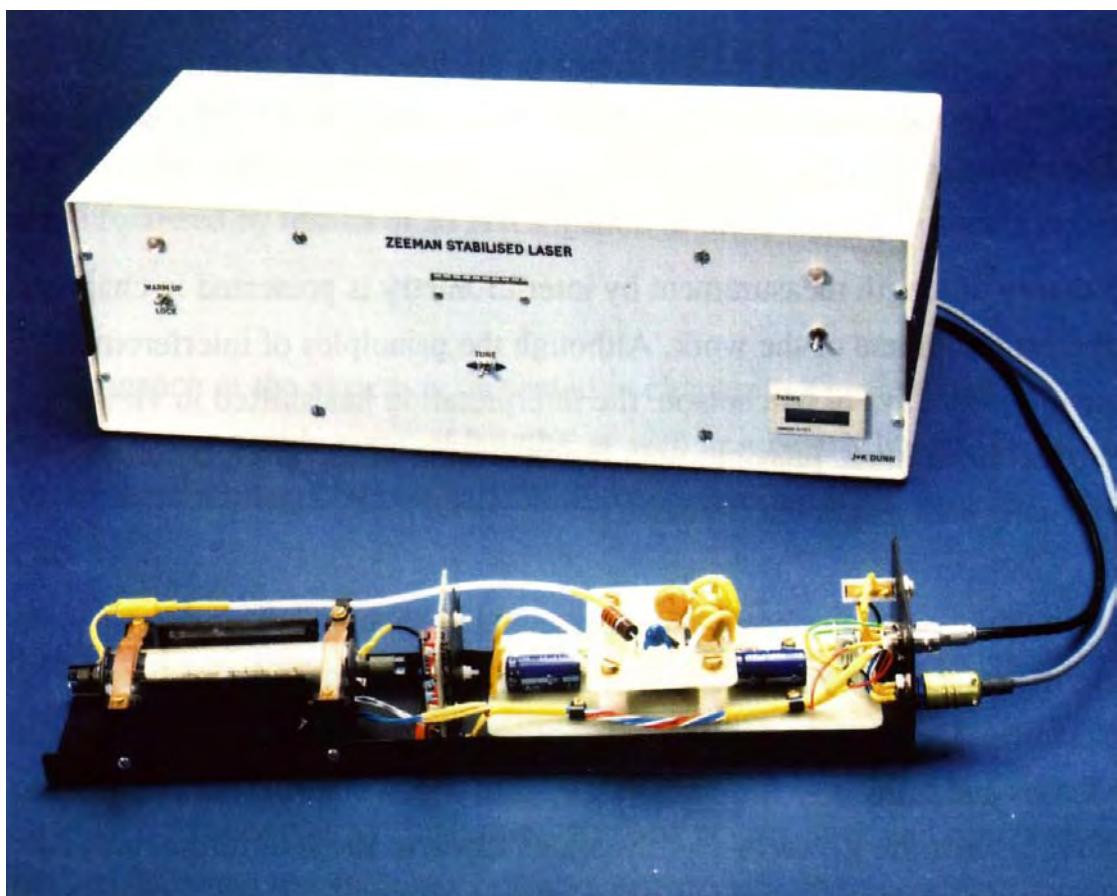


Figure 1.9 - NPL design Zeeman stabilised laser

Annually, NPL measures the length of some 200 length bars and 1200 gauge blocks for customers as well as providing a measurement and audit service for laboratories accredited under NAMAS (National Measurement Accreditation Service). In turn, these calibrated length standards are disseminated throughout UK industry where they are used to demonstrate traceability of length measurement for billions of pounds worth of trade.

As well as demonstrating national traceability, frequent European intercomparisons ensure that different countries' measurement systems are compatible and demonstrate traceability of length measurement to the internationally agreed definition of the metre, detailed previously in this chapter. As an example, a recent EUROMET intercomparison of gauge block measurements showed agreement between a number of European standards labs to within each lab's uncertainty budget for gauge block calibrations. For NPL this meant that the difference between gauge block lengths measured by NPL and the mean of all of the results was less than 60 nm for a 100 mm gauge block.

1.5 CONTENTS OF THE THESIS

This thesis describes the work of the author in the design, construction, commissioning and testing of the National Primary Length Bar Interferometer, together with research into the theory of length standards and interferometry.

A brief theory of length measurement by interferometry is presented in chapter 2, as this forms the central theme of the work. Although the principles of interferometry have not changed since the days of Michelson, the interpretation has shifted in viewpoint. A new approach has been made possible with the invention of the laser, an almost ideal light source for long path length interferometry.

The design of the new interferometer is presented in chapter 3. This describes the overall design of the instrument, together with some more detailed aspects such as the optics and the mechanics. The instrument itself is a collection of many components ranging from simple lenses and mirrors to the computer system which controls the entire measurement procedure. Details are given of modifications made to the interferometer to allow it to make double-ended measurements of length bars without the need for wringing.

The techniques for aligning the interferometer are dealt with in chapter 4, together with the results of preliminary checks made on the system, and various known optical defects and discrepancies.

Chapter 5 examines the analysis of interference fringes with particular emphasis on phase-stepping techniques. The techniques of fringe skeletonisation, temporal heterodyning, spatial heterodyning, Fourier Transform, phase-locking and phase-stepping (or phase-shifting) interferometry are examined. Error sources are identified and the limitations and benefits of the techniques explained. Phase-stepping interferometry is selected as the most suitable for analysis of fringes in the Primary Length Bar Interferometer. The techniques of 3, 4 and 5-position phase-stepping interferometry are examined. The 5-position technique that is selected belongs to the set of “N+1 symmetrical” algorithms. A demonstration of the errors of the technique is given, followed by a description of the implementation of the algorithm in the interferometer.

The data analysis and information processing are covered in chapter 6, including the software algorithms for phase-unwrapping and multiple-wavelength analysis. Overviews are given of the hardware and software, together with some example results.

The topic of refractive index correction of the laser wavelengths is the subject of chapter 7. This includes the operation of an air refractometer used to verify the

performance of the empirical equation used to calculate the refractive index and an assessment of the stability of the refractive index within the chamber.

Chapter 8 deals with the thermal control of the interferometer chamber and the subject of thermal expansivity measurements. The design of the temperature control system is presented followed by results of an investigation of the temperature stability of the air and length bars.

The performance of the system is presented in chapter 9. This includes results of length measurement over a wide range of lengths, as well as repeatability and results of thermal expansion measurements. A comparison of measurements with other NPL equipment is reported.

An error analysis of the complete instrument is presented in chapter 10. This will form the basis of the uncertainty budget for the completed instrument.

The final chapter, chapter 11, forms an overview of the work and draws some general conclusions on length measurement by interferometry and some specific conclusions with regard to the National Primary Length Bar Interferometer.

Appendix A is a list of the optical and mechanical equipment used to construct the interferometer with details of the optical testing of the equipment in appendix B. Appendix C deals with the bending of bars supported horizontally due to their own weight and the compensation of the platen's weight. Appendix D is a calculation of the compression of length bars when standing vertically. Appendix E is a list of the electrical connections and connectors. Copies of papers published by the author are collected in appendix F.

REFERENCES FOR CHAPTER 1

- [1] Bailey A E Units and standards of measurement *J. Phys. E: Sci. Instrum.* **15** (1982) 849-855
 - [2] Michelson A A & Benoit J R Détermination expérimentale de la valeur du mètre en longueurs d'ondes lumineuses *Trav. et Mem. BIPM* **11** (1895) 1
 - [3] Barrell H The Metre *Contemp. Phys.* **3** (1962) 415-434
 - [4] Giacomo P The new definition of the metre *Am. J. Phys.* **52** (1984) 607-613
 - [5] Kaye G W C & Laby T H *Tables of Physical and Chemical Constants* 15th edn (1989) (Harlow, Essex: Longmans) 1-13
 - [6] The International System of Units (1986) (translation of a BIPM publication), HMSO ISBN 0 11 887527 2
 - [7] Petley B W New definition of the metre *Nature* **303** (1983) 373-376
 - [8] Rowley W R C The definition of the metre: from polar quadrant to the speed of light *Phys. Bull.* **35** (1984) 282-284
 - [9] Terrien J International agreement on the value of the velocity of light *Metrologia* **10** (1974) 3
 - [10] Franks A Nanometric surface metrology at the National Physical Laboratory *Metrologia* **28** (1992) 471-482
 - [11] Kunzmann H Nanometrology at the PTB *Metrologia* **28** (1992) 443-453
 - [12] Petley B W The role of the fundamental constants of physics in metrology *Metrologia* **29** (1992) 95-112
 - [13] Klein H A, Bell A S, Barwood G P & Gill P Laser cooling of trapped Yb⁺ *Appl. Phys. B* **50** (1990) 13-17
 - [14] Gibble K & Chu S Future slow-atom frequency standards *Metrologia* **29** (1992) 201-212
 - [15] Peggs G N Creating a standards infrastructure for co-ordinate measurement technology in the UK *Ann. CIRP* **38** (1989) 521-523
 - [16] British Standard BS 888 (1950) (London: British Standards Institution)
 - [17] British Standard BS 4311 (1968 & 1993) (London: British Standards Institution)
 - [18] British Standard BS 5317 (1976) (London: British Standards Institution)
 - [19] British Standard BS 1790 (1961) (London: British Standards Institution)
 - [20] International Standard ISO 3650 (1978) (Geneva: International Organisation for Standardisation)
 - [21] German Standard DIN 861 (1980) (Berlin: Deutsches Institut für Normung e.V.)
 - [22] Pugh D J & Jackson K Automatic gauge block measurement using multiple wavelength interferometry *Proc. SPIE* **656** (1986) 244-249
 - [23] The Financial Times (London) 3 February 1986, page V
-

CHAPTER 2

REVIEW OF LENGTH MEASURING INTERFEROMETERS

“Lux, etsi per immunda transeat, non inquinatur.”
(“Light, even though it passes through pollution, is not polluted.”)
St. Augustine

2.1 THEORY OF INTERFEROMETRIC LENGTH MEASUREMENT

The basic principles of length measurement using interferometry were demonstrated in the reverse sense by Michelson in 1892 when he measured the wavelength of the red light from cadmium in terms of the International Prototype Metre.

The principle of length measurement by interferometry is straightforward - it is the comparison of a mechanical length (or a distance in space) against a known wavelength of light. This may be expressed in a simple equation

$$L = (N + f)\lambda \quad (2.1)$$

where L is the length being measured, λ is the wavelength, N is an integer and f is a fraction ($0 < f < 1$). Commonly the optics are arranged such that the light beam measures exactly double the required length (*i.e.* it is a double-pass system), in which case the measurement units, ‘fringes’, are half-wavelengths.

$$L = (N + f) \frac{\lambda}{2} \quad (2.2)$$

i.e. one interference fringe corresponds to a distance or length equal to $\lambda / 2$. By using a light source for which λ is known (*e.g.* laser, gas discharge lamp), measurement of N and f leads directly to a value for L .

When using sources of visible light the wavelength of the light is small, typically 400 - 700 nm and hence the basic ‘unit’ of measurement, one fringe, is 200 - 350 nm in size. Hence knowledge of N alone provides a measurement resolution of 200 - 350 nm. For

comparison, a typical atomic spacing is of the order of 0.5 nm, so the tick marks on an ‘interferometric ruler’ are spaced 400 - 700 atoms apart.

By careful analysis of these interference fringes, it is possible to sub-divide them (and hence measure f) to a resolution of 1/100 to 1/1000 of a fringe - the minor tick marks on the ‘ruler’ are represented by single atoms.

By using sources with much smaller wavelengths (*e.g.* x-rays), the size of each fringe can be reduced to 0.2 nm, and fringe sub-division yields a measurement resolution in the picometre range [1]. Alternatively, the complications of x-ray optics and fringe detection may be avoided by using a ‘many-pass’ arrangement of optics in which the measurement beam traverses the object length many times, by multiple reflection from slightly tilted mirrors. (This should not be confused with multiple-beam interferometers such as the Fabry-Perot design. In the former, each point in the interference pattern is the summation of two beams, which have travelled in many passes, whereas in the latter (Fabry-Perot), each point is the summation of many beams, which have travelled different path lengths - see § 2.4.2).

2.2 BASIC INTERFEROMETER TYPES

There are many types of interferometer using a variety of techniques, mostly using lasers as their light source. Each design of interferometer is suited to a particular situation and has certain advantages and disadvantages for interferometry of length bars. These will now be examined.

The first distinction that can be made between interferometer types is whether they are dynamic or static. Dynamic interferometers are usually fringe counting interferometers, often with a small beam diameter. Static interferometer designs, on the other hand, are often large field, and are typically used for optical testing.

2.3 REVIEW OF SMALL FIELD, DYNAMIC INTERFEROMETERS

2.3.1 Fringe counting interferometry

In the simplest type of fringe counting interferometer, one of the mirrors is moved whilst the other remains stationary. As the interference fringes cross a detector, typically a photodiode, the number of maxima or minima in the signal are counted by a simple circuit. No attempt is made to sub-divide the fringe count. The resolution of this

type of interferometer is thus limited to the size of one interference fringe, typically 300 nm.

In fringe counting interferometers with interpolation, the output beam of a laser is split into two. After travelling along the reference and measurement paths the beams interfere with each other and are split again. One beam is given an extra phase difference of $\pi/4$ to produce two orthogonal outputs for example by using a quarter-wave plate, as in figure 2.1, or by using a specially coated beamsplitter [2]. Two detectors view these two outputs which are in phase quadrature and the signals from these two detectors, after suitable processing, are used to drive a bi-directional counter. By using a fringe interpolator using look-up tables or computer processing, each fringe can be sub-divided to the 1/1000 fringe level to provide the potential for more accurate measurement.

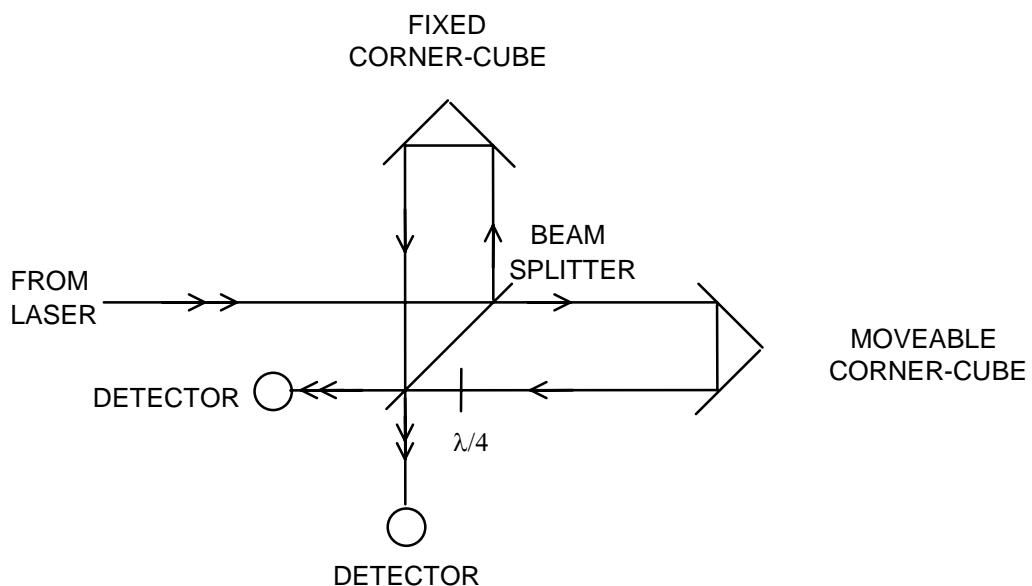


Figure 2.1 - Schematic diagram of a fringe counting interferometer

The main disadvantage of using a fringe counting system is that it provides only relative measurements of distance and requires accurate setting at two defined points to define a length. In the measurement of absolute lengths, the optical path difference of the fringe counting interferometer must be increased by a distance equal to the length of the object to be measured. This can be done by traversing a mirror on a linear stage between two points coincident with the two ends of the object to be measured, see figure 2.2, or by using the interferometer to monitor the position of the probe of a co-ordinate measuring machine.

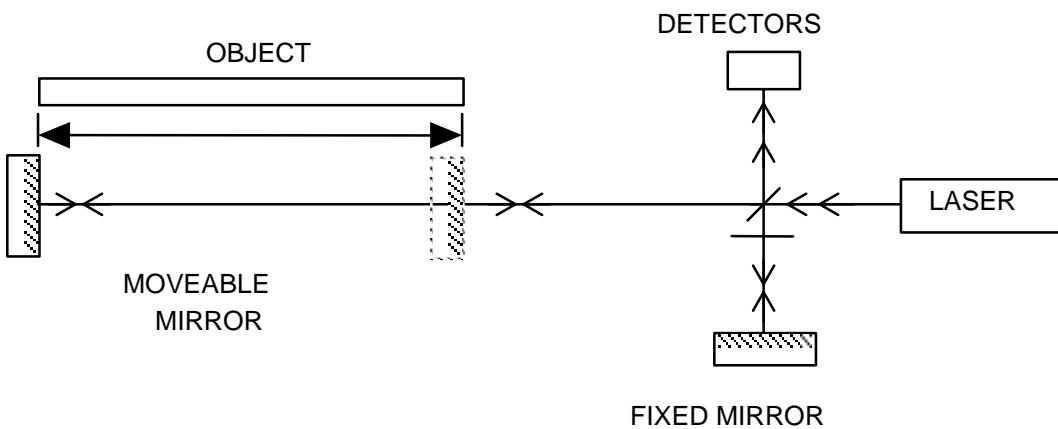


Figure 2.2 - Use of fringe counting to measure the length of an object

This system is then prone to alignment errors. There are three axes which must be coincident: the central axis of the object, the axis of the linear stage and the axis of the interferometer beam. Any mis-alignment of one of these will lead to a length dependent error.

The fringe interpolation is also prone to errors. If the amplitudes of the two orthogonal signals are not equal or their phase difference is not exactly $\pi/2$, then the interpolation will be incorrect. Table 2.1 gives typical values for fringe interpolations errors for a commercial fringe counter with 1/1000 fringe interpolation. Use of a computer curve fitting algorithm to fit an ellipse to the intensity data can sometimes be used to increase the accuracy [3].

Parameter	Error in parameter	Error in fringe fraction
Gain mis-match	5 %	0.004
DC offset	10 mV	0.001
Phase difference	5 °	0.010

Table 2.1 - Typical fringe interpolation errors (FT Technologies FT612AS)

The speed of measurement is also important to the overall accuracy. Any fringe counting system which relies on the translation of a moveable mirror or mechanical probe will require a finite amount of time for the probe to move between the two ends of the object. The speed of movement is not usually limited by the ability of the fringe counter to track the fringes as the probe is moved since fringe counting rates up to 20 MHz may be easily achieved, corresponding to a velocity of 6 m s^{-1} . However it takes a few seconds to accelerate and decelerate the probe and mount. With a settling time of a few seconds before the fringe interpolation can be made accurately this requires a total

of approximately 10 seconds during which any movement in the bar or optical components, any drift in refractive index or thermal expansion of the bar will cause an error in the measurement.

2.3.2 Heterodyne fringe counting interferometry

The field of heterodyne interferometry has been made accessible due to the very narrow linewidths of lasers which allows the observation of beats between two laser frequencies. The technique has resulted in a commercial design of interferometer (Hewlett-Packard 5528A) which can be used for distance measurement of path lengths up to 60 metres [4]. The instrument contains a He-Ne laser in which a longitudinal magnetic field splits the output beam into two components (Zeeman splitting) which are separated in frequency by 2 MHz. These two components have opposite circular polarisations which are converted into orthogonal linear polarisations after passage through a quarter wave plate. A polarising beamsplitter directs one polarisation to a fixed corner-cube which acts as a reference whilst the other beam is directed to the moveable corner-cube, see figure 2.3. The beat frequency between the two beams is detected and combined with a reference frequency in a differential counter. When the mirrors are stationary these two frequencies are equal, however if one of the corner-cubes is moved there is a frequency difference signal which can be used to monitor the change in path length between the two arms.

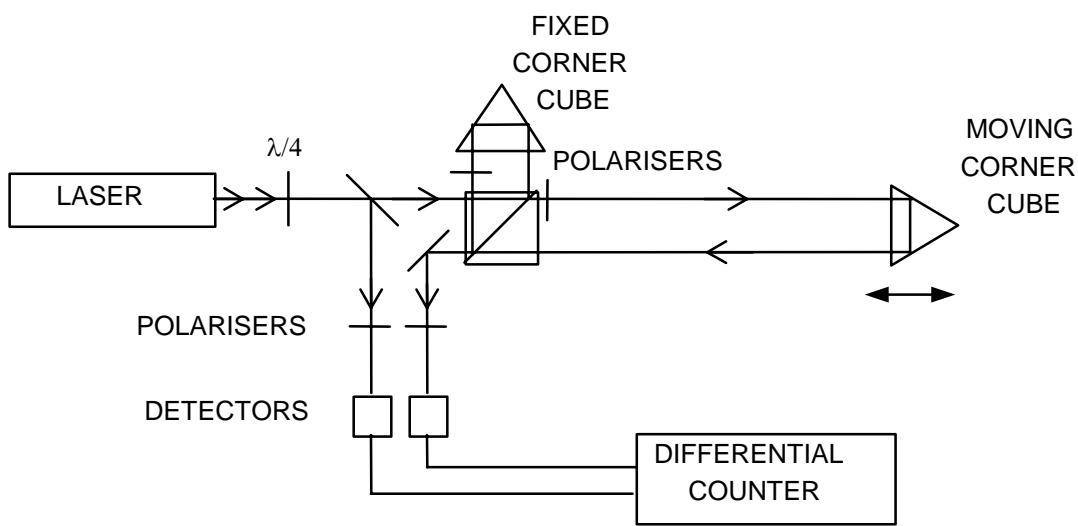


Figure 2.3 - Schematic diagram of a heterodyne fringe counter

The conventional arrangement of the Hewlett-Packard system offers a resolution of 10 nm, though this can be improved by using multiple-pass interferometry where the beam traverses the measurement path many times.

The drawbacks with this system are similar to those of the simple fringe counting system above. This type of interferometer is used in many industrial calibration laboratories where the wavelength of the laser is calibrated to provide traceability.

2.3.3 Two-wavelength fringe counting interferometry

By illuminating an interferometer sequentially with two wavelengths λ_1 and λ_2 , the effective range is that which would be obtained with an effective wavelength λ_e

$$\lambda_e = \frac{\lambda_1 \lambda_2}{|\lambda_1 - \lambda_2|} \quad (2.3)$$

Measurements have been made with carbon dioxide lasers [5] where the laser output wavelength is switched rapidly between two wavelengths as one of the mirrors of the interferometer is moved. Matsumoto [6] has investigated measurement of distances up to 100 m to an accuracy of 1 part in 10^7 . Error sources for this technique include the alignment of the paths, the accuracy of the laser wavelengths (including refractive index effects) and the accuracy of the fringe interpolation performed in the computer.

2.3.4 Other fringe counting systems

Kubota [7] employed a frequency modulating interferometer using semiconductor lasers where polarised beams were combined to produce fringe patterns with intensities in phase quadrature. The absolute measurement of distance was performed by sweeping the frequency of the laser by ramping the supply current. Detection of the beat frequency provided information about the path difference between the two arms of the interferometer. The accuracy of the system is limited by the accuracy of the wavelength measurement whilst sweeping.

Other authors have used direct modulation of the laser output intensity using feedback. In this scheme the moveable mirror reflects the beam back into the laser cavity and thus acts as an extended external cavity [8,9,10]. The drawback with this scheme is that the feedback affects the ability of the laser to stabilise at a single mode and hence the absolute value of the laser wavelength cannot be maintained.

In order to overcome the limitations described above, a system for the measurement of length bars must be static, *i.e.* the measurements must be performed with little or no

movement of the optical components or mechanical parts of the system including the object or length bar to be measured. A drawback of using a dynamic fringe counting interferometer is that it does not directly measure the positions of the end faces of the bar, rather it measures the positions of the interferometer mirrors, which require accurate setting at the opposite ends of the bar. Better accuracy can be achieved by measuring the bar directly and using its flat polished end faces as the optical mirrors in the interferometer.

The system must therefore be able to detect, image or otherwise measure the positions of the two ends of the bar simultaneously. The laser source must operate at an accurately known wavelength such that the distance between the ends of the bar can be measured in terms of interference fringes of a known, fixed size. The interference pattern generated by the interferometer must be amenable to analysis to allow fringe sub-division to the nanometre level. These conditions are satisfied by using a large aperture interferometer. The effect of laser beam diffraction is also minimised by increasing the aperture of the interferometer.

2.3.5 The effect of laser beam diffraction on measured length

Due to the spread of wavevectors in a gaussian laser beam, there is a slight alteration to the longitudinal propagation speed of parts of the wavefront. This can be estimated as follows.

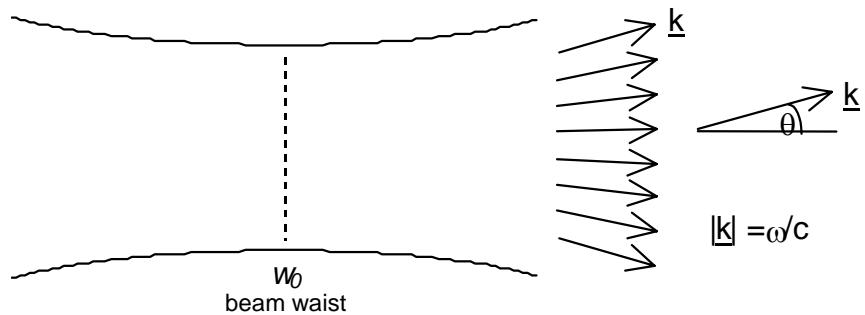


Figure 2.4 - Laser beam waist - alteration of effective propagation speed

The effective propagation speed is reduced by

$$\approx |\underline{k}|(1 - \cos \theta) \quad (2.4)$$

with

$$\theta \approx \frac{\lambda}{w_0} \quad (2.5)$$

where λ is the wavelength and w_0 is the minimum beam diameter (waist). Therefore the correction to the measured distance is approximately

$$\frac{\theta^2}{2} \quad \text{or} \quad \frac{\lambda^2}{2w_0^2}$$

For an expanded laser beam used in an interferometer with aperture 80 mm, this evaluates to $\sim 3 \times 10^{-11}$, and is thus negligible. A more detailed derivation is given by Rowley [11] and is reproduced here, to confirm the approximate estimate. The combined effect of the wavefront curvature and the propagation phase shift on D , the distance measured by the interferometer, is given by

$$D' = D - \frac{\lambda}{2\pi} \arctan \left(\frac{D\lambda}{2\pi w_0^2} \right) \quad (2.6)$$

$$D' \approx D \left(1 - \frac{\lambda^2}{4\pi^2 w_0^2} \right) \quad (2.7)$$

Substituting $\lambda = 633$ nm, $w_0 = 80$ mm (beam diameter), the result is

$$D' = D(1 - 1.6 \times 10^{-12})$$

i.e. negligible. This is the same result as that of Dorenwendt and Bönsch [12]. Mana [13] obtained the result

$$\frac{\Delta\lambda}{\lambda} = \frac{\lambda^2}{2} \frac{\pi^2}{w^2} \quad (2.8)$$

which is approximately 10^{-10} . Thus by using an expanded beam in the interferometer, rather than the conventional unexpanded laser beam size of 1 - 2 mm diameter, diffraction effects due to the beam waist can be made negligible.

2.4 REVIEW OF LARGE-FIELD INTERFEROMETER DESIGNS

2.4.1 The Fizeau interferometer

The most common large field interferometer is the Fizeau interferometer. This requires the minimum of optical components: a light source, collimating lens, reference flat and test surface, see figure 2.5. The source and return beams are coincident unless the

source is positioned slightly off-axis which then leads to an obliquity effect (see § 4.1.3). The profile of the fringes in a Fizeau interferometer with tilt is determined by the reflection and transmission coefficients of the optical surfaces used.

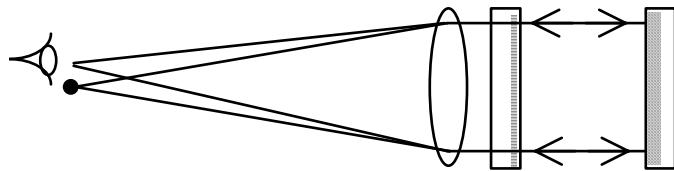


Figure 2.5 - Standard Fizeau interferometer

To overcome the problem of separating the source and return beams it is common to use an extra beamsplitter in the collimator path to re-direct the output beam, figure 2.6.

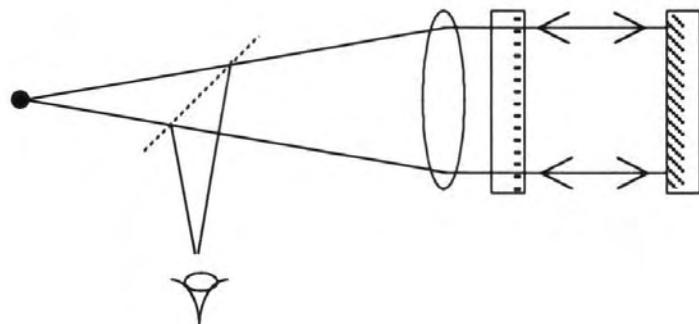


Figure 2.6 - Modification to Fizeau interferometer

Care has to be taken to minimise extra reflections from the beamsplitter secondary surface, *e.g.* by using an anti-reflection coating. Typical fringe intensity profiles can be seen in figure 2.7. These were calculated from equation 2.9, for values of R from 0.1 to 0.9, where $R = R_1R_2$, and $R_1 = R_2$.

$$I = 1 - \frac{(1-R_1)^2(1-R_2)^2}{1+(R_1R_2)^2 - 2R_1R_2 \cos(2m\pi)} \quad (2.9)$$

R_1 and R_2 are the reflectivities of the two surfaces and m is the order of interference.

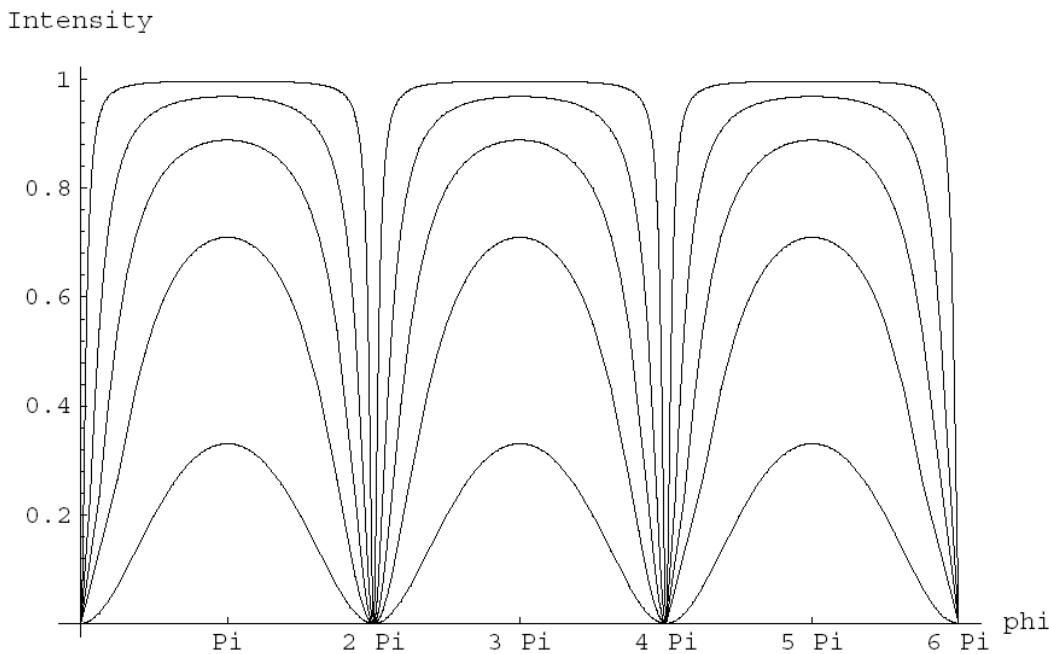


Figure 2.7 - Normalised intensity profiles of fringes in a Fizeau interferometer for values of the reflectivity (R) ranging from 0.1 (lowest curve) to 0.9 (uppermost curve), according to equation 2.9

2.4.2 The Fabry-Perot interferometer

In fact the Fizeau interferometer may be regarded as a version of the Fabry-Perot interferometer which uses multiple reflection to achieve narrow profile fringes especially suited to spectroscopic work.

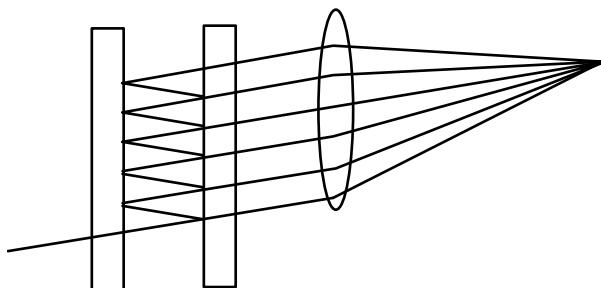


Figure 2.8 - Fabry-Perot interferometer

2.4.3 The Michelson interferometer

By introducing the extra beamsplitter to the Fizeau interferometer, the number of optical components has equalled that found in another common interferometer, the Michelson interferometer. Here a beamsplitter at 45° is used to split the incoming beam into two (amplitude division) which are directed to two mirrors: the reference mirror and the test mirror, and then recombined.

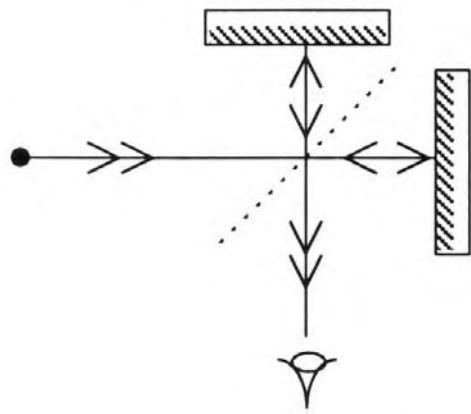


Figure 2.9 - Michelson interferometer

The Michelson interferometer produces different types of fringes according to the path difference between the two arms and whether or not any tilt has been introduced. It is common to use a compensating plate in the Michelson interferometer to account for the path difference between the two beams since one of them passes through the glass of the beamsplitter 3 times whereas the other passes through only once.

2.4.4 The Twyman-Green interferometer

When the Michelson interferometer is used with a collimated beam of light the arrangement is called the Twyman-Green interferometer [14].

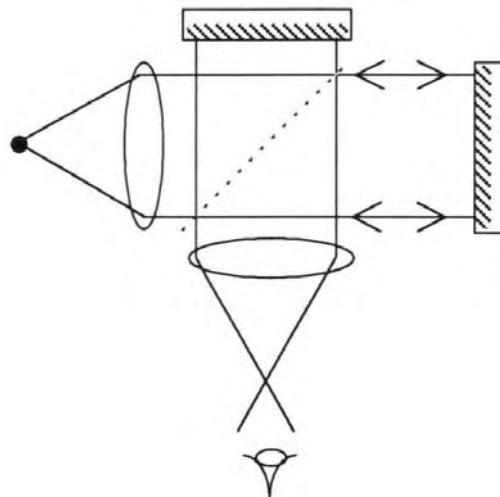


Figure 2.10 - Twyman-Green interferometer

The Twyman-Green interferometer is ideal for measurements of surface shape and surface texture as it offers a large field with sinusoidal fringes of good contrast and has the added advantage that by careful positioning of the reference mirror, twice the coherence range of the Fizeau interferometer can be obtained when imaging extended

objects, *i.e.* by positioning the reference mirror at an optical path equal to half the length of the extended object to be imaged, the useable coherence is doubled, see figure 2.11.

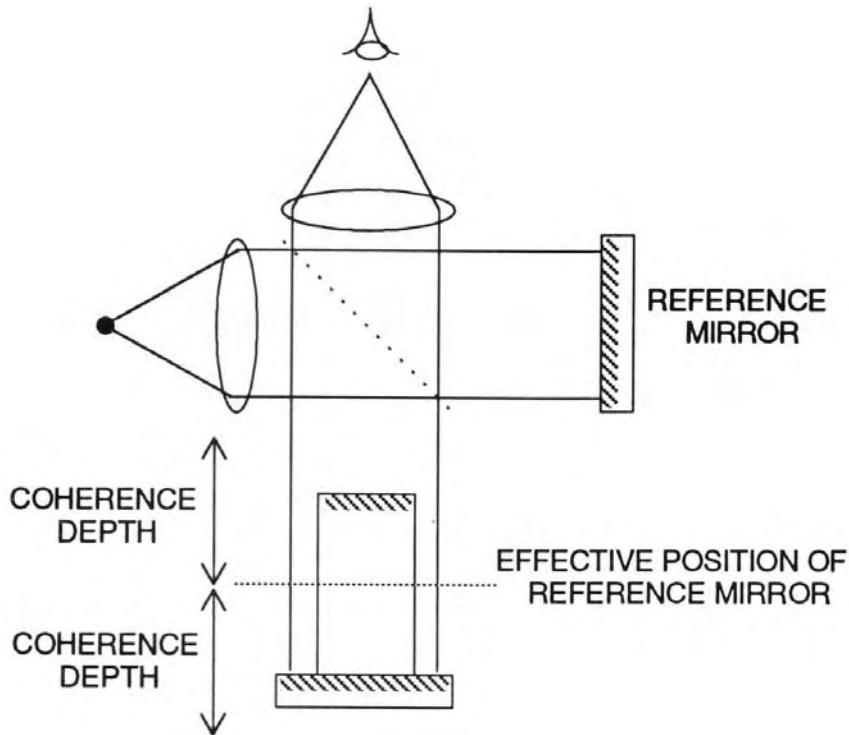


Figure 2.11 - Coherence depth doubling by positioning of reference mirror

Thus for a source of coherence length 50 mm, by using a Twyman-Green arrangement as opposed to a Fizeau system, objects up to 100 mm can be measured.

Whilst the narrow fringes of the Fizeau/Fabry-Perot interferometer are ideal for fringe locating and tracking algorithms, the sinusoidal nature of the fringes in a Twyman-Green interferometer are ideal for analysis by phase-stepping techniques which offer potentially greater accuracy of fringe interpolation approaching 1/100 to 1/1000 fringe (see chapter 5). Although it is possible to use phase-stepping with a Fizeau interferometer (see § 5.3.7.3), the analysis is more complex and requires symmetrical fringe profiles.

2.4.5 Other designs of length measuring interferometer

Previous interferometers for the measurement of end standards of length have included the NPL Automatic Gauge Block Interferometer (Twyman-Green), NPL-Hilger interferometer (Fizeau) and the Kösters-Zeiss interference comparator which used a precision Kösters prism as the beamsplitter/combiner, shown in figure 2.12.

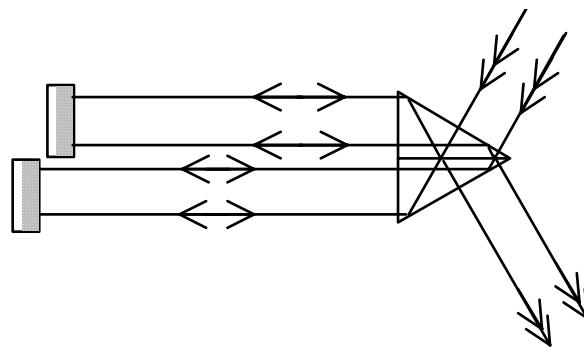


Figure 2.12 - Kösters-Zeiss interference comparator

2.5 PRIMARY LENGTH BAR INTERFEROMETER - BASIC INTERFEROMETER TYPE

Of the designs investigated, the Twyman-Green design was chosen as the basis for the Primary Length Bar Interferometer. The large field makes it possible to view interference fringes over the whole of the surface of the bar and hence measure surface form. The sinusoidal fringe intensity is ideal for analysis by phase-stepping interferometry (see chapter 5), the alignment is straightforward and the coherence-depth sufficient (given a relatively small source).

To measure the length of the bar, it is attached to a reference flat or platen, as shown in figure 2.13.

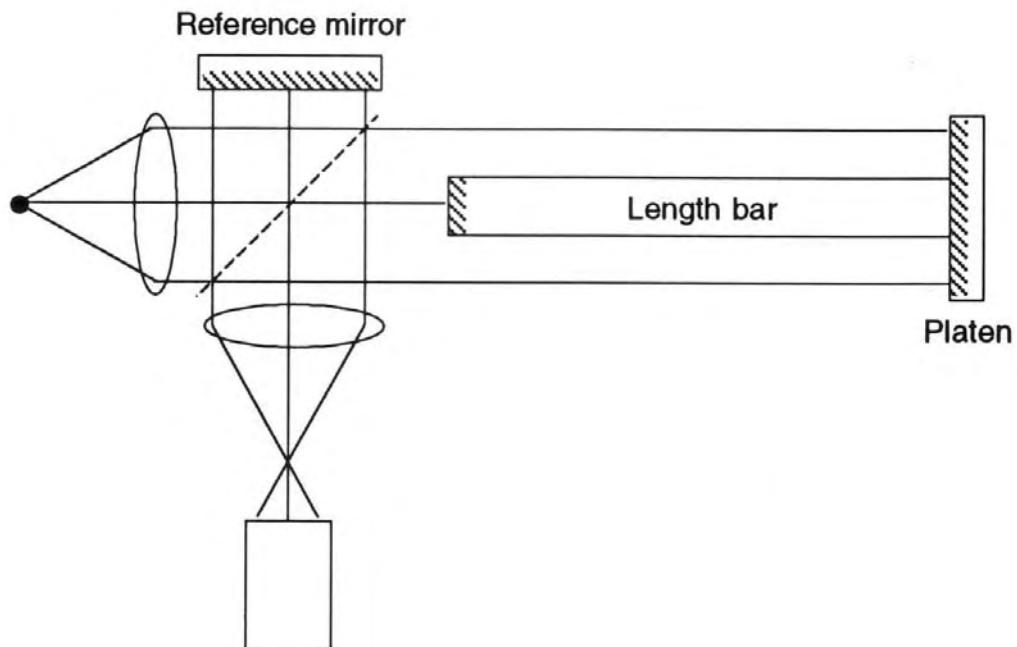


Figure 2.13 - Twyman-Green interferometer for the measurement of length bars

2.5.1 The technique of ‘wringing’

The platen is attached to the bar by the process of ‘wringing’. In this process the surfaces of the bar and platen are thoroughly cleaned using acetone to remove all presence of oils, dirt and dust. A small drop of a solution of ‘wringing fluid’ (liquid paraffin diluted 1:50 in Ark lone (1,1,2 Trichloro-1,2,2 Trifluoro-Ethane)), is smeared over the surface of the platen and polished off with a tissue until none can be seen by the naked eye. The platen is then slid onto the end of the bar whilst the two are in plane contact. During this process, the bar rests on temporary supports to allow access to the end of the bar.

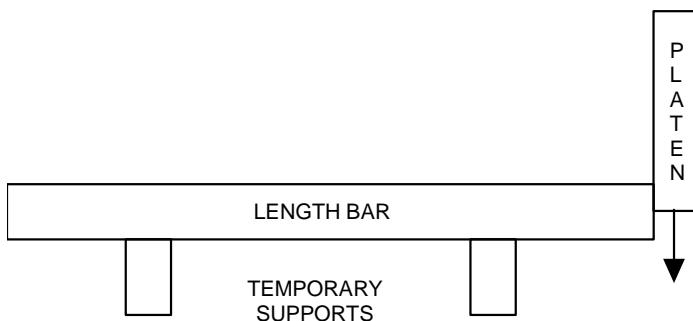


Figure 2.14 - Wringing a platen onto the end of a length bar

If the surfaces of the bar and platen are sufficiently flat and polished/lapped, molecular attraction takes place between the two surfaces and the molecular layer of paraffin between them. This attraction is strong enough to support the weight of the platen. Previous work [15] has shown that the ‘wringing thickness’ - the apparent separation between the two surfaces, is typically 7 nm (± 5 nm), and is not dependent on the amount of wringing fluid or its composition. The wringing was also proved to be a definite attraction rather than the effects of air pressure holding the two surfaces together (wringing can take place in a vacuum). The effects of surface tension have also been shown to be minimal in this situation. The wringing is much stronger when a wringing fluid is used on slightly rough surfaces, *i.e.* surfaces which are macroscopically flat, but not highly polished. It is thought that this is due to penetration of the wringing fluid into the sub-microscopic interstices of the platen leading to a larger surface area in good contact with the bar.

The size of the wringing film layer is taken into account when defining the length represented by a length bar. BS 5317 states:

Length ... is defined, with the bar mounted horizontally and referred to the standard reference temperature of 20 °C ... as the distance from the centre of one of its faces to a flat surface in wringing contact with the opposite face, measured normal to the surface.

There is an advantage in using a wrung length standard in that the length can be represented by a step height which allows measurement by probing in one direction; the probes can be optical, capacitive, linear variable differential transformers (LVDTs) or touch-trigger probes of CMMs. The penetration of the probe into the surface of the bar and platen due to phase effects (optical probes) or material compression (mechanical probes) is then accounted for, if the platen and bar are of the same material. This is not true if the object is probed from different directions.

REFERENCES FOR CHAPTER 2

- [1] Franks A Nanometric surface metrology at the National Physical Laboratory *Metrologia* **28** (1992) 471-482
 - [2] Raine K W & Downs M J Beam-splitter coatings for producing phase quadrature interferometer outputs *Opt. Acta* **25** (1978) 549-558
 - [3] Birch K P Optical fringe subdivision with nanometric accuracy *Precis. Eng.* **12** (1990) 195-199
 - [4] Dukes J N & Gordon G B A two-hundred-foot yardstick with graduations every microinch *Hewlett-Packard J.* **21** (1970) 2-8
 - [5] Gillard C W & Buholz N E Progress in absolute distance interferometry *Opt. Eng.* **22** (1983) 348-353
 - [6] Matsumoto H Synthetic interferometric distance measuring system using a CO₂ laser *Appl. Opt.* **25** (1986) 493-498
 - [7] Kubota T, Nara M & Yoshimo T Interferometer for measuring displacement and distance *Opt. Lett.* **12** (1987) 310-312
 - [8] Ashby D E T F & Jephcott D F Measurement of plasma density using a gas laser as an infrared interferometer *Appl. Phys. Lett.* **3** (1963) 13-16
 - [9] Dandridge A, Miles R O & Giallorenzi T G Diode laser sensor *Electron. Lett.* **16** (1980) 948-949
 - [10] Yoshino T, Nara M, Mnatzakanian S, Lee B S & Strand T C Laser diode feedback interferometer for stabilization and displacement measurements *Appl. Opt.* **26** (1987) 892-897
 - [11] Rowley W R C Laser wavelength measurements and standards for the determination of length *Precis. Meas. & Fund. Constants, NBS Spec. Publ.* **617** (1984) 57-64
 - [12] Dorenwendt K & Bönsch G Über den Einfluß auf die interferentielle Längenmessung *Metrologia* **12** (1976) 57-60
 - [13] Mana G Diffraction effects in optical interferometers illuminated by laser sources *Metrologia* **26** (1989) 87-93
 - [14] Twyman F & Green A (1916) British Patent 103832
 - [15] Rolt F H *Gauges and Fine Measurement* (MacMillan & Co.: London) (1929)
-

CHAPTER 3

THE DESIGN OF THE INTERFEROMETER

“A whole is that which has beginning, middle and end.”
Aristotle

3.1 OPTICAL DESIGN OF THE INTERFEROMETER

3.1.1 Description of the interferometer

The optical design of the interferometer is shown in schematic form in figure 3.1. Details of the components may be found in Appendix A with results of optical testing in Appendix B. Appendix E contains details of the electronics equipment rack connections. Apart from the addition of path-folding mirrors, the design is that of a Twyman-Green interferometer. The collimator focal length is 1.5 m, the de-collimation length is 1.0 m. The overall size of the interferometer is approximately 1.2 x 2.4 x 0.3 m. The optical axis is 120 mm above the baseplate and the reference and measurement beams have a diameter of approximately 80 mm. Figure 3.2 shows the components of the interferometer with a perspective view of the optical layout in figure 3.3 with further details in figure 3.25.

The light sources for the interferometer are 3 frequency-stabilised helium-neon lasers operating at wavelengths of approximately 633 nm (red), 543 nm (green), and 612 nm (orange). The output of each laser is focused into a single mode optical fibre of 2.8 - 3.6 μm core diameter. The 3 fibres have their other ends terminated in a small ferrule and form a 3 spot light source. The small diameter of the fibre has two advantages: it decreases the effective angular diameter of the source at the collimator lens (see § 4.1.3) and it allows only one mode to propagate in the fibre, preventing the formation of large-scale speckle in the interference pattern [1]. The 3-fibre light source can also be used as an alignment aid when setting up the interferometer (see § 4.1.2.3).

The light exits the fibres and diverges with an angular spread which is dependent on the numerical aperture of the fibres, which is 0.12. The divergent beam from each fibre over-fills the collimator lens (diameter 100 mm) which subtends a numerical aperture of 0.03 at the source. The ends of the fibres are positioned at the focal point of the collimator lens, so the beam emerges from the lens fully collimated.

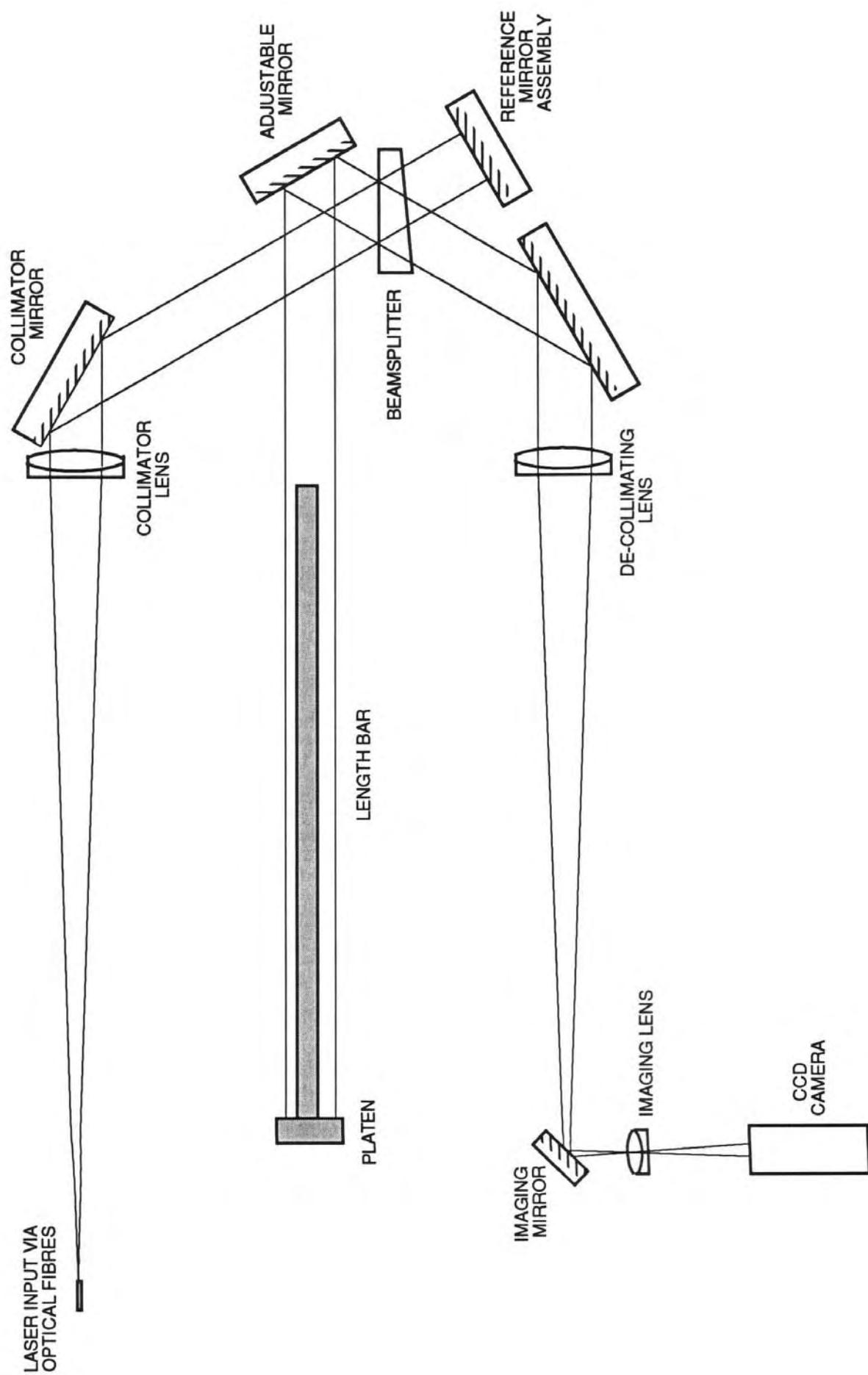


Figure 3.1 - Schematic design of interferometer optics

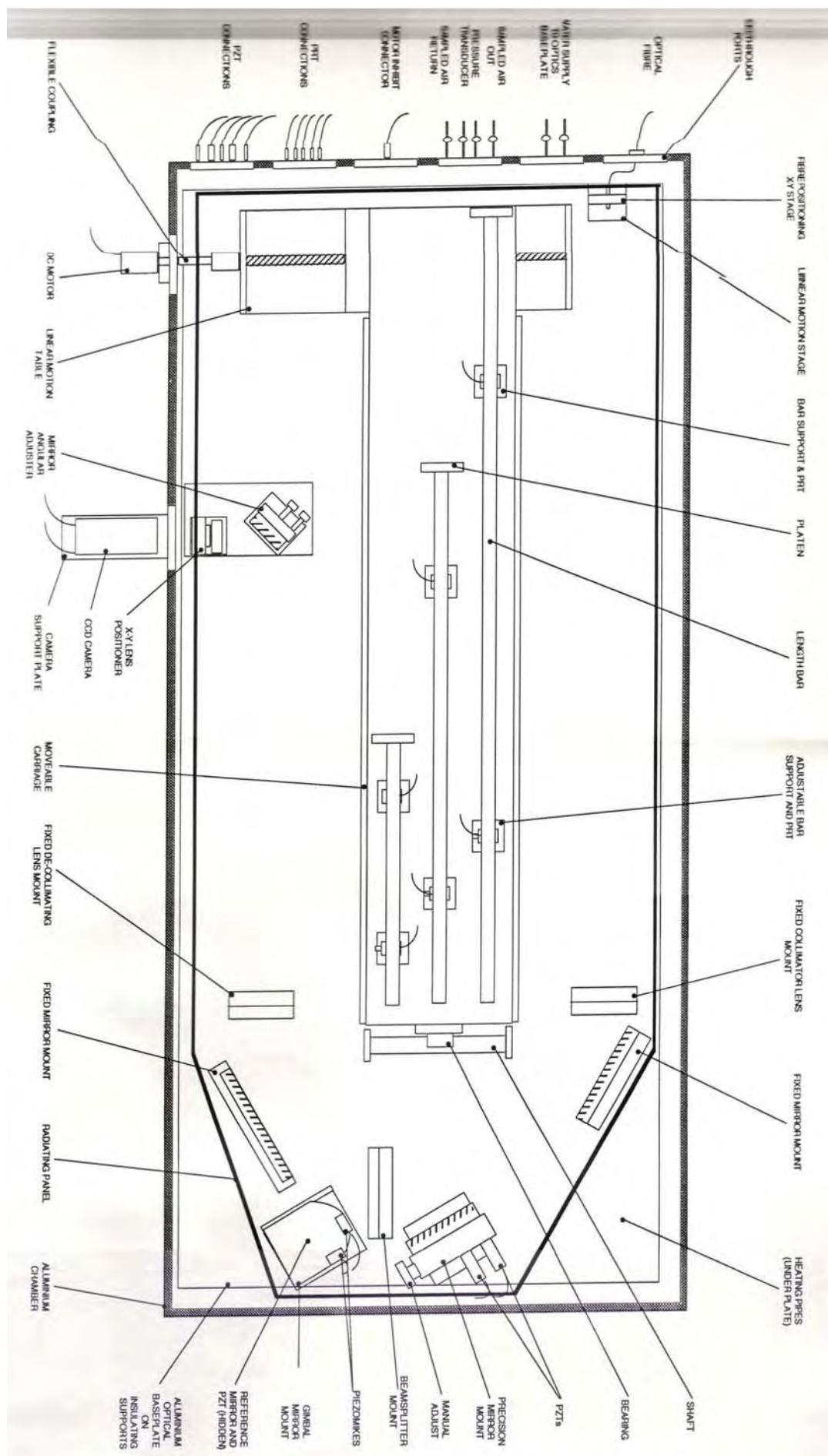


Figure 3.2 - Diagram showing the opto-mechanical components of the interferometer

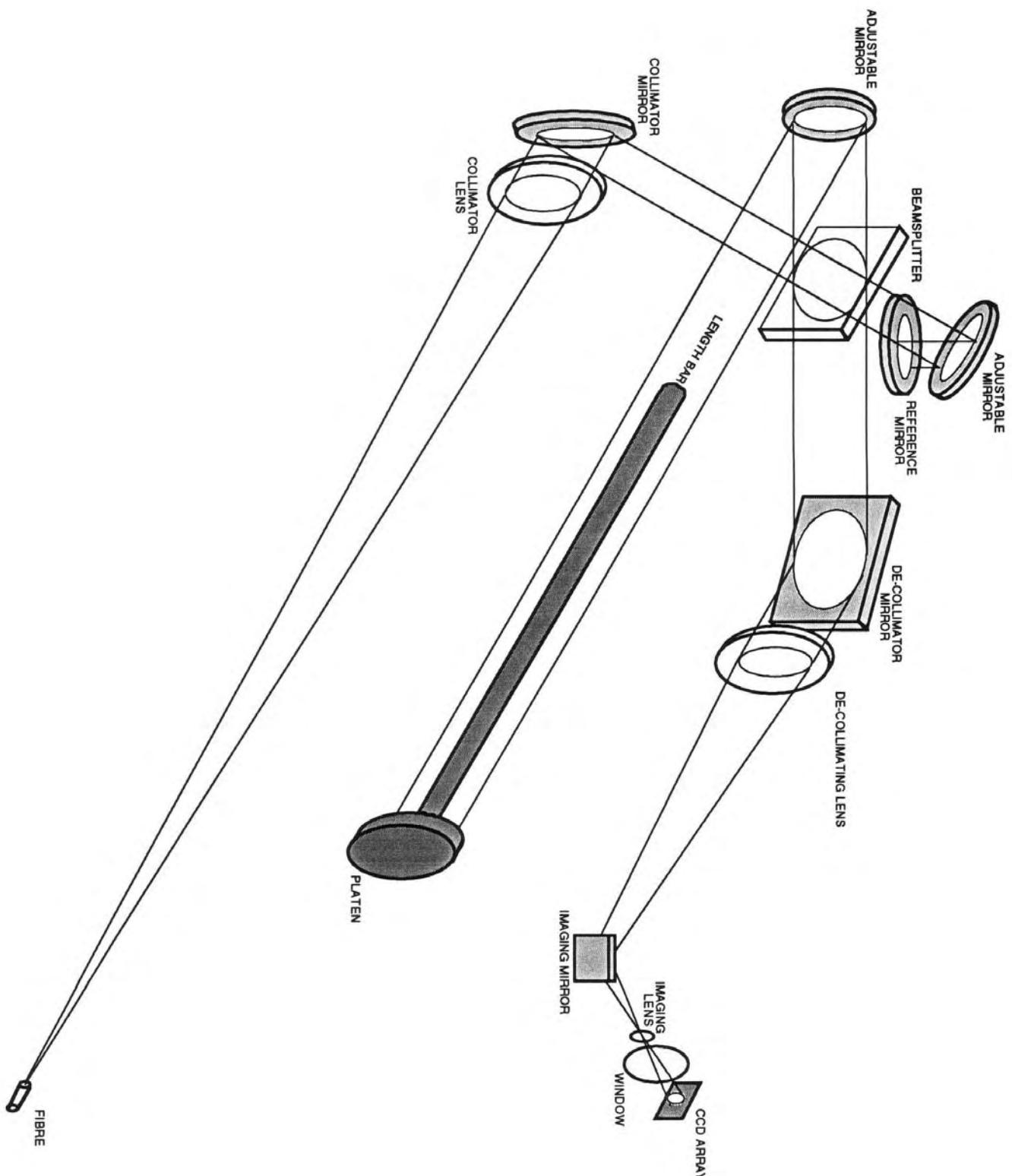


Figure 3.3 - Perspective view of optical components and directions of beams

Proper focusing of the collimator can be checked using either a shearing plate interferometer (see § 4.1.4) or by using an optical flat to re-direct the beam back to the source and then adjusting the focus until the spot size is a minimum. The focal length of the collimator was chosen to be 1.5 m for two reasons. Firstly, the larger the collimation length, the smaller the obliquity effect due to the size and positioning error of the source (see § 4.1.3) and secondly there was a commercially available achromatic doublet of focal length 1.5 m, suitable for the instrument (see optical testing in Appendix B).

The collimated beam is directed by a 150 mm diameter mirror (hereafter called the ‘collimator mirror’) onto the beamsplitter. The beamsplitter is coated for a nominal 50/50 transmission/reflectance and is wedged at 0.5° along its major axis to prevent secondary reflections from interfering with the primary transmitted and reflected beams.

The beam transmitted by the beamsplitter is the reference beam and is directed by a mirror angled at 45° (not shown in figure 3.1 for clarity) onto the reference mirror. The reference mirror actually lies in a horizontal plane, *i.e.* with its primary face horizontal. This face is coated with chromium which has a similar reflectivity (see Appendix A) to the measurement faces of length bars. In this way the contrast of the fringes is maximised since the reference beam and measurement beam have equal intensities.

The beam reflected by the beamsplitter is directed onto the length bar and platen by a remotely-adjustable mirror. Part of this measurement beam is reflected from the front surface of the bar and part by the platen surface.

The measurement and reference beams re-combine at the beamsplitter and are directed via two mirrors, a de-collimating lens and an imaging lens onto a CCD camera. The interference fringes formed by the interaction of the reference and measurement beams are located at infinity once they have emerged from the de-collimating lens and are imaged onto the CCD array by the imaging lens. This lens is also used to focus the image of the end of the length bar onto the CCD array. The image viewed by the camera is thus a superposition of the interference fringes, the image of the bar, and the partially out of focus image of the platen surface.

3.1.2 The interferometer chamber

Because the accuracy of the interferometer is dependent on knowing accurately the refractive index of the air in the measurement beam and because accurate fringe fraction measurements require stable fringes (and no air turbulence), the entire

interferometer (except for heat sources such as the camera and lasers) was constructed inside a sealed, aluminium-walled chamber. This ensures that the refractive index inside the interferometer is uniform and relatively stable and not affected by possible external contaminants such as exhaled water vapour and CO₂ or vapour from solvents such as acetone or alcohol used for cleaning bars and platens.

An alternative would have been to measure the bars in a vacuum. This idea was rejected for two reasons. Firstly, the length of a metal bar is shorter when in vacuum compared to in air due to the absence of air pressure compressing it. For example, a 1 m bar would shorten by approximately 20 nm under vacuum conditions compared to normal atmospheric pressure. The exact magnitude of the shortening would depend on material properties of the bar, such as the exact value of Young's modulus. Without accurate determination of these properties, especially the transition from un-hardened to hardened material near the end faces, the exact change in length of the bar could not be calculated, and hence the length of the bar at 1 atmosphere would not be known. Secondly, the absence of any conductive or convective medium in the chamber would lead to an increase in the thermal soaking time required for the bars to reach thermal equilibrium at each temperature. Thus the use of a vacuum system was rejected in favour of a sealed chamber at atmospheric pressure.

The pressure inside the chamber varies slowly (the chamber is not pressure-sealed) with external atmospheric pressure variations. A totally sealed chamber was considered, but this would have required a substantial stainless steel construction which would have had safety implications. The chamber is temperature-controlled (see Chapter 8) further reducing variations in refractive index. (For the effect of these variations on refractive index, see § 7.3).

The chamber is a 3 unit construction and consists of the optical table surface, a 'collar', and a lid. The optical table has a surface of non-magnetic 300-series stainless steel, with sealed tapped holes and a specially lapped area for sealing against a silicon seal on the bottom face of the collar. The thermal expansivity of the steel is $16.6 \times 10^{-6} \text{ K}^{-1}$. The collar is made of welded aluminium, 16 mm thick, 350 mm high. The collar encloses an area of 2250 x 1050 mm. On one side of the collar there are 6 feed-through ports for connecting equipment inside the chamber to the outside world, together with 2 ports on a long side wall for rotary drive to the translation table and a window for the output beam.

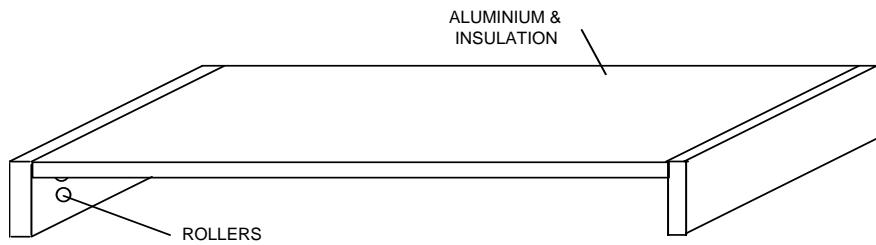


Figure 3.4 - Lid of chamber

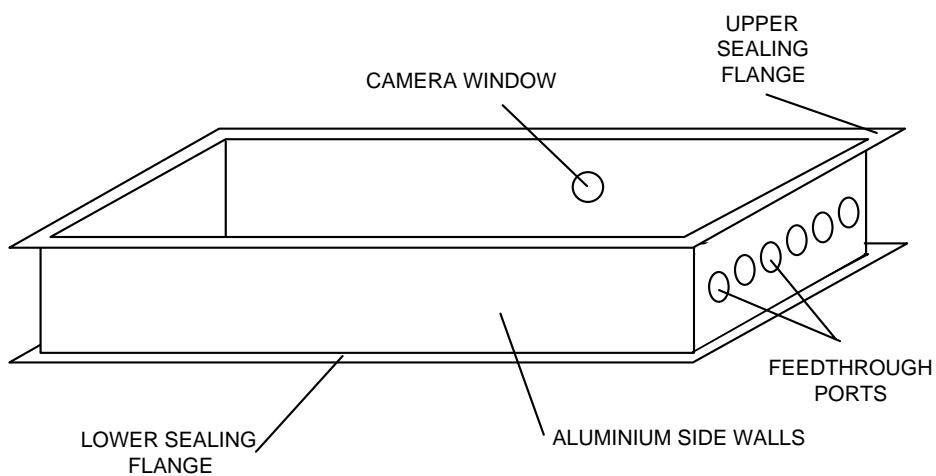


Figure 3.5 - Aluminium 'collar' used as side walls of chamber

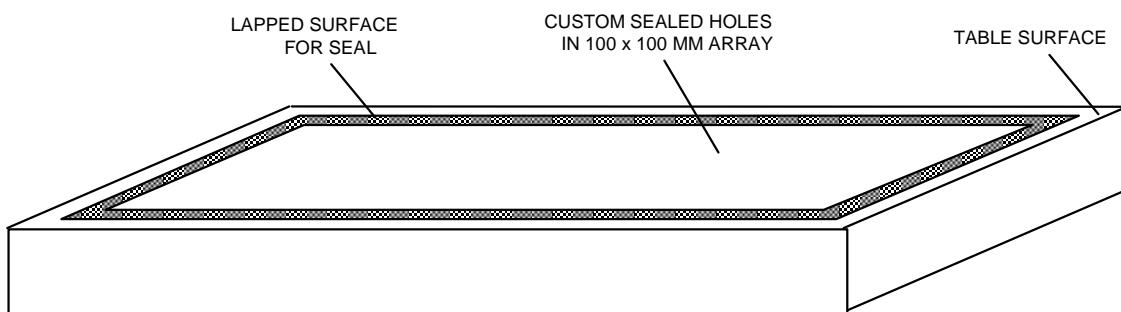


Figure 3.6 - Optical table used as base of interferometer chamber

Further details of the heating and insulation of the interferometer chamber can be found in chapter 8.

3.2 DETAILS OF INTERFEROMETER COMPONENTS

Figure 3.2 (fold-out) shows a plan of the interferometer showing most of the components. Technical specifications and dimensions of the optical and optomechanical components can be found in Appendix A. Some of the components and systems have been specially designed for the interferometer and are detailed below.

3.2.1 Design of lasers used with the interferometer

Because the use of the single mode lasers is vital to the interferometer, the design of the lasers will be examined, starting with the background theory of the operation of the helium-neon gas laser.

3.2.1.1 Helium-neon laser theory

The continuous wave helium-neon (He-Ne) gas laser contains a mixture of approximately 8 parts of helium to 1 part of neon at a total pressure of a few millibars. The laser consists of an optical cavity, similar to that of a Fabry Perot etalon, formed by

a plasma tube with optical quality mirrors at the ends. The gas in the tube is excited by a high voltage discharge of approximately 1.5 to 2.5 kV, at a current of approximately 5 to 6 mA. The discharge creates a plasma in the tube which emits radiation at various wavelengths corresponding to the multitude of allowed transitions in the helium and neon atoms.

The coherent radiation emitted by the He-Ne laser at approximately 632.8 nm corresponds to the $3s_2 - 2p_4$ transition in neon. The excited $3s_2$ level is pumped by energetic $2s_1$ helium atoms colliding with the neon atoms; the $2s_1$ helium energy level is similar in energy to the $3s_2$ level of neon and the lighter helium atoms are easily excited into the $2s_1$ level by the plasma discharge (see figure 3.7). The excess energy of the collision is approximately equal to kT , *i.e.* it is easily removed by the atoms in the plasma as kinetic (thermal) energy.

The collisional pumping of the $3s_2$ level in neon produces the selective excitation or population inversion which is required for lasing action. The $2p$ neon state decays in 10^{-8} second to the $1s$ state, maintaining the population inversion. The $1s$ state relaxes to the ground state by collision with the walls of the plasma tube. The laser gain is quite small and so losses at the end mirrors must be minimised by using a high reflectance coating, typically 99.9%. The output power is limited by the fact that the upper lasing state reaches saturation at quite low discharge powers, whereas the lower state increases its population more slowly. After a certain discharge power is reached, further increase in the power leads to a decrease in the population inversion, and hence lower light-power output.

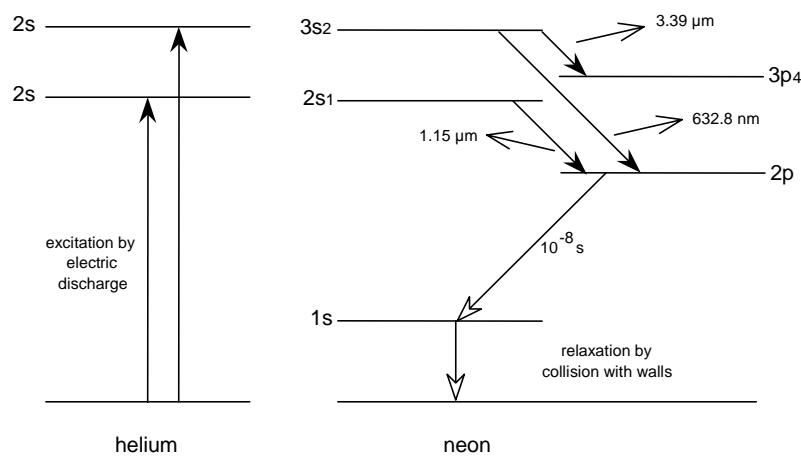


Figure 3.7 - Energy levels in the He-Ne gas laser for 632.8 nm radiation

The 632.8 nm (approximate) operating wavelength is selected by the spacing of the end mirrors, *i.e.* by the total length of the optical cavity, l_c . The length of the cavity must be such that waves reflected by the two end mirrors are in phase. The wavelengths of successive axial modes are then given by

$$2l_c = m\lambda \quad (3.1)$$

These modes are separated in wavenumber by

$$\Delta\sigma = \frac{1}{2l_c} \quad (3.2)$$

or in terms of frequency

$$\Delta\nu = \frac{c}{2l_c} \quad (3.3)$$

This would lead to a series of narrow lines of similar intensity in the spectrum, if it were not for the effects of Doppler broadening and the gaussian distribution of atoms available for stimulated emission. The availability of atoms for stimulated emission is given by

$$A(\sigma) = \frac{1}{a\sqrt{\pi}} e^{-\frac{\sigma^2}{a^2}} \quad (3.4)$$

where $a^2 = \frac{2kT\sigma_0^2}{Mc^2}$ (3.5)

k is Boltzman's constant, T is the temperature of the gas, M is the molecular weight of the gas, and σ_0 is the line centre. This distribution has a half width at $w_\sigma = 2a\sqrt{\ln 2}$.

Also, when a particular mode is oscillating, there is a selective depopulation of atoms with specific velocities (laser cooling) which leads to a dip in the gain profile. For modes oscillating away from the centre of the gain curve the atomic populations for the two opposite directions of propagation are different due to the equal but opposite velocities. For modes oscillating at the centre of the gain curve, the two populations become a single population of effectively stationary atoms. Thus a dip in the gain profile occurs at the centre of the gain curve - the so called "Lamb dip". Some early laser stabilisation schemes locked to the Lamb dip [2] but the position of the dip is dependent on other parameters of the laser such as the position of the gain curve and so is unstable.

For early lasers with a typical cavity length of 1 m the mode spacing was 0.5 m⁻¹, with a gain profile width of approximately 5.5 m⁻¹. Thus several axial modes were present in the gain profile with gains sufficient for laser action, and so two or more modes would operate simultaneously, making the laser unsuitable for interferometry. By using a shorter tube and then carefully lowering the power of the discharge and hence lowering the gain curve, it was possible to achieve single mode operation.

Having selected a single mode, the fundamental limitation to the linewidth (and thus to the temporal coherence) is spontaneous noise fluctuations which broaden the line into a Lorentzian function [3]

$$w_\sigma = \frac{8\pi hc^2 \sigma w_c^2}{P} \quad (3.6)$$

or

$$\Delta v = \frac{\pi h v}{P} (\Delta v_c)^2 \quad (3.7)$$

where P is the operating power of the laser, h is Planck's constant, c is the speed of light and w_c is the linewidth due to the cavity. However a more common-place limitation is the thermal vibration of the tube size which directly affects the wavelength of the oscillating mode. For the lowest frequency stretching mode of the laser tube

$$w_\sigma = \sigma \sqrt{\frac{2kTV}{E}} \quad (3.8)$$

where V is the volume of the tube material of Young's modulus E .

3.2.1.2 Single mode laser wavelength stabilisation schemes

To allow a laser to be used in interferometry with coherence lengths above a few millimetres it must operate in a single mode. There have been many proposed schemes for laser stabilisation [4].

As mentioned above, the Lamb dip was used in an early stabilisation scheme. Here the intensity of the output beam was monitored as the length of the cavity was modulated, for example by Piezo Electric Transducers (PZTs). Another scheme used mirrors external to the laser cavity which were modulated, with the output intensity being monitored and the laser locked to the centre of the Lamb dip [5]. The reproducibility of lasers locked to the Lamb dip is limited by the shift of the Lamb dip centre as the pressure of the gas inside the laser tube varies and also by a discharge current dependent shift. The large width of the Lamb dip itself (about 5×10^{-7} of the laser frequency) also limits the frequency stability obtainable from this technique.

Use has also been made of tuneable Fabry-Perot etalons in a similar system. Other groups have locked the output of one laser to the frequency of a second stabilised laser. Others have used neon discharge absorption cells [6] where the laser was locked to the absorption spectrum of neon in an external tube, the theory being that the un-excited neon would have a narrower linewidth than the neon in the laser discharge.

3.2.1.3 Laser frequency-stabilisation using saturated absorption

The technique with the greatest stability is used in the Primary Reference lasers which represent the UK's Primary Length Standard and involves controlling the length of the cavity to alter the wavelength and locking the wavelength to an absorption line in saturated iodine vapour [7,8,9]. This is a more stable technique since the absorption takes place from a thermally populated energy level which is free from the perturbing effects of the electric discharge in the laser tube.

If the output beam from a laser is passed straight through an absorption cell, then absorption takes place over a Doppler broadened transition. However if the cell is placed in a standing wave optical field then the high intensity laser field saturates the absorption and a narrow dip appears at the centre of the absorption line corresponding to molecules which are stationary or moving perpendicular to the direction of the beam. This dip produces an increase in the laser power in the region of the absorption line. This increase can be detected and the laser frequency stabilised to the centre of the absorption line. The absorption line is reproducible and insensitive to perturbations. The line width is dependent on the absorber pressure, laser power and energy level lifetime. Saturated absorption linewidths are typically less than 1×10^{-8} of the laser frequency.

An evacuated quartz cell containing a small iodine crystal is placed in the laser cavity and temperature controlled to 23 °C. The laser mirrors are mounted on PZTs and the end plates are separated by invar bars to ensure a thermally stable cavity. A small frequency modulation is applied to the laser via one of the PZTs. This leads to an amplitude modulation in the output power which is detected using a phase sensitive detector and fed back to the other PZT as a correction signal. The frequency control system employs a photodiode, low noise amplifier, coherent filter and phase sensitive detector followed by an integrating filter.

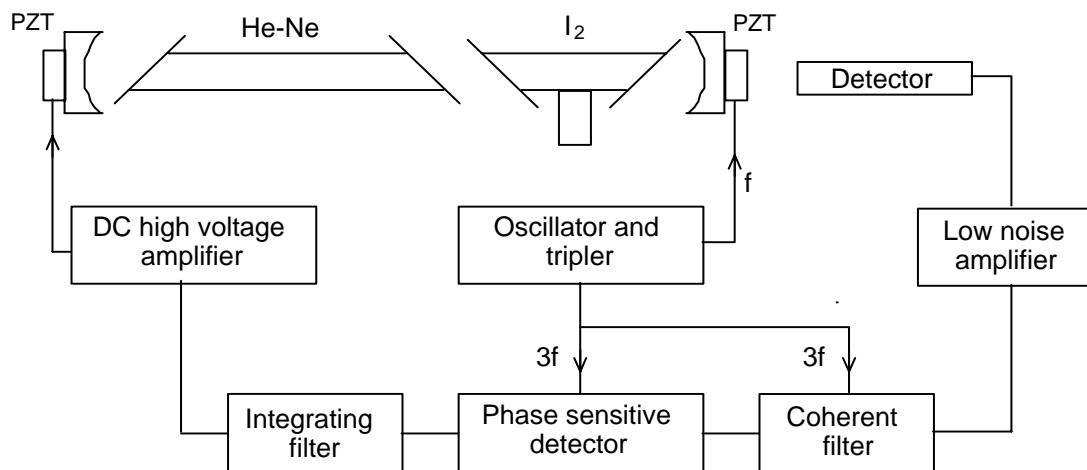


Figure 3.8 - Schematic diagram of an iodine-stabilised He-Ne laser

Detection of the absorption signal at the laser modulation frequency results in a first derivative scan which shows the hyperfine components superimposed on the sloping background of the neon gain curve. The laser may be servo-locked to any of these lines, the frequency of which has been fixed internationally at the time of the re-definition of the metre in 1983 in terms of the speed of light [10] (see § 1.3.5).

Iodine-stabilised He-Ne lasers can achieve frequency stability [11] of a few parts in 10^{13} over a period of a few minutes with long term reproducibility of a few parts in 10^{11} .

3.2.1.4 Zeeman-Stabilised 633 nm Lasers

An alternative technique to saturated absorption is used in the commercial lasers used in the Primary Length Bar Interferometer. The method of stabilisation used for these lasers is based on the Zeeman effect [12,13,14]. A longitudinal magnetic field is applied to a single mode He-Ne laser tube, splitting the normally linearly polarised mode into two circularly polarised modes which are oppositely polarised. A field strength of 0.02 T is sufficient to split the modes, which remain locked together at low B-field, to produce the linear polarisation. These two modes differ in frequency by 300 kHz, around a mean frequency corresponding to the original linear mode [15].

The wavelength difference between the two modes is due to each of the two modes experiencing a different refractive index and therefore different optical path length, in

the He-Ne mixture. This arises due to magnetic splitting of an atomic state of neon, shown in figure 3.9.

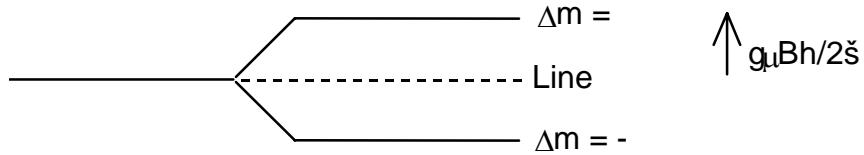


Figure 3.9 - Magnetic splitting of neon - g is the Landé g factor, μ the Bohr magneton, B the field strength, and h is Planck's constant

The $\Delta m=+1$ mode couples with the left polarised mode and the $\Delta m=-1$ mode couples with the right polarised mode. The relative frequencies of the polarisation modes are given by

$$\omega_{\pm} = \frac{cN}{2Ln_{\pm}} \quad (3.9)$$

where c is the speed of light, L is the cavity length, n_{\pm} the refractive index for the mode, and N the axial quantum number [16]. The shape of the refractivities ($n_{\pm}-1$) of the two components and their difference is shown in the following figure.

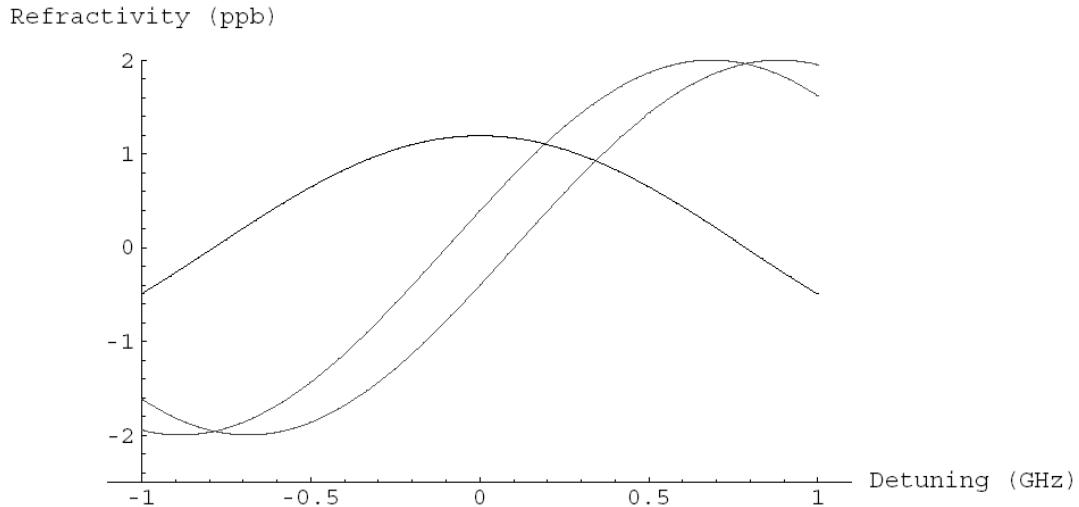


Figure 3.10 - Refractivities of two Zeeman modes in 632.8 nm laser mode: darker line is difference between the two refractivities depicted by the two lighter curves

Note: the exact shape of the difference near w_0 depends on the strength of the Zeeman splitting and the amount of saturation.

The important feature of the Zeeman split gain curve is that the position of w_0 does not vary with B field strength - it remains at the original (un-split) line centre, and is thus a

very stable lock point. If one combines the 2 oppositely polarised components, one observes the beat frequency between them.

from

$$\omega_{\pm} = \frac{cN}{2Ln_{\pm}}$$

$$\Delta\omega = \omega_{+} - \omega_{-} = \frac{cN}{2L} \left(\frac{1}{n_{+}} - \frac{1}{n_{-}} \right) \quad (3.10)$$

which is proportional to

$$\omega_0 [\chi_{+}(\nu) - \chi_{-}(\nu)]$$

where $c_{+}(n)$ and $c_{-}(n)$ are dispersion functions for the left and right polarised modes, respectively. For a more complete derivation see Tomlinson *et al* [17]. As the laser is tuned by altering the cavity length L , the beat frequency will pass through a peak which corresponds to the laser frequency being tuned to ω_0 .

This tuning curve can be used as an error signal for controlling the laser frequency. The particular method chosen to modulate the laser cavity length is thermal expansion. A thin foil heater is attached to the laser tube, connected to a square-root power amplifier. Two magnets are fixed onto the tube to provide the axial magnetic field. A polarising beamsplitter is used, together with a photodetector and amplifier to detect the ~ 300 kHz beat frequency. This error signal is fed to various stages of counters and amplifiers and then to the heater.

The laser tube requires a period of approximately 10 minutes to reach the correct temperature corresponding to the required tube length for operation at frequency ω_0 . A phase-locked loop circuit then fine-tunes the temperature and length of the cavity, to stabilise the laser at the correct frequency. This last process takes only a few seconds to achieve lock. The frequency stability of the laser is 5×10^{-10} for 1 s averages and is white-noise limited for averaging times between 100 ms and 10 min. The day to day resettability of the laser frequency is typically $\pm 5 \times 10^{-10}$. There is also a linear drift of frequency with the amount of time for which the laser has operated. This is due to clean up of the helium-neon mixture, whilst undergoing discharge. The rate of drift is unique to each laser, but is stable with respect to time, and can be ascertained after a few calibrations of the laser frequency. As an example, Sasagawa [18] showed drift rates of 0.3 to 5.7 ± 0.5 MHz per calendar year, though these were for frequency against date, rather than against operational time. Rowley [16] reported a drift rate of -1×10^{-11} per hour of operation.

The operation of the green (543 nm) and orange (612 nm) lasers are roughly similar, although some of the electronic techniques used are different. The frequency stability of the green and orange lasers is similar to that of the red, though the short term fluctuations of the orange laser are larger than those of either the red or green lasers.

3.2.1.5 Calibration of the Zeeman-stabilised 633 nm laser

The calibration of the laser is achieved by launching the light from the Zeeman stabilised laser into a multi-mode optical fibre, which terminates near to the iodine stabilised laser in a different laboratory. The beam combines with that from the Primary laser via a beamsplitter. The beat signal between the two frequencies is measured with a photodetector.

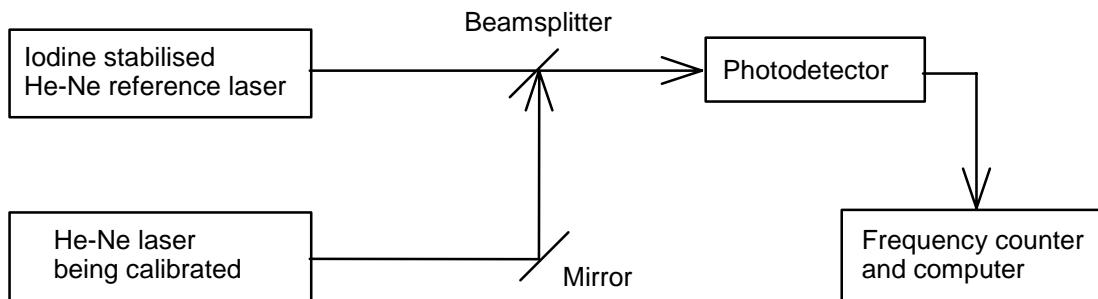


Figure 3.11 - Calibration scheme for Zeeman stabilised laser

This beat frequency is easy to detect, and to monitor via a computer, over a number of hours. A typical value of the beat signal is 260 MHz, with the iodine standard (g component) at approximately 473 612 345 MHz.

The Zeeman laser emits two polarisations (the 2 Zeeman components) which are separated by 0.5 MHz. The 2 components are present in the beam at the interferometer which thus measures with a wavelength which is the mean of these two components. During the laser calibration, beats between each of these frequencies and the iodine frequency are observed. The mean of these is deemed to be the calibrated wavelength of the Zeeman laser.

The Zeeman-stabilised laser incorporates a modulation signal, which is applied to the heaters to achieve the lock point. This signal is a square wave, of frequency 3-5 Hz. This causes a periodic contraction and expansion of the laser tube, and thus the frequency (and wavelength) of the red laser undergoes a sinusoidal modulation. The depth of this modulation is 7 MHz peak to peak, and is approximately 70° out of phase with the square wave modulation signal.

The heater signal is inverted and is used as a synchronisation signal for the timing of the phase stepping digitisation. Thus it is important to calibrate the laser at the synchronisation point, which occurs at the rising edge of the inverted (TTL level) signal (falling edge of the heater signal).

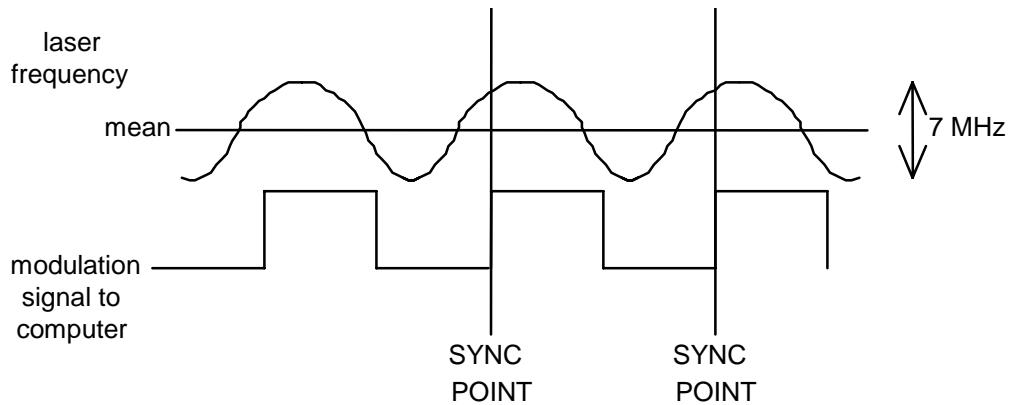


Figure 3.12 - Synchronisation diagram for red Zeeman stabilised laser

By using a frequency to voltage converter, it is possible to synchronise the frequency determination to the correct edge in the modulation signal.

As an example, a red laser, serial number U5, was calibrated after 4103 hrs of use. The mean frequency of the laser was found to be 473 612 605 MHz, and the correction for the modulation signal was +3 MHz, making the effective frequency 473 612 608 MHz, and the vacuum wavelength 632.990 872 57 nm.

3.2.1.6 Calibration of 612 nm and 543 nm lasers

The orange (612 nm) and green (543 nm) lasers are also calibrated by beat frequency comparison with iodine-stabilised reference lasers. There is no modulation signal and hence no correction for these lasers. The 612 nm reference laser is also an internationally accepted realisation of the metre (see Table 1.3). The 543 nm laser is currently being proposed (at the CCDM) as an additional realisation.

3.2.2 The optical fibre illumination delivery system

The lasers are sources of heat and are thus mounted away from the interferometer. The output beam of each laser, after passing through the acousto-optic modulator, is focused using a pair of lenses into the core of a single mode optical fibre.

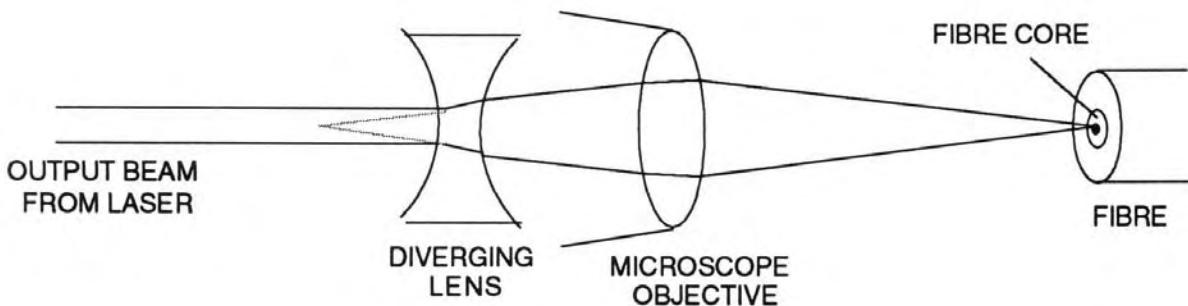


Figure 3.13 - Focusing of laser beam into fibre core

From Fourier-diffraction theory, the size of the focal spot is dependent on the numerical aperture of the input beam. The diverging lens is required to decrease the size of the focused spot so that it fits within the $3 - 4 \mu\text{m}$ diameter of the fibre core. This increases the energy coupling between the beam and the fibre core. The coupling efficiency is also dependent on beam-fibre geometry and the surface form and finish of the fibre. With good quality fibre surfaces the limiting factor is the aberration induced by the diverging lens. To keep the dimensions of the coupling optics small, the power of the lens is relatively high, this leads to higher aberrations and lower coupling efficiency. With longer focal lengths [19] it is possible to increase the efficiency of the coupling to approximately 80% compared to the 20% efficiency obtained with the optics used. Typical coupling efficiency was measured using an optical power meter.

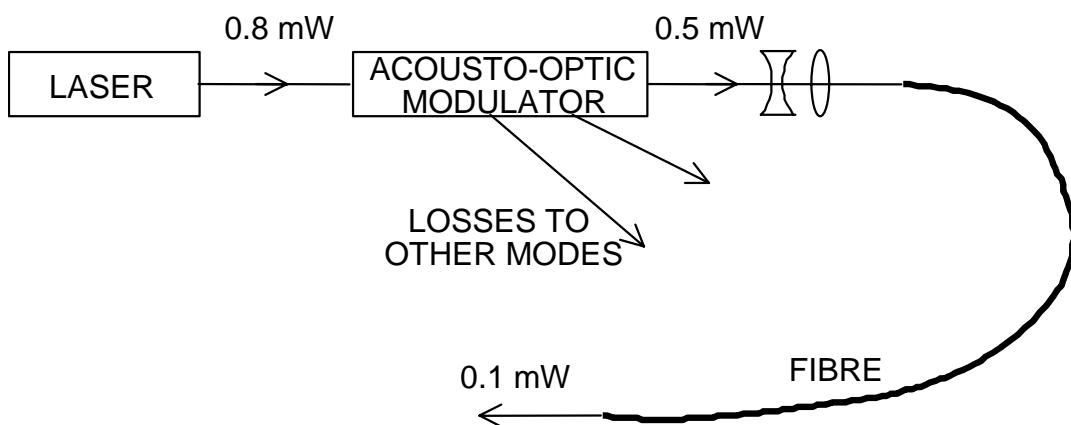


Figure 3.14 - Coupling efficiency of fibre launch

Fortunately 20% coupling efficiency is sufficient for the interferometer. Higher power in the collimator would result in the camera's automatic gain control operating.

Single mode fibres are not as easy to handle or launch light into as multi-mode fibres, however they support only one transmission mode. This prevents the formation of

large-scale speckle in the image which can be a problem unless use is made of electronic speckle interferometry (ESPI) algorithms. Various systems of speckle removal have been tried [20,21]. These rely on mixing the modes, *e.g.* by shaking the fibre or shining through a rotating round glass screen, such that over the averaging period of the detector, the image appears uniform.

For the interferometer it was necessary to have a small diameter source with a relatively small numerical aperture (NA). Most multi-mode fibres have core diameters of 50 µm and greater, and usually relatively large NAs (0.3). Single mode fibres operating in the visible spectrum have approximately 4 µm diameter cores, and small NAs (0.12), making them ideal. The chosen fibres have a core diameters of 2.8 and 3.6 µm and an NA of 0.12.

Each laser beam is focused into one single mode fibre. It is common to couple multiple sources into a single fibre by using spliced fibre couplers.

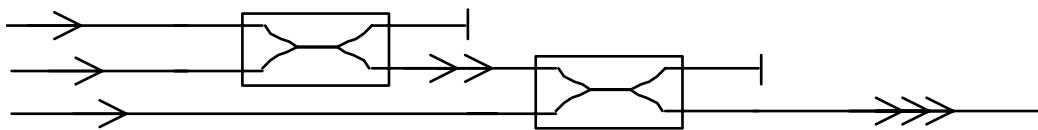


Figure 3.15 - Three source coupling using 2:2 couplers

With common 2:2 couplers a proportion of the input energy is lost via the unused port. Even with 2:1 couplers there are losses in the forwards direction and splitting of the reverse beam decreases the power available for detection when aligning the interferometer (see § 4.1.2.3).

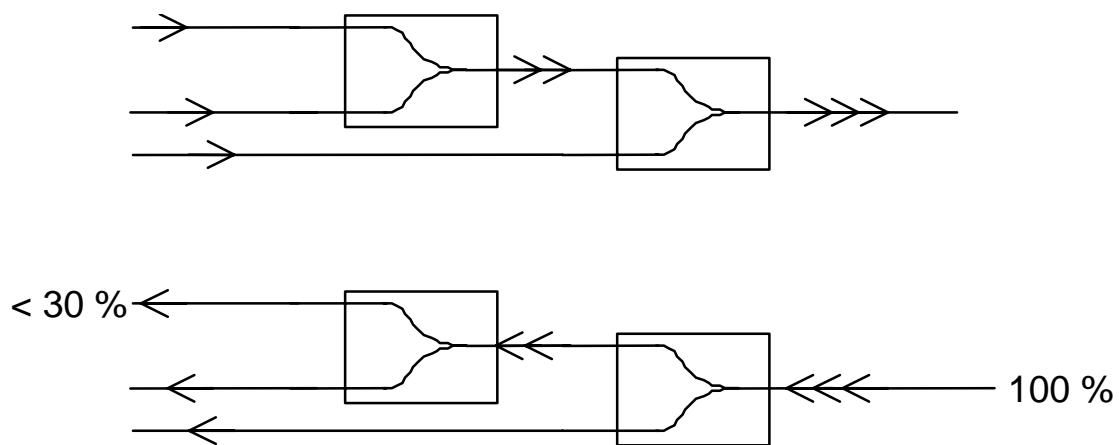


Figure 3.16 - Three source coupling and return spot detection using 2:1 couplers

3:1 couplers can also be used, though these are generally custom-made, expensive and prone to losses, especially single mode devices.

The solution adopted for the NPLBI is to simply use 3 separate fibres and join them mechanically at their output ends [22]. The 3 fibres have the outer buffer coating removed, leaving the 3 - 4 μm core and the 80 - 125 μm primary cladding. The 3 stripped fibres are cemented into a 6 mm diameter ferrule to form a 3-spot light source. Each source thus has an effective diameter of 3 - 4 μm with the sources separated by approximately 100 μm . The emerging beam diverges from each fibre at an angle of 7°, determined by the numerical aperture of the fibre. Thus at the collimator lens, 1.5 m away from the fibres, the beam has expanded to a diameter of over 200 mm, overfilling the 100 mm diameter collimator lens. The lens is thus illuminated approximately uniformly by the central gaussian peak of the beam.

Two different fibres are used in the 3-spot system, a 3M EOTEC FS-SN-3221 for the 633 nm and 612 nm wavelengths, and a 3M EOTEC FS-VS-2211 for the 543 nm wavelength, however in practice there is very little loss whichever fibre is used for each wavelength, and the 3 wavelengths are similar enough to prevent multiple mode operation in any of the fibres.

3.2.3 Reference mirror assembly

The reference mirror assembly consists of a mirror mounted at 45° to the vertical and the reference mirror mounted on a PZT translator. The mounts are attached to a rigid side panel for stability (not shown in figure 3.17).

The PZT is a commercial design digital piezo translator (DPT) (Queensgate Instruments AX100) incorporating a capacitance micrometer at the head of the PZT. PZTs can be prone to various problems during movement. These include hysteresis, non-linearity, and drift with time [23]. These errors can cause errors in the phase-retrieval due to improper phase-stepping [24]. The usual method of overcoming these problems is to use a self-calibration check of the PZT's performance before each measurement is made [25]. This is time consuming and does not always guarantee adequate performance. Alternatively, other phase calculation algorithms can be chosen which are less sensitive to phase-stepping errors [26].

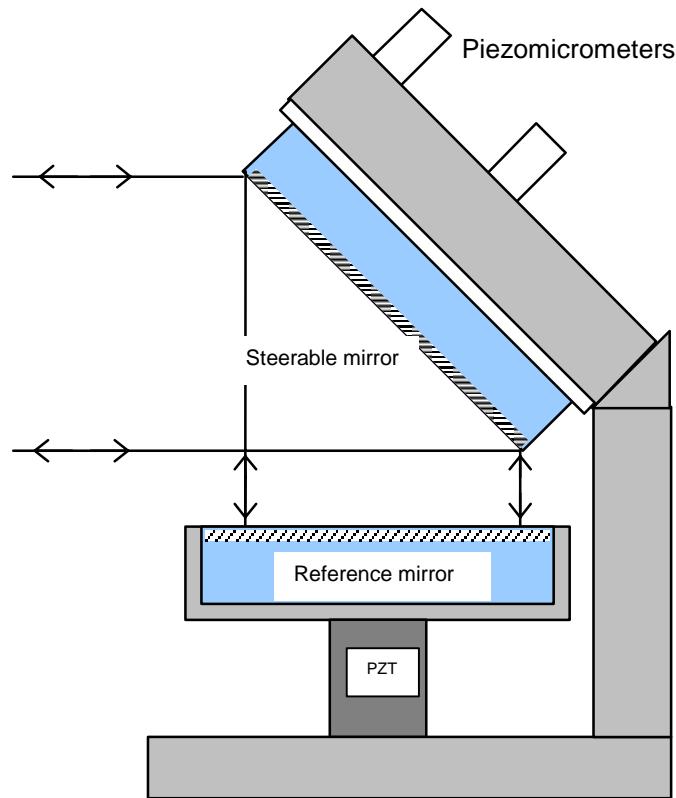


Figure 3.17 - Reference mirror assembly

The commercial DPT system used in the Length Bar Interferometer overcomes the above problems by using a capacitive sensor [27] mounted in the DPT case which responds linearly to elongation of the PZT. The signal from this sensor is used in a feedback loop to correct for any inaccuracies in the DPT movement. The position sensor has picometre sensitivity and the system as a whole has a noise level less than 1 nm RMS (0.003 fringe at $\lambda = 633$ nm).

The DPT case is made of ZERODURTM (a low thermal expansion coefficient glass) and INVARTM to minimise temporal and thermal drift. The whole DPT system is microprocessor controlled allowing digital control of the movement of the DPT, with readout of position attained. Once the readout and the movement of the DPT have been initially calibrated, there is no need to re-calibrate before each measurement, and over the four years that this device has been operating, there has been no noticeable drift in the DPT calibration, *i.e.* the digital phase step size used in the phase-stepping has not been altered from its initial value.

The AX100 positioning controller is addressed by the computer via an IEEE-488 interface. Using a simple set of instructions the mirror can be positioned at any of 16384 points along its operating range. The range is controlled by the resistance of a precision resistor in the input to a summing amplifier which sums positional control

signals from the IEEE interface, an external voltage input, and the front panel potentiometer.

The step size of the DPT system was calibrated by digitising interference fringe intensity at a single point whilst repeatedly stepping the DPT in single steps. The step size was calculated to be 1.066 nm. When performing the phase stepping, the mirror must move in steps of 1/4 fringe ($\lambda/8$) - this corresponds to DPT steps of sizes 74, 63 and 72 for the 633 nm, 543 nm and 612 nm wavelengths respectively.

The angled mirror is adjusted by 2 piezo-micrometers (Physik Instrumente P854.00 Piezomikes). These allow a manual adjustment of 6 mm and fine, remote adjustment using a PZT with 30 μm range. Other designs of mounting the reference mirror were tried, with the mirror mounted vertically, *i.e.* with the PZT stepping the mirror horizontally. These designs used springs or counterweights to try to constrain the mirror. With all these designs it was not possible to prevent tilting of the mirror during phase-stepping, which resulted in a linear variation of phase-step angle down the image, and hence errors in the calculated phase-maps. Since an angled mirror was necessary in the reference arm for steering purposes, it was a simple solution to mount this mirror at 45° and direct the beam downwards onto the reference mirror.

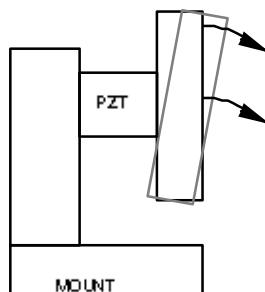


Figure 3.18 - Exaggerated tilting of reference mirror when translating horizontally

3.2.4 Design of the imaging optics

Ignoring path-folding mirrors, the imaging optics of the interferometer are shown in schematic form in figure 3.19. Because the collimator lens is properly focused on the source, the optics in the interferometer are telecentric. A small achromat is used at the focal plane of the de-collimating lens to focus the image of the end of the bar onto the CCD.

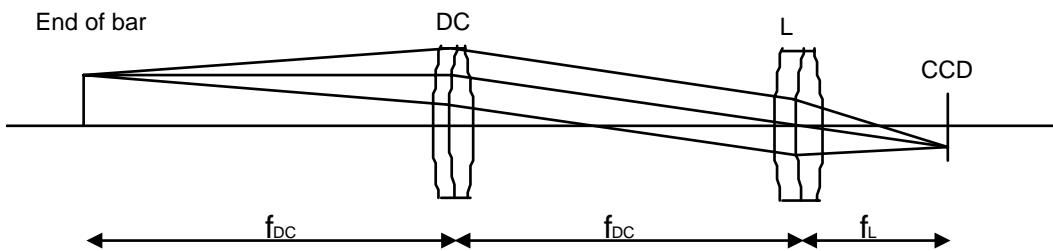


Figure 3.19 - The imaging optics of the NPLBI

Because the Twyman-Green interferometer uses a collimated light source, the source image rays are parallel to the length bar. This means that the images of the end of the bar and the partially out of focus image of the platen will have the same scale in the CCD image plane.

The system of imaging optics is a compromise. There are 2 surfaces of interest in the measurement beam: the exposed face of the bar and the surface of the platen. These are separated spatially by up to 1.5 m (or 3.0 m path difference). Although the use of a small source increases the focal depth of both the fringe and object imaging optics, it is still necessary to focus at some point in the object path.

For consistency, the plane that was imaged out of the interferometer was chosen to be the front (un-wrung) face of the bar. Bars positioned in the interferometer are thus placed such that their front faces are in the focal plane of the de-collimating lens. It is the end surface of the bar rather than the platen which is of particular interest for flatness measurements. The surface of the reference flat is effectively smoothed by software to a best fit surface. The imaging system also allows easy adjustment to cope with rectangular long series gauge blocks, which are also measured in the interferometer.

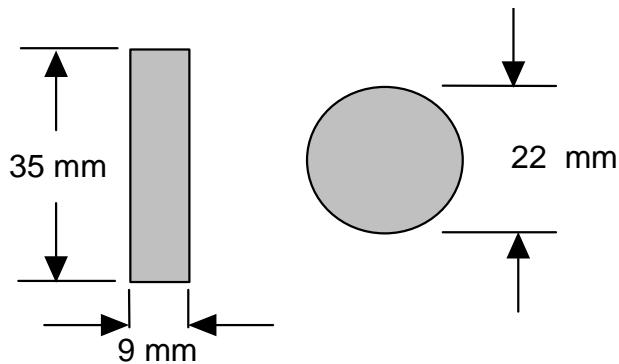


Figure 3.20 - Comparison of end face sizes of length bars and long series gauge blocks

Because long series gauge blocks have to be measured resting on their narrow edge, a different magnification is required to fit the gauge block in the image field. This is achieved by replacing the imaging lens with one of different focal length and re-positioning the camera (two mechanical stops mark the camera positions for gauge blocks and length bars).

Thus the length bar (or gauge block) and platen are both imaged, with the interference fringes, although there is slight degradation of the image at the edge of the bar due to diffraction of the beams at grazing incidence along the bar. This is particularly evident for long bars and necessitates a software ‘mask’ at the edge of the bar in the phase calculations.

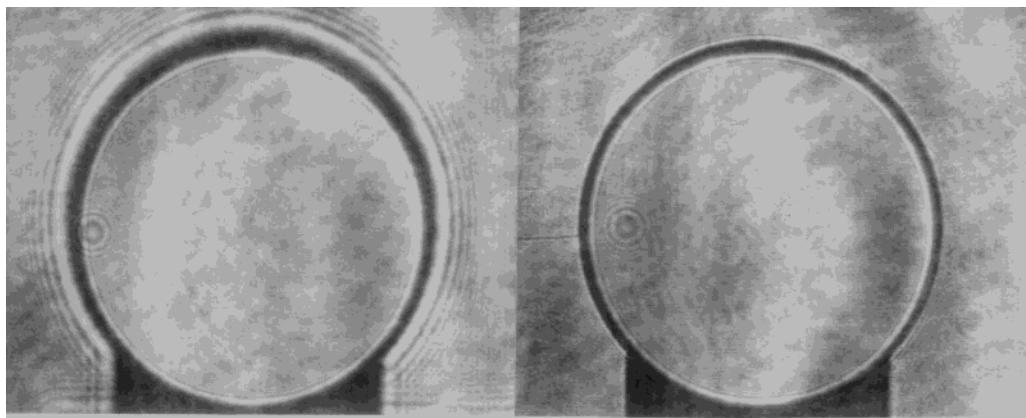
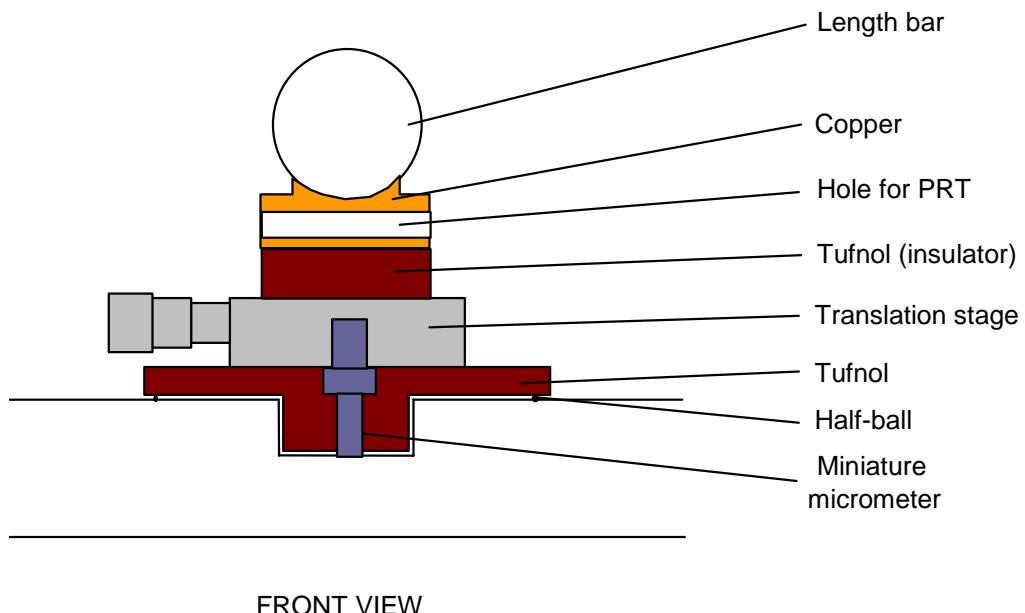


Figure 3.21 - Video prints of images of a 1000 mm length bar (left) and a 225 mm length bar (right) with fringes suppressed, showing different diffraction effects at the edge of the bars due to the different lengths

3.2.5 Design of length bar supports

Due to flexing of length bars under their own weight and the weight of any wrung-on platen, it is conventional to support them at two points along their length in such a way that the end faces remain parallel and vertical - for un-wrung bars these points are called the ‘Airy points’. For bars with platens wrung to one face, the supports points are moved further out, towards the ends of the bar to compensate for the weight of the platen - a full derivation is given in Appendix C.

It is important that there are no further forces acting on the bar other than its weight and the reactions at the two supports otherwise further bending will occur. This prevents attachment of platinum resistance thermometers (PRTs) to measure the bar temperature at points other than the support points. Special supports were designed to include PRTs in such a way that the bar was freely supported at the supports, whilst remaining in good thermal contact with the PRTs.



FRONT VIEW

Figure 3.22 - Length bar supports with integral PRTs, end view

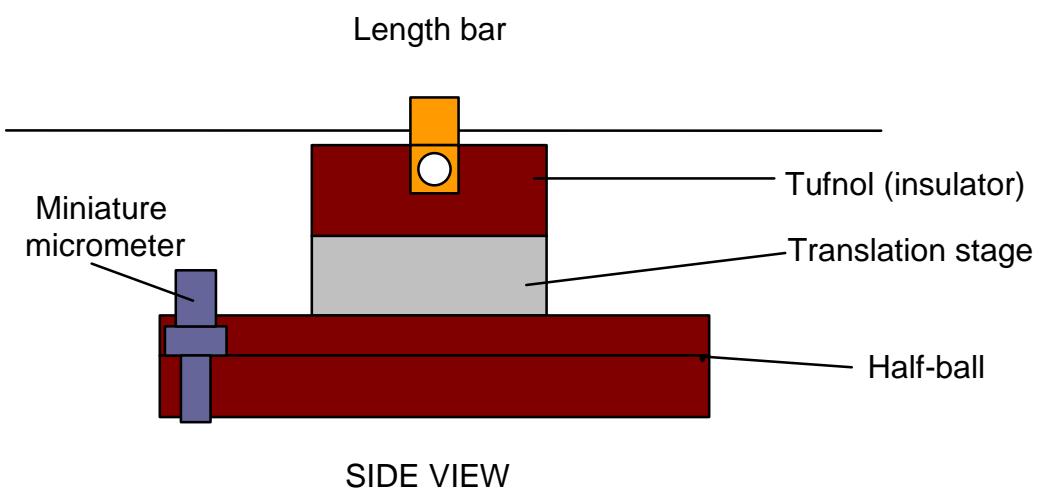


Figure 3.23 - Length bar supports with integral PRTs, side view

Each support consists of a small copper block (chosen for good thermal conductivity), shaped to fit the underside of the length bar or gauge block, with a hole into which a PRT is placed. The copper blocks are fitted tightly into insulating spacer blocks made from TUFNOL™. The supports used at the front end of the bars also contain miniature micrometers and translation stages. These are used to position each bar in the interferometer so that it lies parallel to the measurement beam axis, to within the range of adjustment of the PZT-adjustable mirror. The two PRTs inside the copper blocks are in good thermal contact with the length bar, and hence the temperature of the bar can be measured simply as the mean temperature of these PRTs, assuming any temperature gradients in the bar are small or linear (see § 8.4.1).

The insulating supports are mounted at three vertical-contacting points onto the carriage using three half-balls, one in the slot and two on the top surface. In the case of the adjustable supports, the miniature micrometer takes the place of one of the balls.

3.2.6 Length bar support carriage design

In order to allow more than one bar to be sealed inside the chamber at one time, a translatable carriage was designed to accept up to 3 length bars or gauge blocks. The carriage is made from a block of PERALUMAN (see appendix A) of size 1.5 x 0.26 x 0.02 m with stiffening members attached underneath and at the sides. This forms a solid base for the length bar supports. Three slots were milled into the carriage, 90 mm apart, to a depth of 8 mm, with a mutual flatness and parallelism of 0.5 mm. The length bar (or gauge block) supports fit into these slots and can be positioned at any length along the carriage to cope with bars of different lengths.

The carriage is mounted on a translation stage via a strip hinge. The motion is thus driven from one end via rigid strip which allows the carriage to tilt about only one axis. The other end of the carriage is supported on a bearing which traverses a shaft. The carriage has to be more constrained than a conventional kinematic mount because it must undergo one dimensional translation whilst the bars remain parallel to the measurement beam. If the system is considered as three points in space, corresponding to the bearing and two points at the opposite corners of the strip hinge, then the hinge effectively fixes the x, y and z co-ordinates of two points, and hence fixes rotation about 2 axes, and the bearing fixes y and z of the remaining point. This fixes rotation about the third spatial axis and allows only translation in x (when the x co-ordinates of the other two points translate similarly).

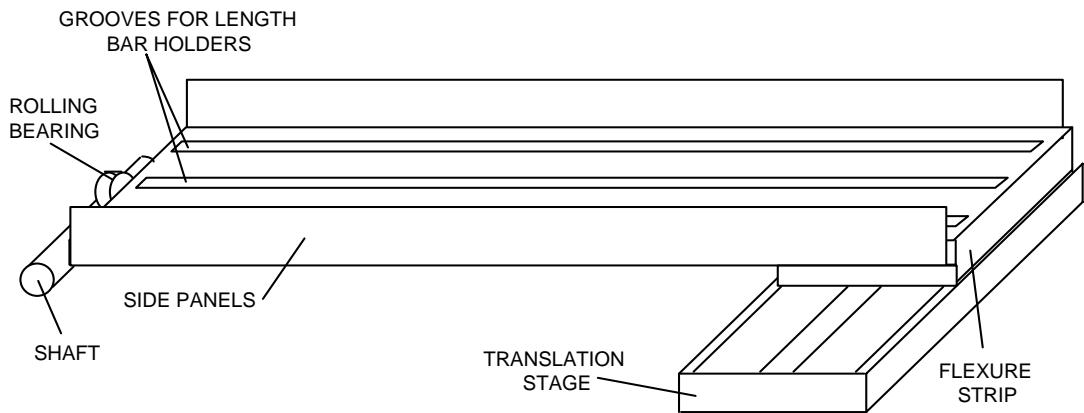


Figure 3.24 - Length bar support carriage

During translation, interference fringes visible in the image remain parallel with the same spacing whilst moving across the image. This indicates that there is no tilting of the carriage, which translates forwards slightly as well as sideways (the translation axis is not perfectly orthogonal to the measurement axis) - this is however no problem. The carriage and supports can be seen in figure 3.25 which shows a close-up view inside the chamber.



Figure 3.25 - View of carriage, supports, length bars and optics inside chamber

3.2.7 Design of the adjustable mirror in the measurement arm

Fine adjustment of the alignment of the measurement beam with the axis of the bar to be measured is achieved using a mirror mounted on a 2-directional flexure. The flexure is tilted by PZT control and is mounted in a commercial high-quality mirror mount for coarse, manual adjustment. The flexure can be tilted in two orthogonal directions about axes which pass through the centre of the flexure stage.

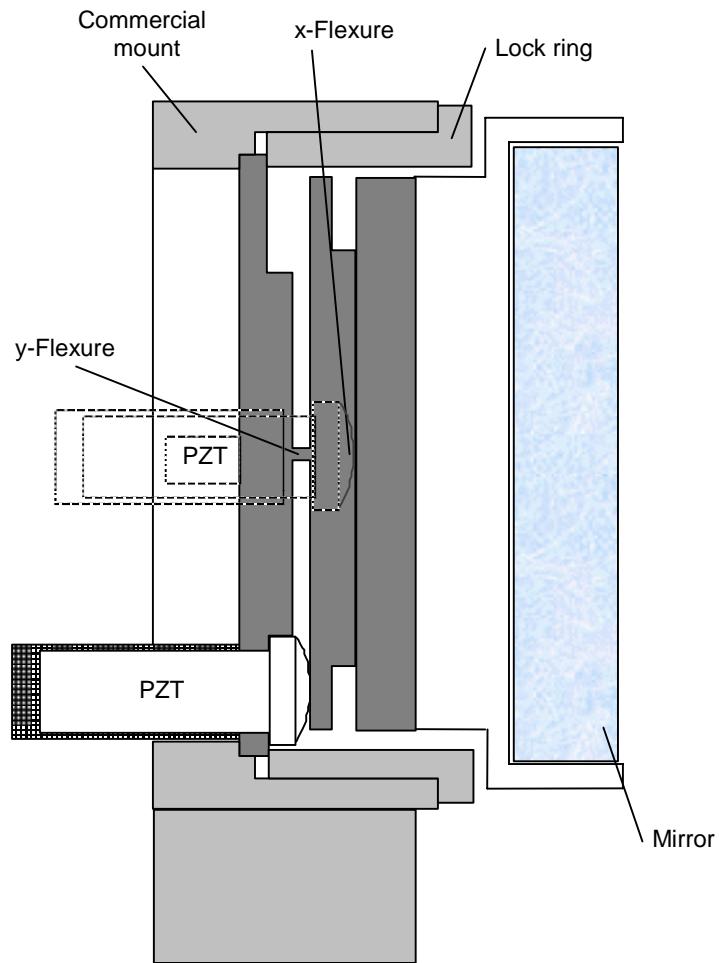


Figure 3.26 - Flexure system for tilting mirror

3.2.8 External equipment rack and computing equipment

The interferometer chamber is connected to an equipment rack which houses the electronic equipment used to perform measurements inside the interferometer. These units are: a Druck pressure transducer, a Michell hygrometer transducer & displays, a CO₂ meter and display, a Tinsley precision resistance bridge with 15 channel selector switch and standard resistor, PZT power supplies, a PTFE re-circulating sample pump,

the controller for the Queensgate AX100 DPT system, a power supply for the carriage motor, the CCD camera power supply and a residual current circuit breaker.

In turn, signals from the rack are routed to the control computer and display monitor. The signals are TTL level IO signals, IEEE bus, video signal and video genlock signal. The water temperature controller is connected with flexible PVC tubing to the lid and baseplate pipes.

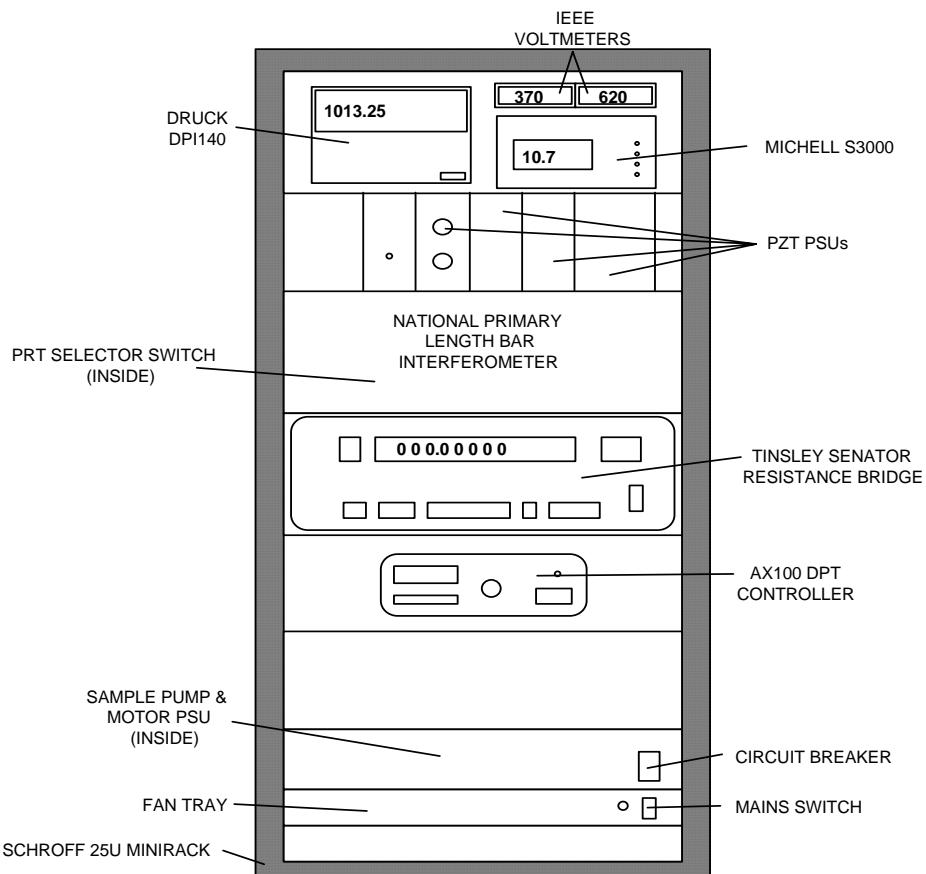


Figure 3.27 - Schematic front view of equipment rack

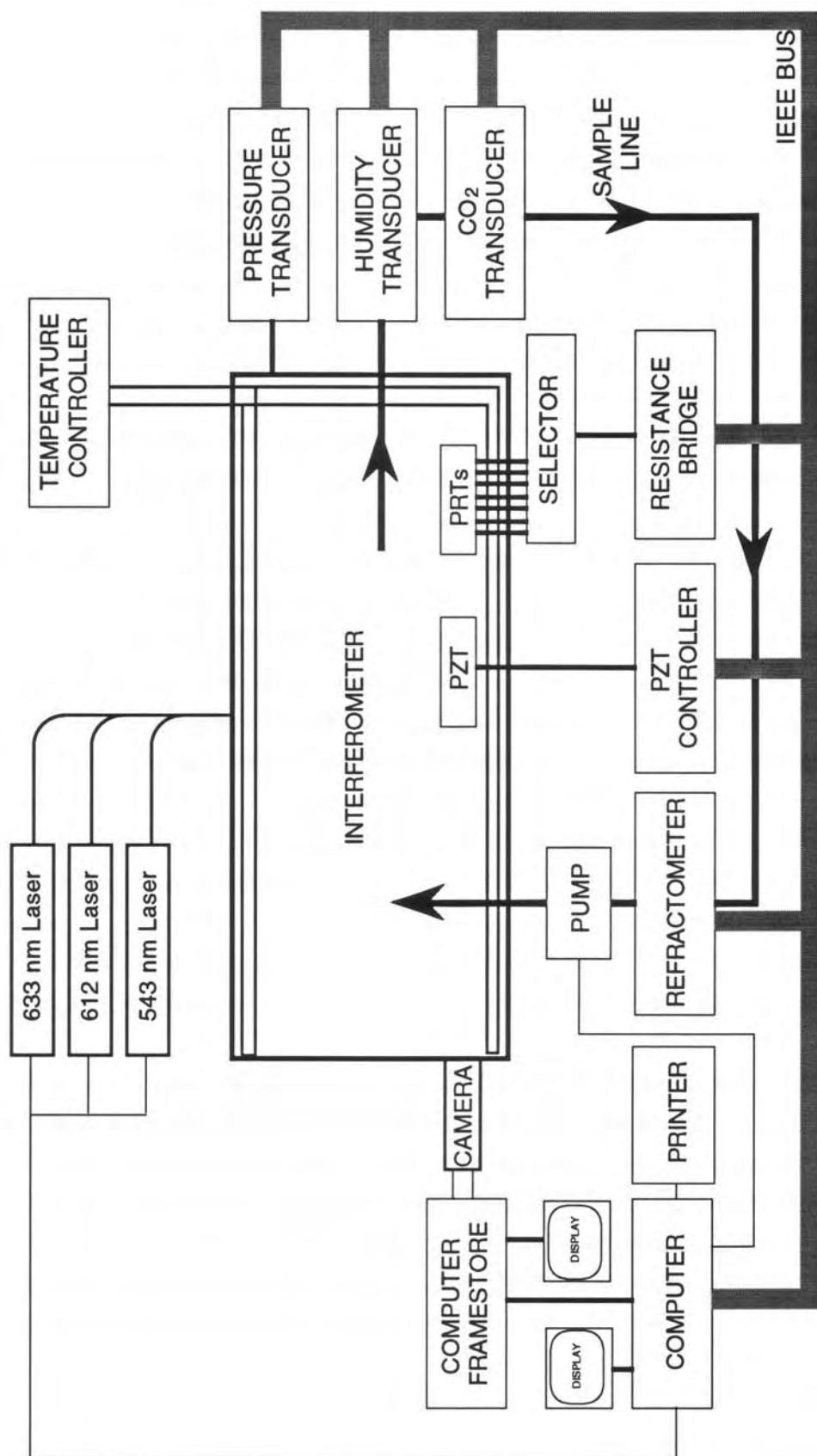


Figure 3.28 - Schematic diagram of interferometer & external equipment

3.3 DOUBLE-ENDED INTERFEROMETRY

This addition to the interferometer was not part of the original design but was conceived as an additional experimental technique to enhance the capabilities of the interferometer and to provide a useful research tool for future measurements concerned with the ‘wringing film’ thickness and the ‘phase correction’ (see definitions of these terms which follow).

When measuring the length of a bar wrung to a platen, the result obtained by an interferometer may be different to that obtained using a stylus or probing instrument because of four factors: the thickness of the ‘wringing film’, the surface roughness, the ‘phase correction’ and the geometrical form of the measuring faces of the bar.

By measuring the length of the bar without being wrung to a platen, the effect of the wringing film is removed and thus measurements of wringing film thickness can be attempted. There is a complication however in that there will be a different size phase correction since the beam will be reflecting off both ends of the bar rather than reflecting off one end of the bar and also the platen surface. A second advantage is that the double-ended system will allow measurements of both end faces simultaneously. With interferometry using wrung bars, one face is obscured by the platen, and any surface effects which would affect the parallelism measurements are averaged out by fitting a function to the surface of the platen.

3.3.1 The wringing film thickness

When being measured interferometrically according to BS 5317, a length bar should be measured with a platen wrung to one face as shown in figure 2.13 and described in § 2.5.1. The measured length then includes the thickness of a ‘wringing film’ - the molecular layer that separates the surfaces of the length bar and platen. However, in some situations a length bar will be used as an artefact, *e.g.* for validating the performance of a Coordinate Measuring Machine (CMM), where it will be used in the un-wrung state. For accurate measurements, the thickness of the wringing film should then be subtracted.

3.3.2 Surface roughness

The effect of the surface roughness of the platen and length bar surfaces depends on the size of the probe used to probe them. For a conventional mechanical probe (as used on

the NPL Length Bar Machine), the probe radius is typically 1-2 mm, thus the probe will contact the ‘peaks’ - the high points of the surfaces. For interferometry, the probe may be considered to be the light beam. In this case the effective size of the probe will be of the order of the wavelength of the light used: 543 - 633 nm. This will penetrate areas of the surface and will reflect from both the peaks and valleys, appearing on average to reflect from a surface mid-way between the peaks and valleys. BS 4311 (1993) for gauge blocks specifies that the platen should have a flatness of 0.025 µm (25 nm) over the measuring area - this is similar to the roughness expected of the gauge block and length bar surfaces. Thus the effect of surface roughness on interferometrically-measured length ought to cancel since the platen and length bar (or gauge block) should have similar surface roughnesses.

However, this is not the case, as the surface roughness of gauge blocks has been found to vary by approximately 20 nm within a single set of gauges [28]. Because length bars are made of a similar material, a similar variation is expected.

There are a number of techniques for the measurement of surface roughness [29,30] some of which NPL hopes to research in the next three years with a view to reducing the length measurement uncertainty due to surface roughness for gauge block and length bar calibrations. The most common technique used currently for measurement of surface roughness of gauge blocks uses an integrating sphere (Ulbricht’s sphere). In this technique, white light is reflected off a surface to be measured and by means of an integrating sphere with photoelectric detector, the ratio of the diffuse (R_d) to specularly-reflected light (R_s) is measured. It is assumed that the roughness of the test surface is proportional to the square-root of this ratio (R_d/R_s). This technique requires a calibrated surface, for reference.

Because the new interferometer does not yet have a means of measuring the surface roughness of the platen and length bar surfaces, the difference between the two values will lead to a length measurement error. Over the years 1989-1993, the overall correction applied to measurements of steel gauge blocks on steel platens at NPL has ranged from - 40 nm to + 15 nm, with the mean at - 14 nm. One can expect a similar, or larger value for measurements of length bars on platens. Thus a figure of $-14 \text{ nm} \pm 25 \text{ nm}$ will be assumed for later uncertainty calculations.

3.3.3 Phase change on reflection

When light reflects at normal incidence off the surface of a dielectric of higher refractive index than the incident medium, there is a π radians phase shift on reflection

due to the continuity of the E-field at the boundary between the two media. If the reflecting medium is a non-dielectric, it has a complex refractive index and by solving the equations for the amplitude reflection coefficients at the surface it can be shown that there is a phase shift of between 0 and π radians, depending on the properties of the media. This is the case when the beam in the interferometer reflects off the surfaces of the bar and platen. It can be shown [31] that the phase shift, δ , is given by

$$\tan \delta = \frac{2n_1 k_2}{n_1^2 - n_2^2 - k_2^2} \quad (2.10)$$

where n_1 and n_2 are the refractive indices of the incident and reflective media respectively and k_2 is the extinction coefficient in the reflecting (semi-absorbing) medium.

From reference texts [32], typical values for these variables for steel are $n_2 = 2.4$, $k_2 = 3.4$ ($\pm 20\%$). For reflection in air, $n_1 = 1$, this gives a value $\delta = -23^\circ$ (or $180^\circ - 23^\circ = 157^\circ$). Thus the light beam appears to reflect off surfaces approximately 20 nm outside the gauge and platen surfaces, even if the surfaces are microscopically smooth. For a 20% variation in k_2 , δ varies by 5° , corresponding to a variation in measured length of 4.4 nm, for $\lambda = 633$ nm. This variation in k_2 is probably typical for the range of steels used for gauge blocks, length bars and platens. The worst possible difference in material properties could therefore give rise to a length measurement error of ± 8.8 nm.

Techniques for measuring the phase change on reflection range from multiple-reflection interferometry [33] comparing the phase shift on reference and test surfaces to the phase-stepping interferometer of Ishikawa *et al* [30] which used a coupled Twyman-Green-Fizeau interferometer. In this scheme, the ‘optical’ surface of the gauge block, corresponding to the mean surface from which the light is reflected and the surface of the flat, form a Fizeau interferometer. The Fizeau interferometer is one arm of a Twyman-Green interferometer. This design is used to measure the apparent position of the optical plane of the gauge block by phase-stepping the Twyman-Green interferometer to obtain information about the phase shift. Unfortunately, this technique assumes a zero wringing film thickness and zero surface roughness of the optical flat, which may not be the case.

The effects of the surface roughness and the phase change on reflection are often combined into one correction termed the ‘phase correction’.

3.3.4 Effect of surface form errors of measuring faces

In the interferometer, the length of the bar is measured as the perpendicular separation between the centre of the exposed measurement face and the surface of the platen wrung to the other end. This requires an interpolation from the phase data measured on the platen to the area covered by the bar. If the wrung measurement face of the bar is not flat, the interpolation will not accurately reflect the shape of the bar. If the other measuring face is flat, then when measured with the other face wrung, the length measurement will be different. However, both measurements are in accordance with the standard, which is why it is important to measure the flatness of the measuring faces of length bars (see § 6.5).

3.4 OTHER DOUBLE-ENDED DESIGNS

In the double-ended interferometer of Dorenwendt [34] the measurement beam of a Michelson interferometer is split and extended with additional mirrors, forming a triangularly closed path.

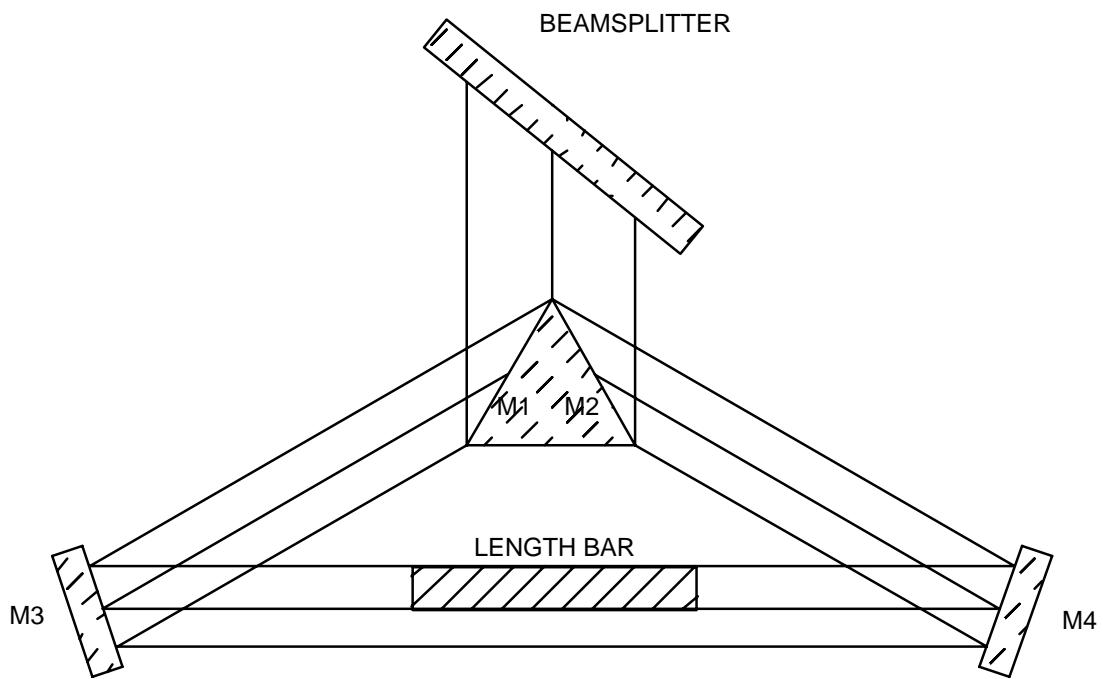


Figure 3.29 - Double-ended interferometer of Dorenwendt

In this design, four extra surfaces are introduced, increasing the complexity of the alignment and introducing further aberrations into the system. In the analysis of this interferometer, Dorenwendt arrived at two values which can be measured in the interferogram which he called ‘fictitious step heights’ L_1 and L_2

$$L_1 = \frac{L}{2} - a \quad L_2 = \frac{L}{2} + a \quad (3.11)$$

where L is the length of the bar and a is the deviation in the direction of measurement from the symmetrical position of the length bar in the optical path. Thus the sum of the two measured step heights represents the optical length of the bar in the un-wrung state.

Dorenwendt further stated that the effects of the phase change experienced by the light in reflecting from the extra mirrors will cancel, provided that the surface characteristics of the pairs of mirrors are similar. This will now be examined.

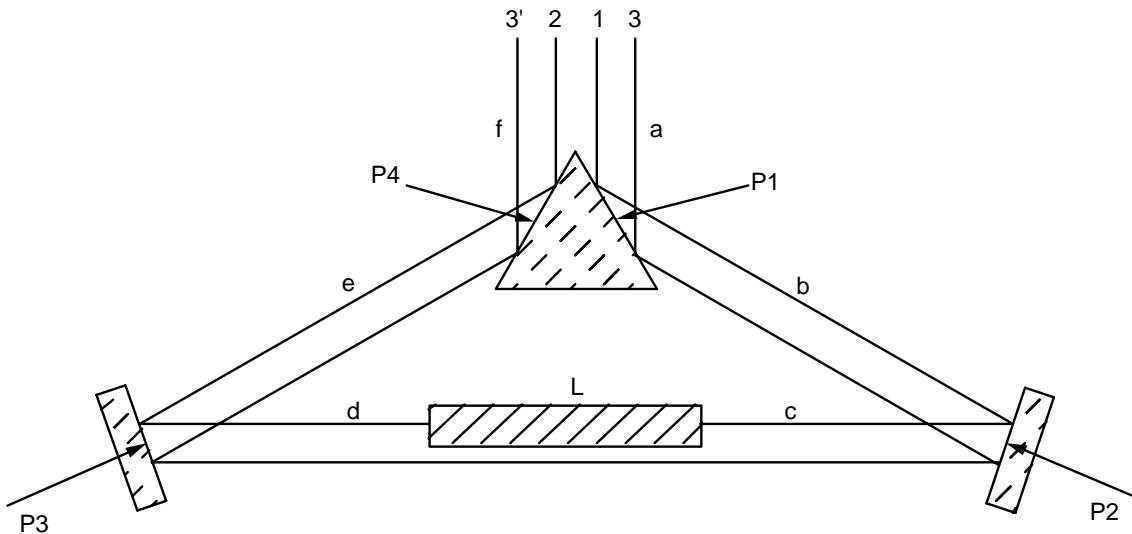


Figure 3.30 - Analysis of triangular interferometer of Dorenwendt

Let the optical path length of each of the three beams (right face, left face, reference) be given by ϕ .

$$\begin{aligned}\phi_1 &= 2(a + b + c + P_1 + P_2) \\ \phi_2 &= 2(f + e + d + P_3 + P_4) \\ \phi_3 &= (a + b + c) + (d + e + f) + L + (P_1 + P_2) + (P_3 + P_4)\end{aligned}\quad (3.12)$$

Now the measurable quantities in the image are the phase differences $\phi_3 - \phi_1$ and $\phi_3 - \phi_2$.

$$\begin{aligned}\phi_3 - \phi_1 &= L - (a + b + c) - (P_1 + P_2) + (d + e + f) + (P_3 + P_4) \\ \phi_3 - \phi_2 &= L + (a + b + c) + (P_1 + P_2) - (d + e + f) - (P_3 + P_4)\end{aligned}\quad (3.13)$$

Thus the sum of the two measured phase differences

$$(\phi_3 - \phi_1) + (\phi_3 - \phi_2) = 2L \quad (3.14)$$

From equations (3.13) the quantity $(a + b + c) + (P_1 + P_2) - (d + e + f) - (P_3 + P_4)$ can be identified as corresponding to $2a$ in equation (3.11).

Note that these equations do not require that $(P_1 + P_2) = (P_3 + P_4)$ unless it is desired that $a = 0$, for which $(a + b + c) = (d + e + f)$ i.e. the bar is in the symmetrical optical path position. Thus the effects of the surface properties of the additional mirrors on the phase change do not affect the overall length measurement. However aberrations of the mirrors will affect the wavefronts and could lead to an error in the measured length or the surface form of the bar.

3.5 DOUBLE-ENDED INTERFEROMETRY IN THE PRIMARY LENGTH BAR INTERFEROMETER

As part of the Primary Length Bar Interferometer a new double-ended interferometer has been designed and constructed. This design requires only one additional pair of mirrors which are used in the form of a pair of roof mirrors. The bar and carriage are displaced sideways and the roof-mirror system placed behind the bar. A different lens and camera position are selected because a different magnification is required to image both ends of the bar, compared to just one end. The roof-mirror arrangement is shown in figure 3.31.

In the image, there are 3 distinct regions: the front face of the bar, the rear face and the reference surface surrounding the bar. In this case, the reference surface corresponds to the surface of the beamsplitter since the beam travels around the bar twice, i.e. the image in the background area is the interference between the normal reference beam and a sheared version of the measurement beam. This demands not only temporal coherence of double the path length of the interferometer in single-ended mode, but also spatial coherence across the beam, which is sheared.

Corrections must be applied to the length measured in the double-ended mode to take account of the phase change at the surfaces of the bar and their roughnesses - this correction is larger than when the bar is wrung, since in the latter case the surfaces of the platen and the bar face the same direction and should have similar surface roughnesses and phase corrections and only the difference between the corrections for the two surfaces is used, whereas in the un-wrung case, their sum is used.

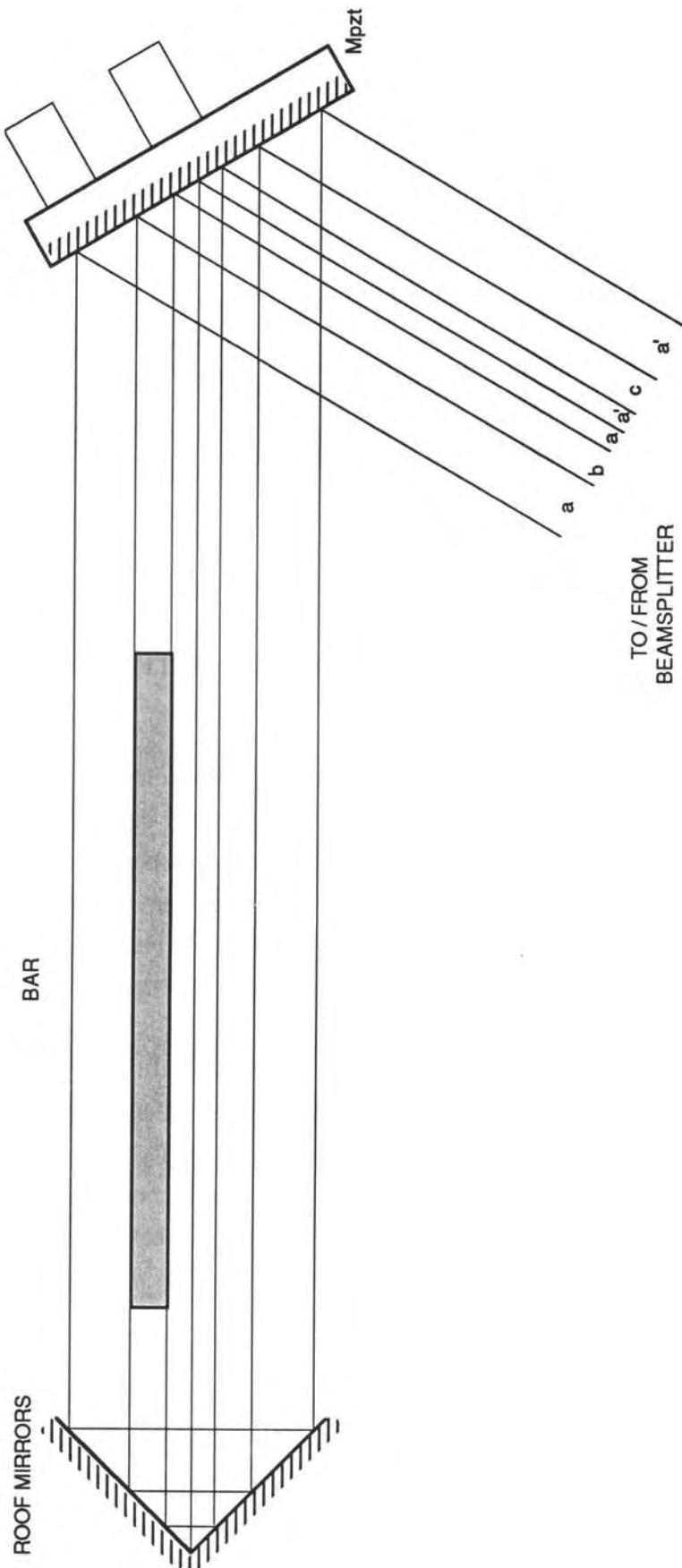


Figure 3.31 - Additional roof-mirror optics for double-ended interferometry

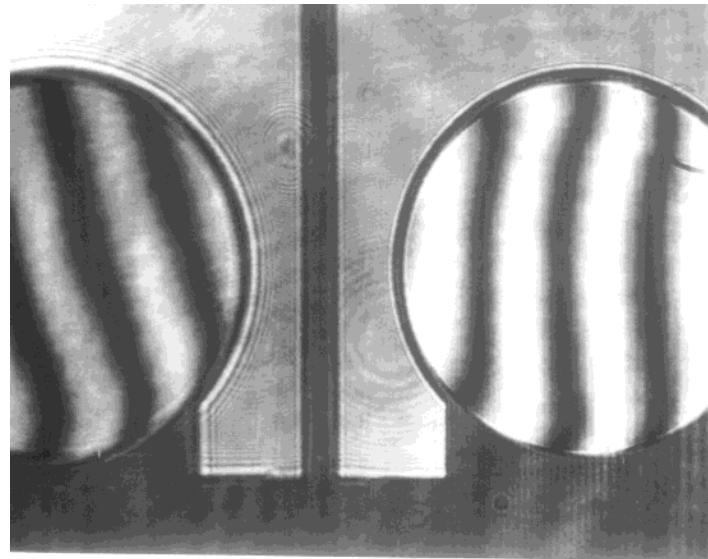


Figure 3.32 - Double ended image

3.5.1 Analysis of new double-ended interferometer

Re-arranging the optics into a linear design simplifies the analysis. The path length of the reference beam of the conventional Twyman-Green interferometer is included in the analysis to show that the interference is still between this reference beam and the beam in the measurement arm.

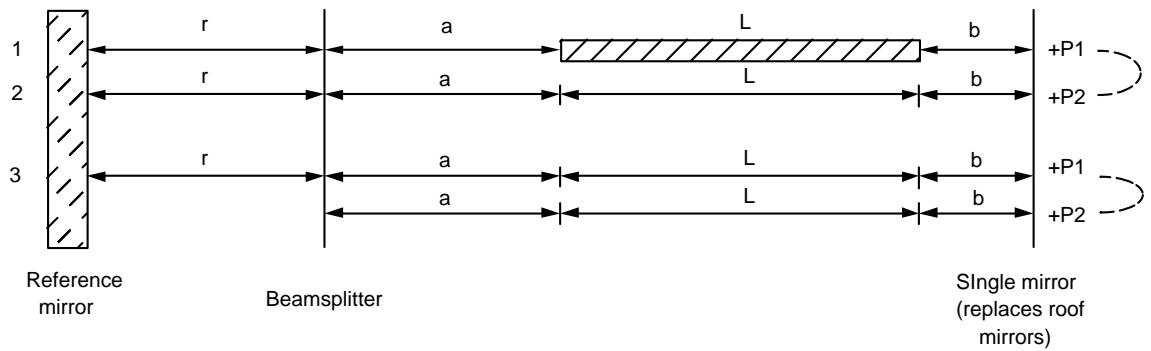


Figure 3.33 - Analysis of new double-ended interferometry

Representing the optical path distances by ϕ gives

$$\begin{aligned}\phi_1 &= 2a - 2r \\ \phi_2 &= 2(a + L + 2b + P_1 + P_2) - 2r \\ \phi_3 &= 2a + 2L + 2b + P_1 + P_2 - 2r\end{aligned}\tag{3.15}$$

Again, the two measured quantities are phase differences

$$\begin{aligned}\phi_3 - \phi_1 &= 2L + 2b + P_1 + P_2 \\ \phi_3 - \phi_2 &= -2b - P_1 - P_2\end{aligned}\tag{3.16}$$

and hence

$$(\phi_3 - \phi_1) + (\phi_3 - \phi_2) = 2L\tag{3.17}$$

As above, the factor of two is due to the double-pass nature of the design (as is the case for single-ended measurement) and means that each fringe is half a wavelength in size. The effects of P_1 and P_2 , the phase changes at the mirror surfaces drop out of the analysis. Thus this system has the advantages outlined above, but uses fewer extra surfaces and is easier to align.

3.5.2 Double-ended phase-stepping

When the double-ended mirrors are adjusted correctly (see § 4.1.1.7), the fringes are straight and continuous across the join of the two mirrors. The fringes are adjusted to have the same tilt as for single-ended measurements (§ 4.1.5). The phase-stepping is performed as for single-ended measurements.

REFERENCES FOR CHAPTER 3

- [1] Lewis A J & Pugh D J Design Note: Interferometer light source and alignment aid using single-mode optical fibres *Meas. Sci. Technol.* **3** (1992) 929-930
- [2] CCDM 5th Session (1973) (*BIPM Sevres, France 92310*) M58
- [3] Jaseja T S, Javan A & Townes C H Frequency stability of He-Ne masers and measurements of length *Phys. Rev. Lett.* **10** (1963) 165-167
- [4] White A D A two-channel laser frequency control system *IEEE J. Quant. Elect.* **1** (1965) 322
- [5] Rowley W R C & Wilson D C Optical coupling effects in frequency stabilized lasers *Appl. Opt.* **11** (1972) 475-476
- [6] White A D, Gordon E I & Labuda E F Frequency stabilisation of single mode gas lasers *Appl. Phys. Lett.* **5** (1964) 97-98
- [7] Brillet A & Cérez P Laser frequency stabilization by saturated absorption *J. de Phys. (France)* **42 (C-8)** (1981) 73-82
- [8] Cérez P, Brillet A & Hartman F Metrological properties of the R(127) line of iodine studied by laser saturated absorption *IEEE Trans. Inst. Meas.* **IM-23** (1974) 526-528
- [9] Cérez P & Bennett S J Helium-neon laser stabilized by saturated absorption in iodine at 612 nm *Appl. Opt.* **18** (1979) 1079-1083
- [10] (Editor's Note) Documents concerning the new definition of the metre *Metrologia* **19** (1984) 163-177
- [11] Allan D W Statistics of atomic frequency standards *Proc. IEEE* **54** (1966) 221-230
- [12] Morris R H, Ferguson J B & Warmiac J S Frequency stabilization of internal mirror He-Ne lasers in a transverse magnetic field *Appl. Opt.* **14** (1975) 2808
- [13] Umeda N, Tsukiji M & Takasaki H Stabilized ^3He - ^{20}Ne transverse Zeeman laser *Appl. Opt.* **19** (1980) 442-450
- [14] Fellman T, Junger P & Stahlberg B Stabilisation of a green He-Ne laser *Appl. Opt.* **26** (1987) 2705-2706
- [15] Baer T, Kowalski F V & Hall J L Frequency stabilization of a 0.633 μm He-Ne longitudinal Zeeman laser *Appl. Opt.* **19** (1980) 3173-3177
- [16] Rowley W R C The performance of a longitudinal Zeeman-stabilised He-Ne laser (633 nm) with thermal modulation and control *Meas. Sci. Technol.* **1** (1990) 348-351
- [17] Tomlinson W J & Fork R L Properties of gaseous optical masers in weak axial magnetic fields *Phys. Rev.* **164** (1968) 480-483
- [18] Sasagawa G S & Zumberge M A Five year frequency stability of a Zeeman-stabilised laser *Appl. Opt.* **28** (1989) 824-825
- [19] Virdee M S *NPL Report MOM 108* (1992) 1-16

- [20] Lowenthal S & Joyeux D Speckle removal by a slowly moving diffuser associated with a motionless diffuser *J. Opt. Soc. Am.* **61** (1971) 847-851
- [21] Bowman M J Two new methods of improving optical image quality *Appl. Opt.* **7** (1968) 2280-2284
- [22] Lewis A J & Pugh D J Design Note: Interferometer light source and alignment aid using single-mode optical fibres *Meas. Sci. Technol.* **3** (1992) 929-930
- [23] Ai C & Wyant J C Effect of piezoelectric transducer nonlinearity on phase shift interferometry *Appl. Opt.* **26** (1987) 1112-1116
- [24] Ohyama N, Kinoshita S, Corneso-Rodriguez A, Honda T & Tsujiuchi J Accuracy of phase determination with unequal reference phase shift *J. Opt. Soc. Am.* **5** (1988) 2014-2025
- [25] Cheng YY & Wyant J C Phase shifter calibration in phase shifting interferometry *Appl. Opt.* **24** (1985) 3049-3052
- [26] Schwider J Phase shifting interferometry: reference phase error reduction *Appl. Opt.* **28** (1989) 3889-3892
- [27] Harb S, Smith S T & Chetwynd D G Subnanometre behaviour of a capacitive feedback piezoelectric displacement actuator *Rev. Sci. Instrum.* **63** (1992) 1680-1689
- [28] Private communication, G Bönsch, PTB Germany
- [29] Bennett J M Recent developments in surface roughness characterization *Meas. Sci. Technol.* **3** (1992) 1119-1127
- [30] Ishikawa S, Bönsch G, & Böhme H Phase-shifting interferometry with a coupled interferometer: Application to optical roughness of gauge blocks *Optik* **91** (1992) 103-108
- [31] Leach R K Investigation into the measurement of the wringing effect and the phase shift at reflection applied to the accurate measurement of end standards *MSc feasibility study*, Brunel University (1993)
- [32] Kaye G W C & Laby T H *Tables of Physical and Chemical Constants* 15th edn (1989) (Harlow, Essex: Longmans)
- [33] Saunders J B A high-sensitivity interferometer for measurement of phase-shift and other applications *NBS Circular* **581** (1957) 51-59
- [34] Dorenwendt K Interferenteile Messung von nicht angeschobenen Endmaßen *PTB Annual Rep.* (1973) 121

CHAPTER 4

ALIGNMENT, COHERENCE AND OPTICAL TESTING

*“Every physicist thinks he knows what a photon is,
I spent my life trying to find out what a photon is, and I still don’t know”
A Einstein*

4.1 ALIGNMENT OF THE INTERFEROMETER

Before the interferometer can be used to perform measurements it must be properly aligned. There are two stages to the alignment procedure - the initial alignment which must be performed whenever the interferometer has been disturbed or a component replaced, and the accurate alignment which is performed periodically or after thermal expansion measurements have been made, when the alignment may have drifted. Although length measurements are possible with the interferometer only approximately aligned, they will be in error due to the Abbe effect: the measurement beam will not be travelling parallel with the length to be measured. The final alignment ensures that the measurement beam travels parallel to the axis of the bar, so minimising this error.

4.1.1 Approximate alignment of interferometer

Depending on how the interferometer has been adjusted or modified, only certain parts of the initial alignment may be necessary.

4.1.1.1 Laser beam launching into fibres

The output beam of each laser is focused via a system of lenses into the core of a single mode optical fibre (see § 3.2.2). This requires careful alignment of the end of the fibre with the focal spot of the microscope objective lens. Firstly, the fibre is positioned by eye close to the focal spot by adjusting the x - y positioners. Next, the correct focusing is achieved by careful longitudinal positioning of the fibre ferrule until the laser speckle observed on the mount of the objective lens is at its greatest. The speckle is due to interference between parts of the beam reflected from the surfaces of the ferrule and the

fibre. The size of the speckle is largest when the beam is focused on the ferrule/fibre surface. Finally, the fine adjusters are used to position the fibre on the focal spot and to adjust the focus. The output beam from the fibre is monitored by eye or by the CCD camera in the interferometer until the output beam reaches maximum intensity. The laser launch is then correctly aligned.

4.1.1.2 Component positioning in interferometer

The baseplate of the interferometer has carefully positioned holes for mounting the optical components and their holders. By following the schematic diagram of the interferometer it is simple to fix the components in the right position by bolting them to the baseplate. Exceptions to this are the final 45° mirror, the TV camera and the reference mirror assembly, which all have some degree of freedom in their positioning to allow for adjustment.

4.1.1.3 Fibre positioning in collimator

The optical fibre bundle inside the brass ferrule must be positioned on the axis of the collimator lens. This is achieved by placing an optical flat against the upright surface of the collimator lens mount and adjusting the position of the fibre until the return spot of this autocollimator arrangement is coincident with the source. The lens holder was machined so that the upright is parallel to the rim against which the lens rests. Provided the source is within 1 - 2 mm of the axial focal point, the beam will be sufficiently collimated.

4.1.1.4 Reference mirror alignment

After aligning the collimator, the reference mirror is aligned with the collimated beam. The collimated beam reflects off the collimator mirror and passes through the beamsplitter and is directed by a 45° mirror onto the reference mirror (see figure 3.17). The 45° mirror is adjusted until the beam reflected from the reference mirror is aligned with the interferometer axis. This axis is the path of the axial ray from the source, through the centre of the collimator lens, through the beamsplitter to the reference mirror, and the reverse path. This axis will thus be a normal to the reference mirror surface, once the latter has been aligned.

4.1.1.5 Measurement beam alignment

Having defined the reference axis by adjusting the reference mirror assembly, the other beam reflected at the beamsplitter is the measurement beam which must be aligned with the bar to be measured. This is easily accomplished since the interference fringes will be ‘fluffed out’ when the reference and measurement beams are co-axially aligned.

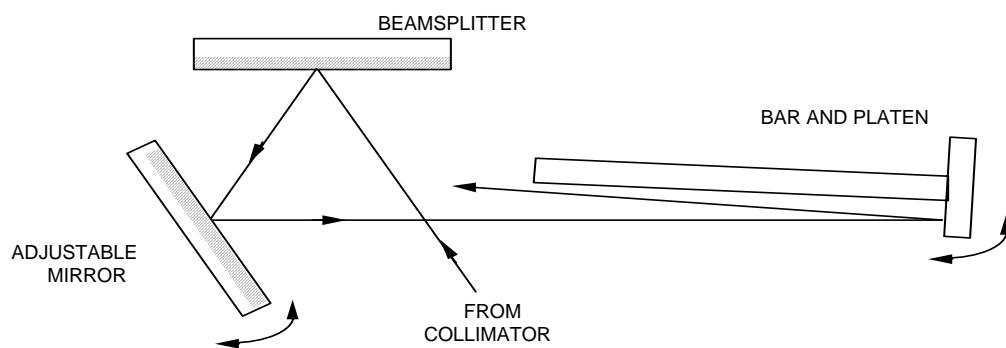


Figure 4.1 - Aligning the measurement beam with the length bar

Once the length bar supports have been adjusted so that the bar is approximately aligned with the measurement beam, the mirror in the measurement beam can be used to accurately align the beam with the axis of the bar, as shown in figure 4.1.

4.1.1.6 Alignment of length bars with measurement beam

Having aligned the measurement beam with the axis of one bar, the support carriage is translated until another bar is in the measurement position. The bar supports are adjusted to approximately align the bar with the measurement beam. Any remaining mis-alignment can be removed at the time of measurement by adjusting the measurement beam mirror using the PZTs.

4.1.1.7 Alignment for double-ended interferometry

For double-ended measurements, the interferometer is aligned as follows.

Firstly, the collimator is aligned by using the autocollimation technique, with an optical flat placed against the collimator mount and the fibre position adjusted until the return spot is incident on the fibre end. When the fibre is correctly positioned, the reverse beam will overfill the end of the fibre due to aberrations and diffraction and will thus illuminate the cladding - this can be observed as an increased glow in the fibre. This technique is also suitable for checking the focal positioning.

Next, the reference arm is aligned with the collimated beam by adjusting the tilt of the reference mirror until the return spot is autocollimated onto the end of the fibre.

The bar is placed on the supports, this time positioned at the Airy points of the bar because there is no platen wrung to one end. The measurement beam is adjusted until the fringes are fluffed out on the face of the bar. The axes of the bar, the measurement beam and the reference beam are now aligned, assuming the bar is not out of tolerance on the squareness of the end face with the axis of the bar.

The roof mirror system is inserted. This produces 3 additional return spots at the source corresponding to the beams returned from the front face of the bar and the two oppositely propagating beams which travel around the bar. The return spots of these latter two beams are symmetrically located either side of the spot from the end of the bar, which will be displaced vertically from the source until the vertical tilt of the two mirrors is corrected. Adjustment of the mutual orthogonality of the mirrors directs the two symmetrical spots onto the source.

When the roof mirrors are not orthogonal, the fringes in the background of the image will exhibit a ‘V-shaped’ characteristic, as shown in figure 4.2. This will give rise to extra tilt in the interferogram of the fringes on the rear face of the bar.

When the roof mirrors are adjusted for mutual orthogonality, the tilt of the pair of mirrors is adjusted until the fringes in the background are fluffed out. The interferometer is then completely aligned for double-ended measurements. Any difference in tilt between the two images of the ends of the length bar is due to parallelism errors in the bar. Once the roof mirrors have been set orthogonal with each other, only tilt of the pair of mirrors from the vertical plane will influence the fringes in the image. The imaging of the front face of the bar is unchanged from the single-ended case, except for the magnification.

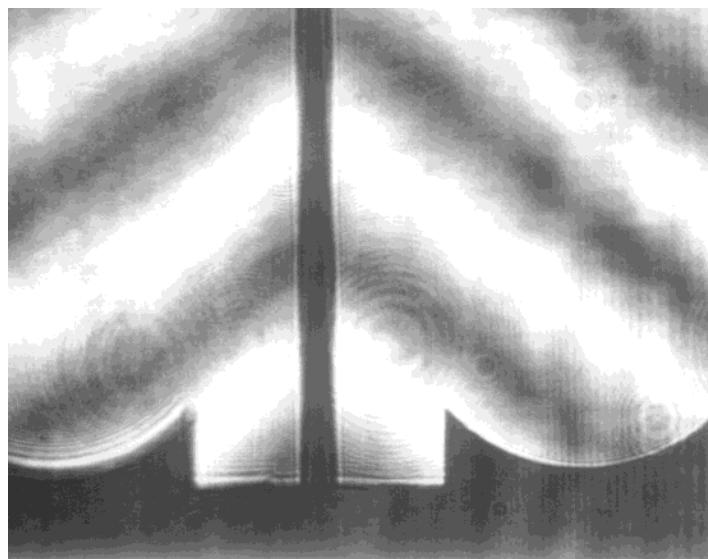


Figure 4.2 - Incorrect adjustment of roof mirror orthogonality (bar removed for clarity)

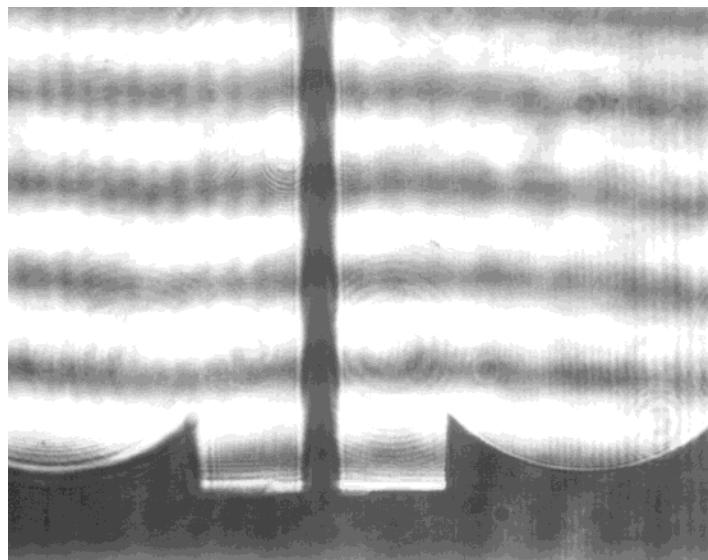


Figure 4.3 - Correct adjustment of roof mirror orthogonality (bar removed for clarity)

4.1.2 Accurate alignment of the interferometer

The accuracy of length measurements is critically dependent on the accuracy of alignment of the interferometer. To produce interference fringes visible on the monitor, the reference and measurements beams must be aligned with each other to within approximately 1 arcmin. However, even with fluffed out fringes, the alignment may still be in error and the length measurement inaccurate due to the reference and measurement beams not being aligned with the axis of the bar being measured.

4.1.2.1 Cosine error due to measurement beam mis-alignment

Consider the bar and platen in figure 4.4.

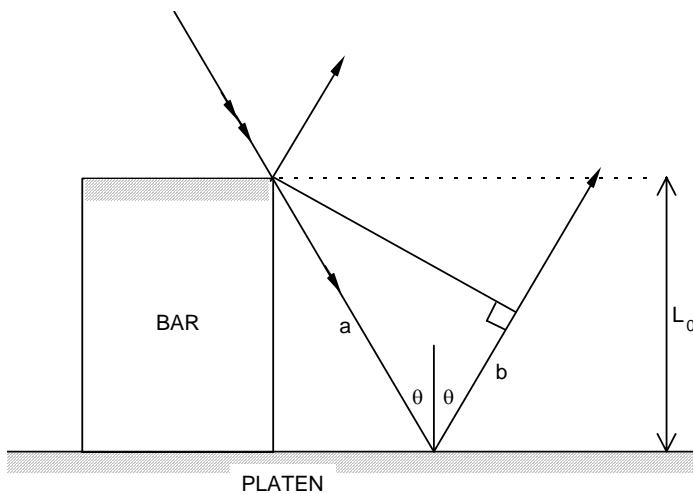


Figure 4.4 - Cosine error of measurement beam

The measured length is half the difference in optical path between the two beams, ΔOPD .

$$\begin{aligned}
 \Delta OPD &= a + b \\
 a = \frac{L_0}{\cos \theta} \quad b = a \cos 2\theta &= \frac{L_0 \cos 2\theta}{\cos \theta} \\
 \therefore \frac{\Delta OPD}{2} &= \frac{a + b}{2} = \frac{L_0}{2 \cos \theta} [1 + \cos 2\theta] \\
 &= \frac{L_0}{\cos \theta} [\cos^2 \theta]
 \end{aligned}$$

measured length = $L_0 \cos \theta$ (4.1)

Hence, unless the measurement beam is aligned with the axis of the bar, the measured length will be subject to a length dependent cosine error. Note, the effect of the end faces of the bar not being perpendicular to the axis of the bar are dealt with in chapter 10 - for now it will be assumed that the bar is a right circular (or rectangular) cylinder.

4.1.2.2 Alignment of the three interferometer axes

There are three axes which must be aligned in the interferometer: the axis of the bar, the reference beam axis and the measurement beam axis. To align all three, a common point of reference must be used: this is the source spot. First, the collimator is aligned using the method described above. Next, the reference arm is aligned with the collimator, followed by alignment of the measurement beam (including the bar). Thus the initial alignment of the reference beam with the collimator is important. A special technique was developed to align this beam.

4.1.2.3 Two-fibre autocollimation technique

The system uses three single-mode optical fibres which have had the buffer coating removed from both ends (see § 3.2.2). At one end the fibres are cemented into a tight bundle. The other end of each of the fibres is individually mounted and polished, see figure 4.5. Each fibre in the bundle can serve two functions; it can act as the light source for the collimator when light from a laser is focused into the fibre core, and secondly the fibre can be used to detect the return spot, when used in an autocollimation arrangement, for which another fibre is used as the light source, see figure 4.6.

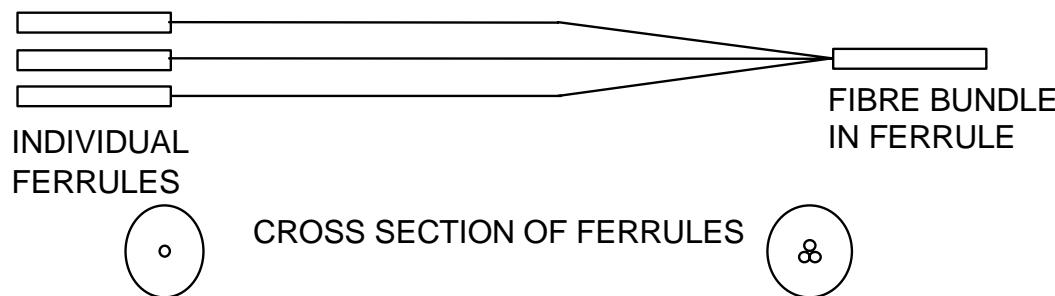


Figure 4.5 - Three fibre system

In principle it is possible to use this technique with just one fibre acting as both source and detector though the extra optical components required, such as beamsplitters or couplers, could introduce losses which would make the detection of the return spot more difficult.

When used in the autocollimation arrangement of figure 4.6, the reference mirror of the interferometer is used to reflect the beam back to the source where one of the fibres in the bundle is used to detect the return spot. The fibre bundle is moved in three orthogonal directions and the intensity of the light incident on the detector fibre is monitored using a photodetector placed at the output end of the fibre, after it has been removed from the laser launch optics. When the detected intensity is maximised, the source and detector fibres are symmetrically positioned on either side of the axis of the interferometer. The off-axis position of the source is then half the separation of the fibres, which is typically $50\text{ }\mu\text{m}$. This technique aligns the collimator with the reference arm of the interferometer. The expected obliquity error from this system is less than 5.6×10^{-10} , or 0.56 nm in 1 m . After alignment, the detector fibre may be replaced and used to launch a third laser source.

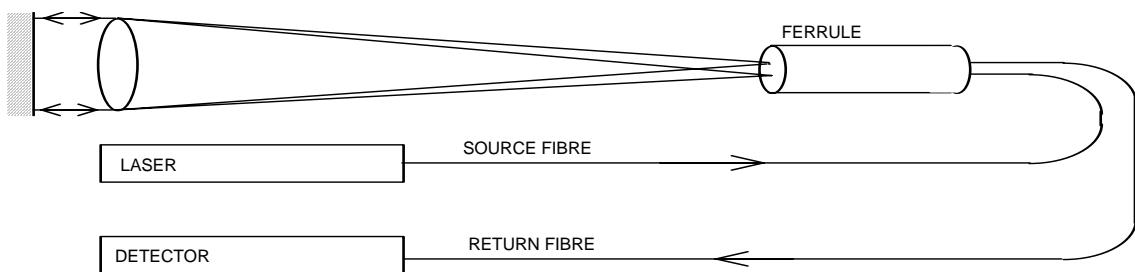


Figure 4.6 - Autocollimation arrangement

Evaluation measurements have been made using this system illuminated by the 633 nm red laser. Figure 4.7 shows the peak in the detected intensity as the fibre bundle was positioned radially and axially. These results were repeatable after coarse adjustment over several millimetres of travel.

Assuming the achromatic collimator lens to be diffraction limited, the expected central maximum (Airy disc) of the return spot diffraction pattern should be $\sim 25\text{ }\mu\text{m}$ in diameter [1] and should result in a peak of width $\sim 20\text{ }\mu\text{m}$ when a $4\text{ }\mu\text{m}$ diameter fibre is scanned across the moving diffraction pattern, as occurs when the fibre bundle undergoes radial motion. This can be seen in figure 4.7(a). It is thought that the non-symmetrical peaks in the observed data are due to cross-talk from the adjacent fibre which becomes partially illuminated.

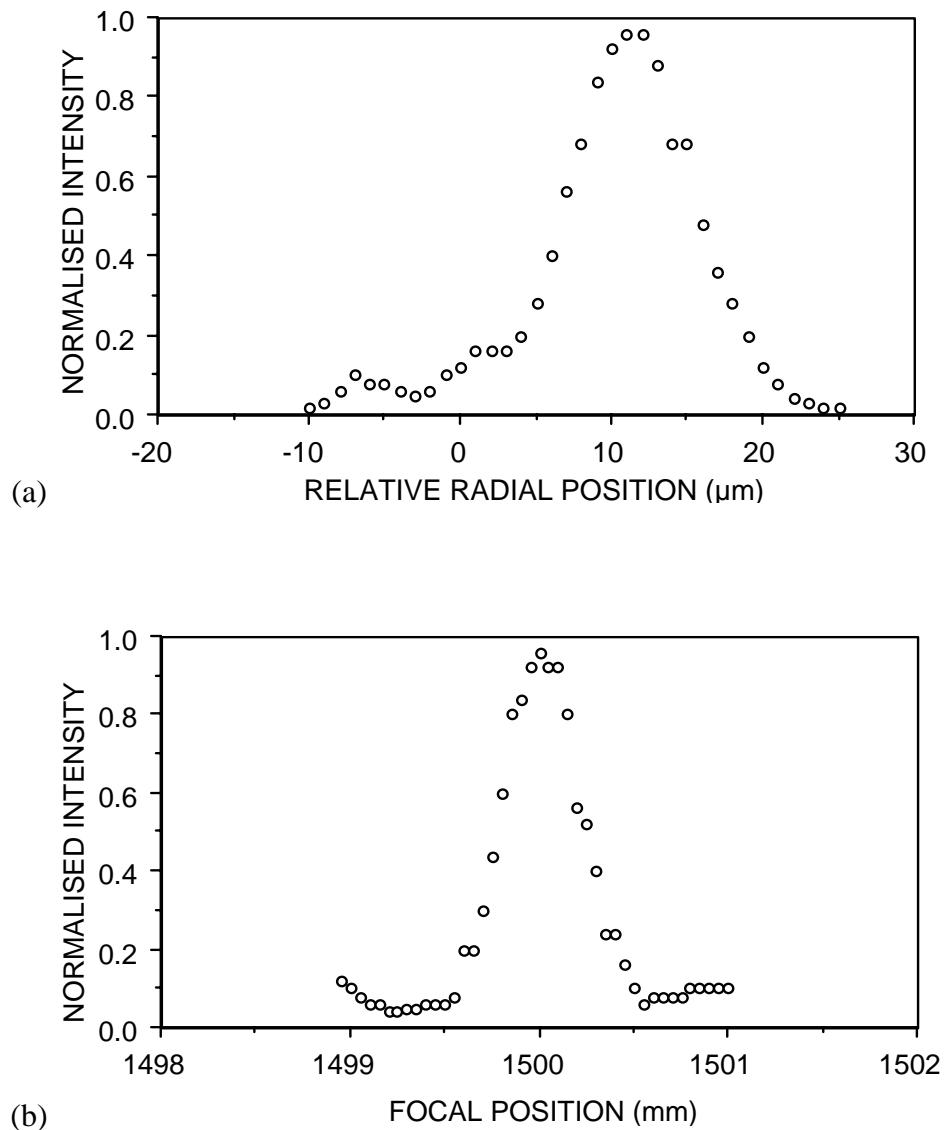


Figure 4.7 - Detected intensity during (a) radial positioning and (b) axial positioning of the fibre bundle (normalised units)

When diffraction theory is applied to an un-aberrated circular pupil with defocus it predicts minima in the diffraction pattern, spaced at 1.1 mm along the focal axis. The results shown in figure 4.7(b) are consistent with the theory.

The single-mode fibre system thus provides a simple, efficient, speckle free light source for an interferometer. The autocollimation arrangement using one fibre as a detector allows accurate repositioning of the light source, allowing the collimator beam to be aligned with the interferometer axis, whilst minimising the obliquity effect due to the source. However, with this system the return intensity is low and requires a sensitive detector. It is easier to observe the additional light scattered in the fibre cladding, as described in § 4.1.1.7.

4.1.3 Obliquity effect due to position and size of light source

The light source for the interferometer is the end of an optical fibre, positioned at the focal point of a 100 mm diameter achromat. The source diameter is that of the fibre, which is approximately 4 μm . The fibre is positioned to be nominally on axis, at the correct focal length. However small errors in the positioning can lead to the light source being off axis. This, as well as the finite size of the source contribute errors in the measurement, which can be regarded as obliquity errors, *i.e.* they cause the apparent length of the measured object to be slightly different from the true length. These effects can be removed by using a correction factor. Here, it is shown that the correction factors for the interferometer are very small, due to the design of the collimator, and can be neglected.

When the source is positioned off-axis, the effect can be thought of as causing a small angular deviation, θ , of the measurement path with respect to the object axis. For small θ , there is a correction factor per unit length, C_1 , which is $(1-\cos\theta)$, which is approximately $\theta^2/2$. For an aperture of negligible size, at distance s off axis, the correction factor is thus

$$C_1 = \frac{s^2}{2f^2} \quad (4.2)$$

where f is the focal length of the lens. Assuming that the positioning error of the twin-fibre system is $\pm 50 \mu\text{m}$, then for the interferometer, the correction factor for off-axis positioning is

$$C_1 = 5.6 \times 10^{-10}$$

There is also an obliquity effect due to the finite size of the source [2]. This can be considered as the sum of the effects of all infinitesimally small elements which constitute the source. This factor, C_2 , is thus

$$C_2 = \frac{\int_0^{2\pi} \int_0^r \frac{x^2}{2f^2} x d\phi dx}{\int_0^{2\pi} \int_0^r x d\phi dx}$$

$$C_2 = \frac{r^2}{4f^2} \quad (4.3)$$

For the single mode fibres used, $r \sim 2 \mu\text{m}$, thus $C_2 = 4.4 \times 10^{-13}$ and is hence negligible. The total obliquity effect is thus 5.6×10^{-10} or a length measurement error of $\pm 0.56 \text{ nm in } 1 \text{ m}$.

4.1.3.1 Obliquity effect due to source size - full derivation

Although it is possible to have a correctly aligned interferometer, the size of the source will affect the measurements of phase performed in the interferometer.

Bruce [2] considers the extra phase shift introduced to the system due to the source size as an obliquity effect. Consider the circular aperture on axis, in figure 4.8.

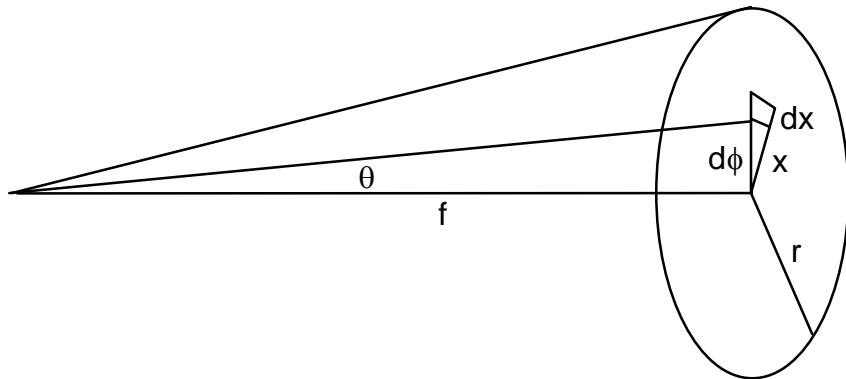


Figure 4.8 - Obliquity effect due to a circular aperture on axis of interferometer

The intensity of the interference fringes is given by

$$I = \cos^2 K$$

where K = phase difference/2

i.e.

$$K = 2\pi L / \lambda$$

where L is the length being measured. For an element of the source at angle θ to the axis, size $d\phi dx$

$$\delta I = \cos^2(K \cos \theta) x d\phi dx \quad (4.4)$$

For the total effect from the source, integrate over aperture

$$I = \int_0^{2\pi} \int_0^r \cos^2(K \cos \theta) x d\phi dx$$

Since θ is small, we can approximate $\cos \theta$ as $1 - \frac{x^2}{2f^2}$

$$\begin{aligned} I &= 2\pi \int_0^r \cos^2 \left(K - \frac{Kx^2}{2f^2} \right) x dx \\ &= 2\pi \int_0^r \frac{1}{2} \left[\cos \left\{ 2K \left(1 - \frac{x^2}{2f^2} \right) \right\} + 1 \right] x dx \end{aligned}$$

now substituting

$$\begin{aligned} u &= 2K \left(1 - \frac{x^2}{2f^2} \right) \\ du &= \frac{2Kxdx}{f^2} \end{aligned} \tag{4.5}$$

$$I = 2\pi \int \frac{1}{2} [\cos u + 1] \frac{f}{2K} du \tag{4.6}$$

$$I = \frac{\pi r^2}{2} - \frac{\pi f^2}{2K} \left\{ \sin 2K \cos \frac{Kr^2}{f^2} - \cos 2K \sin \frac{Kr^2}{f^2} - \sin 2K \right\}$$

Substituting for the area of the source A , the obliquity angle θ and the phase factor Δ

$$\begin{aligned} A &= \pi r^2 & \Delta &= \frac{K\theta^2}{2} & \theta &= \frac{r}{f} \\ I &= \frac{A}{2} \left[1 + \frac{\sin \Delta}{\Delta} \cos(2K - \Delta) \right] \end{aligned} \tag{4.7}$$

Thus the fringes are symmetrical, but displaced by the phase factor Δ from their normal positions. The fringe contrast is also reduced by the factor $\text{sinc}(\Delta)$.

However for small Δ , approximate $\frac{\sin \Delta}{\Delta} \approx 1$, thus

$$\begin{aligned} I &= \frac{A}{2} [1 + \cos(2K - \Delta)] \\ &= A \cos^2 \left(K - \frac{\Delta}{2} \right) \end{aligned} \tag{4.8}$$

This equation describes normal \cos^2 fringes from an aperture area A , but displaced in phase by $\Delta/2$ compared with those obtained using a point source (on axis). This is the same as the factor C_2 above.

For $0 \leq \Delta \leq \pi$ the factor $\sin\theta/\theta$ in equation (4.7) is positive and the fringe displacement is given by $\Delta/2$. At $\Delta = m\pi$ ($m=1,2,3\dots$) equation (4.7) predicts that the fringes vanish. For the interferometer, with $r = 1.8 \mu\text{m}$, $f = 1.5 \text{ m}$, $\lambda = 633 \text{ nm}$, the first zero of fringe visibility occurs at $L = 400 \text{ km}$. Since $L < 2 \text{ m}$, the fringes should have good visibility for all sizes of length bar.

Whilst it may not be obvious that a symmetrical source, positioned on axis can produce non-symmetrical shifts in the interference fringes, these effects have been observed by Bruce [2] with good agreement (0.001 fringe) between the measured and predicted fringe shifts up to a shift of approximately 0.25 fringe for a 0.78 mm diameter pinhole. The results of these calculations have been confirmed by Thornton [3] using a different analysis. It is easier to see the reason for this shift by noting that for any point not on axis, the corresponding beam from this point will travel at an angle to the axis of the bar being measured (obliquity angle) and must therefore measure the length short by the usual cosine factor. Thus all elements of the source lying on an annulus at a particular radius will all contribute an obliquity error of the same sign and magnitude. The results derived above take into account all such annuli.

4.1.4 Collimation check using a shearing plate interferometer

A shearing plate interferometer can be used to translate wavefront curvature into rotation of straight fringes [4]. When placed into a properly collimated beam with no aberrations, parallel straight fringes are observed which are parallel to the reference line of the shearing plate. The radius of curvature, R , of an incorrectly collimated beam can be measured using a shearing plate

$$R = \frac{s\delta}{\lambda \sin \theta} \quad (4.9)$$

where s is the shearing distance, δ is the fringe spacing and θ is the fringe rotation from the reference line, measured on the sheared image. Measured values in the collimated beam of the interferometer were $s = 7 \text{ mm}$, $\delta = 7 \text{ mm}$, for $\lambda = 612 \text{ nm}$. A value for θ was estimated to be 0.08 rad (from trigonometry), this places a lower limit on R of 1000 m. The effect of the wavefront curvature on the obliquity is derived as follows.

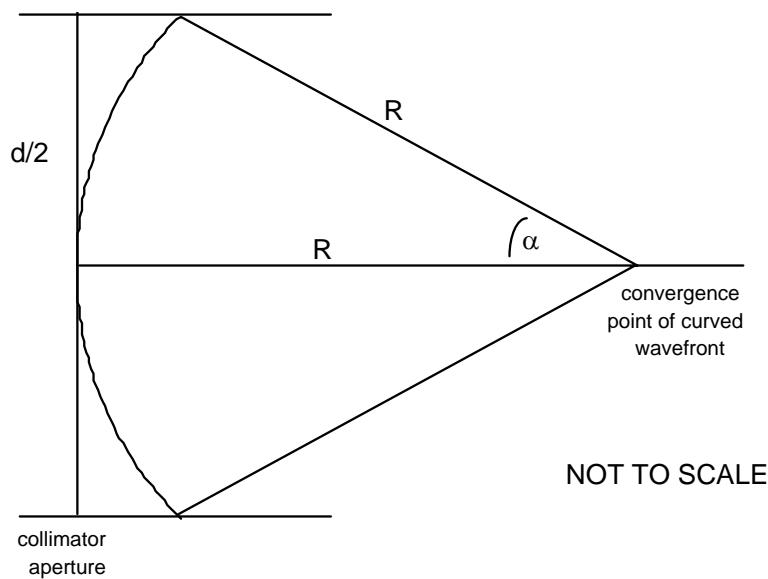


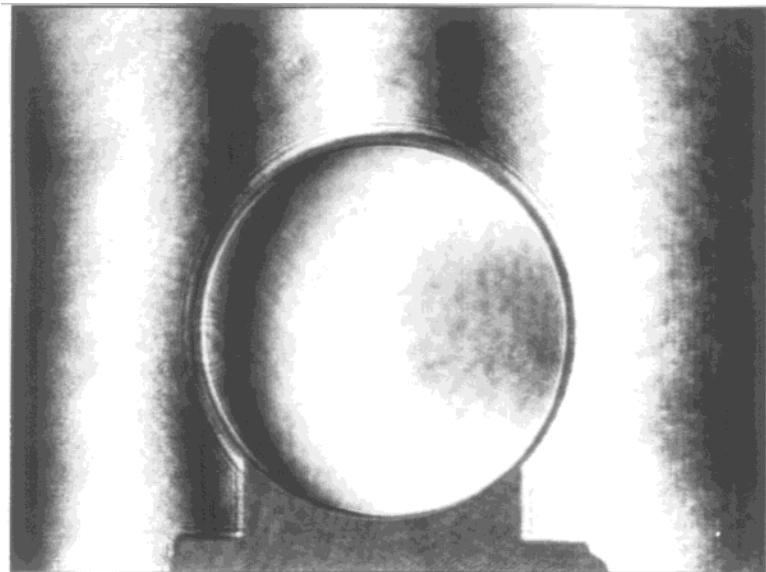
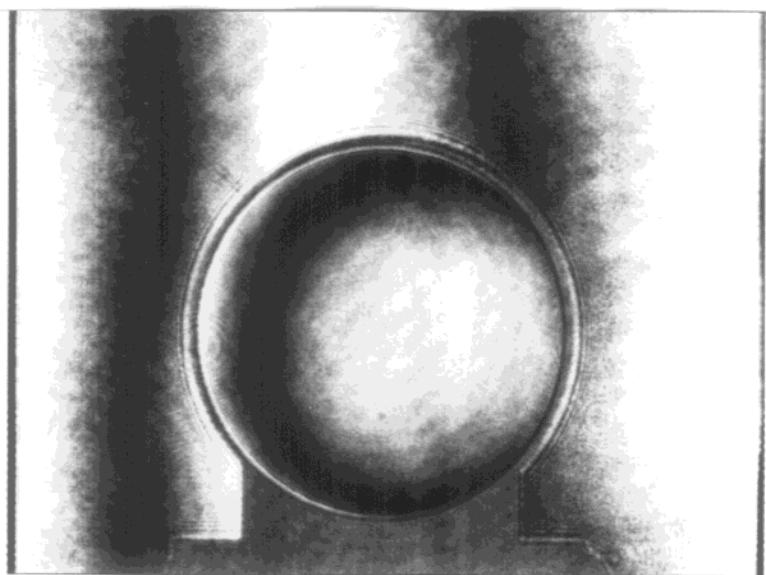
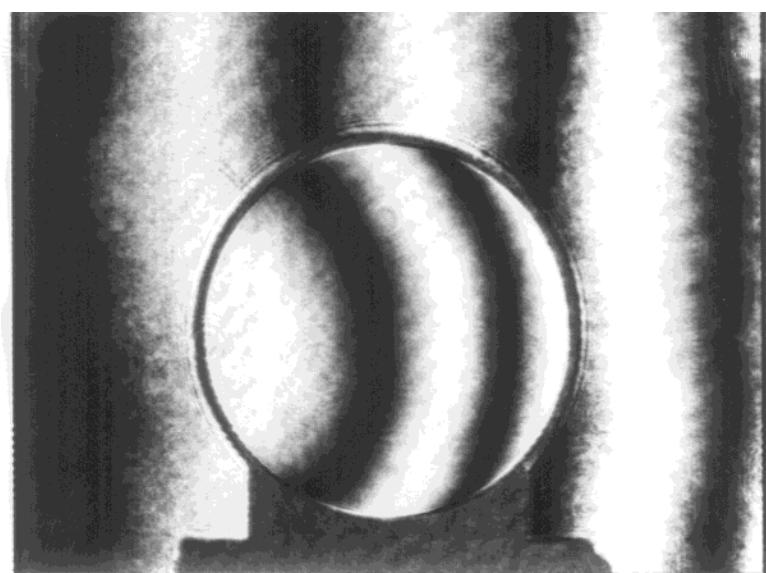
Figure 4.9 - Convergence of un-collimated wavefront

$$\alpha \approx \frac{d}{2R} \quad (4.10)$$

For a value for d of 80 mm, $\alpha = 4.0 \times 10^{-5}$. This causes an obliquity effect of magnitude $\alpha^2/2$, which is 1×10^{-9} . The residual wavefront curvature is due to spherical aberration of the achromatic collimator lens, see § 4.1.7.2.

4.1.5 Tilt in the measurement beam

There is an observed change in tilt in the interferogram of approximately + 6 fringes across the image field, when the green laser illuminates the interferometer compared to the red, but only - 1 fringe between the orange and red illuminations (see figure 4.10). A possible source for this differential tilt has been identified, and an estimate of both the magnitude of the tilt and its corresponding obliquity effect are calculated.

(a) $\lambda = 633 \text{ nm}$ (b) $\lambda = 612 \text{ nm}$ (c) $\lambda = 543 \text{ nm}$ **Figure 4.10** - Difference in tilt between three wavelengths (a) 633 nm, (b) 612 nm, (c) 543 nm

4.1.5.1 Prismatic dispersion at the beamsplitter

The beamsplitter is wedged at 0.5° to prevent multiple reflections from the non-coated side from interfering with the correct reflection. There is no compensator plate in the interferometer, and thus the beamsplitter acts as a dispersive prism of apex angle $\gamma = 0.5^\circ$.

The angular dispersion $d\delta$ for a wavelength change $\delta\lambda$ is given by [5]

$$d\delta = \frac{-2 \sin(\gamma/2) dn'}{\sqrt{1 - n'^2 \sin^2(\gamma/2)}} \quad (4.11)$$

For the material of the beamsplitter, fused silica, dn' can be found from measurements of $n'(\lambda)$:

$$\begin{aligned} n(632.8 \text{ nm}) &= 1.45702 \\ n(546.1 \text{ nm}) &= 1.46008 \end{aligned}$$

thus $dn' = 0.00306$ for $\delta\lambda = 633 - 546 \text{ nm}$, which is approximately the difference between the red and green. This results in

$$d\delta = -2.7 \times 10^{-5} \text{ rad}$$

This additional tilt produces an extra number of fringes across the image area (width approximately 45 mm) given by

$$\frac{2 \times 45 \times 10^{-3} \tan(d\delta)}{633 \times 10^{-9}}$$

$$\sim 4 \text{ fringes}$$

This is only an approximate calculation because values of dn' for the actual material of the beamsplitter will be slightly different from those above, however the direction of the additional tilt which is observed is horizontal. This corresponds to the same plane in which the beamsplitter is wedged and is thus a likely candidate for the extra tilt.

4.1.5.2 Methods for compensation of tilt

The tilt could be corrected by use of a wedged compensator, matched to the beamsplitter. However this would be prone to further spurious reflections and

wavefront aberrations and since the obliquity effect of the additional tilt is negligible, it is not worth correcting in the interferometer, as the software removes any tilt from the final phase maps.

An alternative solution would be to use a system of wedged beamsplitter and compensator plate, angled such that the ghost reflections are trapped inside the beamsplitter by successive total internal reflection [6].

4.1.6 Chromatic aberration - tolerance on collimator focal position

Chromatic aberration in the collimator lens leads to a variation in the position of the focal point with respect to wavelength. The variation between the red and green ends of the spectrum is approximately [7] $f/2000$ where f is the focal length of the collimator achromat, see figure 4.11.

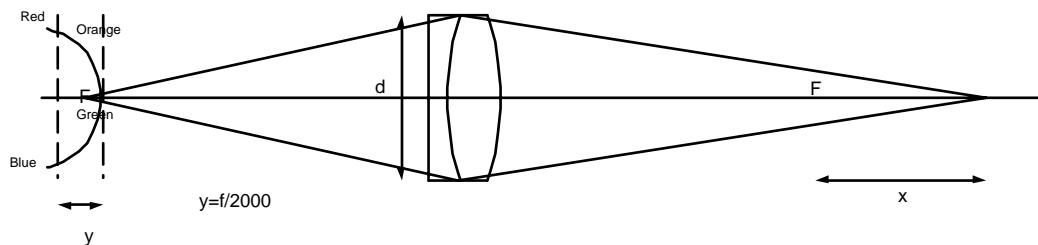


Figure 4.11 - Chromatic dispersion - effect on focal length of collimator

For the achromat used in the Primary Length Bar Interferometer, $f = 1500$ mm. The manufacturer's data states that the variation in focal length between wavelengths 633 nm and 588 nm is 0.47 mm (*i.e.* $y = 0.47$ mm) or $f/3200$. The effect of the afocal positioning of any of the optical fibres can be calculated as follows.

$$xy = f^2$$

and

$$y = f/3200$$

hence

$$x \sim 4800 \text{ m}$$

Consider one ray of a convergent beam, focal length 4800 m, travelling at an angle θ to the axis of a length bar. The error in the measured length of the bar due to the angle of the beam will be given by

$$\Delta L \approx \frac{L\theta^2}{2}$$

and

$$\theta = \frac{d}{2(F+x)}$$

substituting gives

$$\Delta L = 4.4 \times 10^{-11} L$$

If a tolerance is imposed such that $\Delta L < 10^{-9} L$ then it is simple to show that the tolerance on the focus of the collimator is 2.5 mm, which is easily achieved - as shown above, the maximum departure for the achromat used is approximately 0.47 mm.

4.1.7 Optical component quality and spherical aberration

4.1.7.1 Quality of optical components

Measurements of surface quality and subsequent wavefront aberrations of the most important optical components can be found in Appendix B. The majority of the wavefront aberration was found to be spherical aberration.

4.1.7.2 Effect of spherical aberration in collimator

If it is assumed that the wavefront of the interferometer measurement arm contains aberrations, of which the main component is spherical aberration, then the effect of this on the measured length of the length bar can be calculated as follows.

Let the wavefront be of the form $W(x) = ax^4$, where x is a co-ordinate across the wavefront, see figure 4.12.

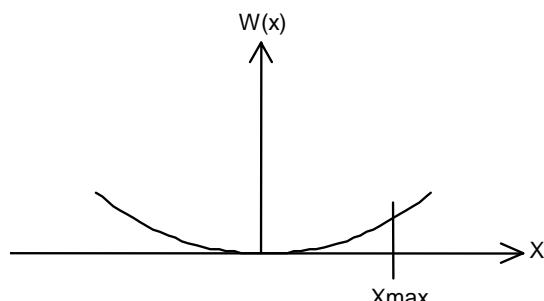


Figure 4.12 - Spherically aberrated wavefront

then the angular aberration will be given by $\theta \approx \frac{\partial W}{\partial x}$

$$\therefore \theta \approx 4ax^3 \quad (4.12)$$

If we now examine the paths travelled by two beams, one un-aberrated, the other aberrated, at an angle θ to the other, where θ is given by the above expression.

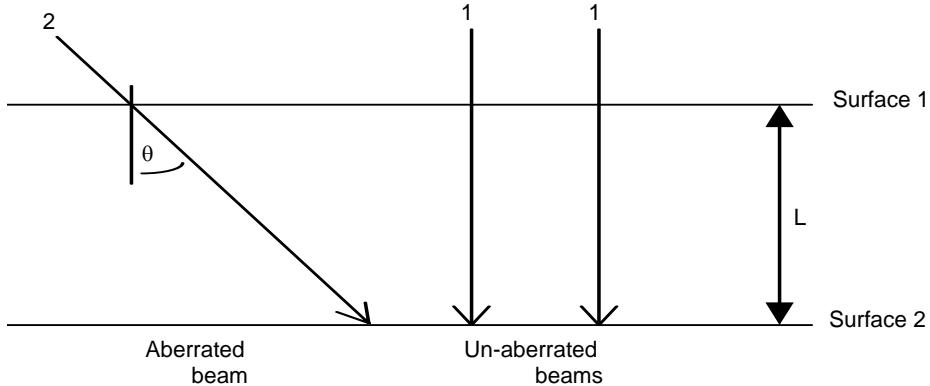


Figure 4.13 - Interference between two spherically aberrated beams

These two beams will have phases

$$\phi_1 = \frac{2\pi}{l} 2L \quad \phi_2 \approx \frac{2\pi}{l} \frac{2L}{\cos q}$$

The phase difference between the two beams will be $\phi_1 - \phi_2$,

$$\phi_1 - \phi_2 = \frac{4\pi L}{l} \left(1 - \frac{1}{\cos q} \right)$$

$$i.e. \quad \Delta\phi \approx \frac{2\pi L q^2}{l} \quad (4.13)$$

From (4.12) and (4.13),

$$a = \sqrt{\frac{\lambda \Delta\phi}{2\pi L}} \frac{1}{4x^3} \quad (4.14)$$

Substituting $\lambda = 633$ nm, and $x = 40$ mm, the radius of the collimated beam, the path length, $L = 1.5$ m (for a 1.5 m bar), and from the Zygo test measurements of the achromat, $W(x) = 0.15 \lambda$, a is found to be 0.0371 m^{-3} . Substituting for a gives a value for $\Delta\phi$ of $\Delta\phi = 0.0013$, or 1/770 fringe. This is equal to 0.4 nm, and hence is a small

systematic error. This value is similar to the value of 1×10^{-9} obliquity error obtained from the estimate of residual wavefront curvature in § 4.1.4.

4.1.8 Effect of squareness of length bar on measured length

The length of a length bar is defined in BS 5317 :

length. This is defined, with the bar mounted horizontally and referred to the standard reference temperature of 20°C , as the distance from the centre of one of its faces to a flat surface in wringing contact with the opposite face, measured normal to the surface.

For a bar which is not perfectly square, *i.e.* the end faces are not both perpendicular to the axis of the bar and also parallel with each other, this can lead to differences between the defined and measured lengths of a bar, depending on how the bar is measured.

Consider a non-parallel, singularly non-square length bar, with defined length L_d and another length L_m .

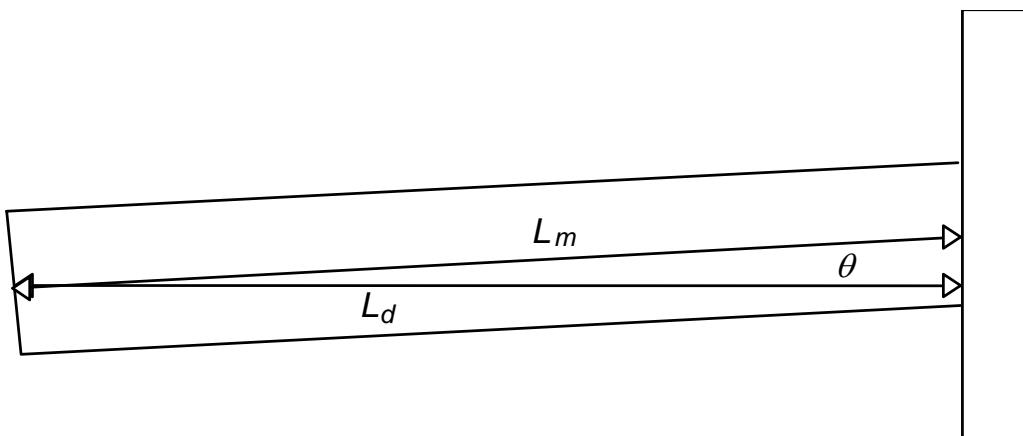


Figure 4.14 - Non-square, singularly non-parallel length bar

Because the interferometer is set up with the platen surface normal to the measurement beam (± 2 to 3 fringes of tilt), the length measured by the interferometer is L_d , the defined length of the bar. The NPL Length Bar Machine measures the mechanical central length, L_m , which will be different to the defined length by a factor of size $\cos\theta$, or approximately $\theta^2/2$.

The size of the angle θ will depend on the squareness of the bar. According to the standard, bars should be within $1.2 \mu\text{m}$ of squareness, for bars up to 400 mm in length, and within $2.5 \mu\text{m}$ of squareness for longer bars. These values are equivalent to 4 and 8 fringes of squareness error, respectively. By converting these values to angles, it can be

shown that these are equivalent to a length measurement error (for the Length Bar Machine) of $1.6 \times 10^{-9} L$ for bars up to 400 mm, and $6.4 \times 10^{-9} L$ for longer bars.

Note that these errors are for the Length Bar Machine and not for the interferometer, as the latter measures the length of bars in accordance with the definition of length in BS 5317.

4.2 COHERENCE IN THE INTERFEROMETER

The length measurements made in the interferometer are measurements of phase across a relatively large aperture, up to 8 cm in diameter. This requires a high level of both temporal coherence and in the case of double ended interferometry, spatial coherence, as well as high quality optical components with minimum wavefront aberrations.

Temporal coherence is of particular importance as the interferometer has relatively large path lengths and the accuracy required of the length measurements dictates a narrow laser linewidth which is closely linked to temporal coherence. Analogous to the coherence time of the light emitted from the source is the coherence length (as opposed to the spatial coherence measured across the beam pupil). The coherence length must be at least equal to the total path length travelled by the beams before striking the detector array surface.

Although spatial coherence is not so important when using the interferometer in its conventional Twyman-Green arrangement due to the common path, non-sheared optical arrangement, it is however very important when making double-ended measurements (see § 3.3.3). In this arrangement interference is formed between different parts of the beams, some of which have been spatially inverted, *i.e.* sheared. This demands a high degree of spatial coherence between all points in the beam as well as the temporal coherence described above. This can be visualised as a *coherence volume* within which the beam must maintain both temporal and spatial coherence. In wave terms both the magnitude and direction of the wave-vector \underline{k} must be well defined and invariant. For the interferometer the beam must be spatially coherent across the maximum shearing distance of approximately 8 cm and along a path length of up to 7 m, making a coherence volume of 0.035 m^3 .

The factors affecting coherence will now be examined and estimates of the spatial coherence and temporal coherence length will be made. For a detailed development of the concept of coherence see Hopkins [8,9,10,11].

4.2.1 Temporal coherence

The temporal coherence of a source is a measure of the spread of frequencies (or wavelengths) emitted by the source. A typical quasi-monochromatic source such as a vapour lamp emits frequencies in the range $\omega_0 \pm \varepsilon / 2$ of equal amplitude and random phase. This can be illustrated through the use of the *temporal coherence function*, $\gamma(\tau)$.

$$\gamma(\tau) = \frac{\overline{A(t)A^*(t+\tau)}}{\sqrt{\overline{|A(t)|^2}\overline{|A(t+\tau)|^2}}} \quad (4.15)$$

where $A(t)$ is the amplitude at time t , $A(t+\tau)$ is the amplitude at time $t+\tau$, and the bar above the symbols represents a mean value over a long interval. For a quasi-monochromatic source γ is unity for small values of τ but then decreases as τ increases. In fact $\gamma(\tau)$ is the Fourier transform of the spectral intensity (Wiener-Kinchine theorem) and is also related to the visibility of interference fringes, given by Michelson's expression

$$V = \frac{I_{\max} - I_{\min}}{I_{\max} + I_{\min}} \quad (4.16)$$

When interference is formed between two beams of equal intensity, with one beam delayed by τ , then $V = |\gamma(\tau)|$.

To obtain good contrast fringes in the interferometer, $\gamma(\tau)$ must be close to unity for τ of the order of 2.3×10^{-8} seconds (time taken for beam to travel 7 m) in other words the coherence time of the source, τ_c must be longer than 10^{-8} s. For most standard light sources such as spectral lamps, the coherence time τ_c is approximately 10^{-9} s due mostly to linewidth broadening. There are two main sources of broadening: Doppler broadening and collision broadening.

The effect of Doppler broadening is to spread the line-shape into a Gaussian profile [12] with half width

$$\sigma = \omega_0 \sqrt{\frac{k_B T}{mc^2}} \quad (4.17)$$

where k_B is Boltzman's constant, m is the mass of the particle undergoing the transition, T is the temperature of the gas (in K) and ω_0 is the central frequency of the radiation, or in terms of wavelengths

$$\sigma = \lambda_0 \sqrt{\frac{k_B T}{mc^2}} \quad (4.18)$$

As an example, for a standard krypton lamp with $\lambda_0 = 5.6 \times 10^{-7}$ m at 80 K, the half-width is 1.6×10^{-13} m, or 3 parts in 10^7 .

The effect of collision broadening is much greater than that of Doppler broadening at atmospheric pressure. This arises due to collisions between atoms in the discharge removing coherence between separate emissions. Thus only in the periods between collisions when each atom is travelling freely will there be coherence. However, most vapour lamps operate at a pressure of a few millibars and under these conditions the effect of collision broadening is less than that of Doppler broadening. Overall, it can be seen that vapour lamps do not possess sufficient temporal coherence required for long path length interferometry.

Fortunately light from a laser is much more coherent, particularly light from a stabilised laser. This is due to the natural coherence exhibited by stimulated emission where the phase of the light emitted by an atom is the same as the wave stimulating it to emit. The limiting factor which determines the linewidth arises from the instabilities of the lasing cavity mirrors together with the small amount of spontaneous emission present in the discharge. A typical linewidth for a He-Ne laser, (see § 3.2.1), is 200 MHz (4×10^{-7}) decreasing to about 50 kHz for a stabilised laser (1×10^{-10}). This is equivalent to a coherence time of 6×10^{-6} seconds which is sufficient for the interferometer.

4.2.2 Spatial coherence

Spatial coherence is a function of the source size: if light from different areas of the source arrives at the image plane with different phases, the visibility of the fringes will decrease due to extra destructive interference.

4.2.2.1 An approximate estimate of the spatial coherence

Consider an incoherent source on axis, illuminating a slit of width x . Behind the slit is a screen. Two points on the screen, A and B are separated by a distance Δ .

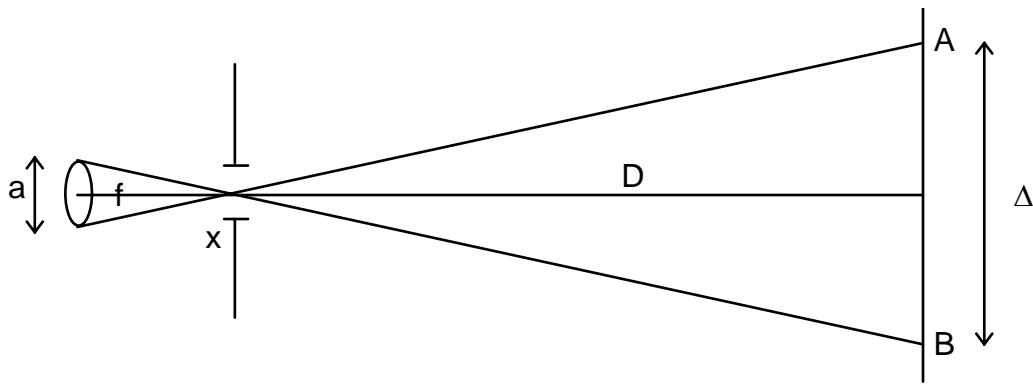


Figure 4.15 - Coherence of an extended source with slit and screen

The extended source produces diffraction in the region between A and B . For an incoherent source, the interference patterns at A and B are not related as their respective source points are un-correlated and so the two interference patterns will on average cancel each other. If the fringes at A and B have period $\sim D\lambda/x$ and if the two sets are in anti-phase, then

$$\Delta = \frac{D\lambda}{2x} \quad (4.19)$$

From similar triangles $\Delta = \frac{aD}{f}$

therefore for fringe cancellation $x = \frac{\lambda f}{2a}$

or $x = \frac{\lambda}{2\theta}$ (4.20)

where θ is the angular size of the source.

Thus since the coherence distance in the plane of the pupil is $\lambda f/2a$ the coherence distance is increased by minimising the angular size of the source, as expected. If the source is of fixed size then the remaining option is to increase the source-pupil distance. The source diameter, a , is 4 μm thus the spatial coherence distance across the wavefront should be approximately 12 cm.

4.2.2.2 Detailed estimate of the spatial coherence

Just as the temporal coherence function $\gamma(\tau)$ is related via a Fourier transform to the spectral intensity $I(v)$, the van Cittert [13]-Zernike [14] theorem states that the spatial

coherence function $\gamma(r)$ is related to the Fourier transform of the intensity distribution of the source, $I(\theta, \phi)$. This will now be derived (after Lipson & Lipson [12]).

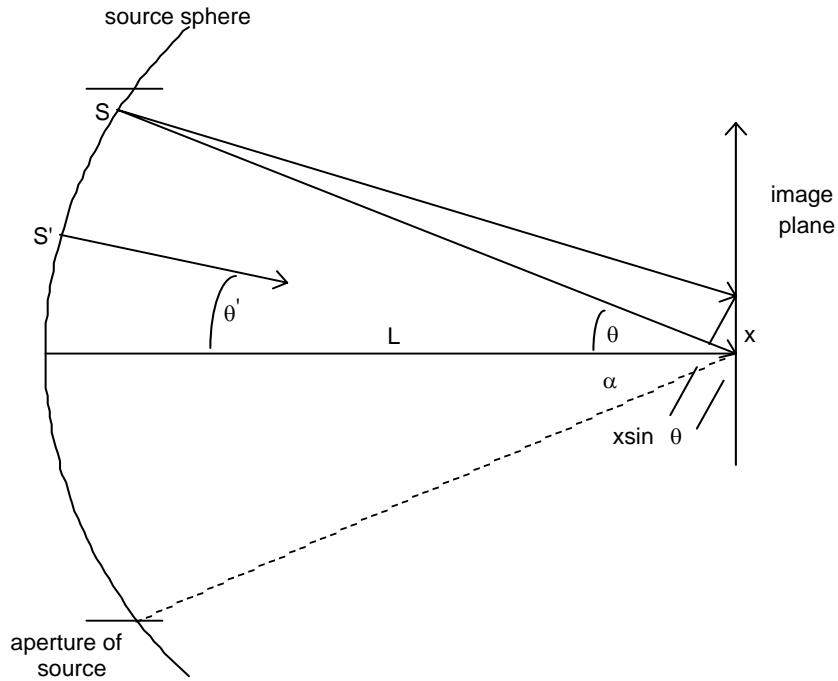


Figure 4.16 - Source sphere centred on image plane origin

Consider the source illustrated in figure 4.16. For an incoherent source the amplitudes at different points on the sphere, $g(\theta)$ and $g(\theta')$ are unrelated.

The amplitude in the image plane at $x = 0$ is $A(0)$

$$A(0) = \frac{1}{L} \int g(\theta) e^{-ikSP} d\theta \quad (4.21)$$

and at $x = x$ is $A(x)$

$$A(x) = \frac{1}{L} \int g(\theta) e^{-ik(SP - x \sin \theta)} d\theta \quad (4.22)$$

Defining $c(x) \equiv A(0)A^*(x)$

$$\begin{aligned} c(x) &= \frac{1}{L^2} \int g(\theta) e^{-ikSP} d\theta \int g^*(\theta) e^{-ik(SP - x \sin \theta)} d\theta \\ &= \frac{1}{L^2} \int g(\theta) g^*(\theta) e^{-ikx \sin \theta} d\theta \end{aligned} \quad (4.23)$$

Now $\gamma(x) \equiv \frac{\overline{c(x)}}{\overline{c(0)}}$ and $g(\theta)g^*(\theta) = I(\theta)$ (4.24)

$$\overline{c(x)} = \frac{1}{L} \int \overline{I(\theta)} e^{-ikx \sin \theta} d\theta$$

For small angles approximate $\sin \theta$ by θ

$$\therefore \gamma(x) = \frac{\int I(\theta) e^{-ikx\theta} d\theta}{\int I(\theta) d\theta} \quad (4.25)$$

i.e. $\gamma(x)$ is the normalised Fourier transform of $\overline{I(\theta)}$.

As an example, if $\overline{I(\theta)}$ is uniform and unity within an incoherent circular source of angular radius α , then $\gamma(x)$ is the normalised Fourier transform of this function which is a Bessel function of the first kind, i.e.

$$\gamma(x) = \frac{2J_1(k\alpha x)}{k\alpha x} \quad (4.26)$$

(For an alternative treatment see for example Mandel and Wolf [15]).

$\gamma(x)$ falls to zero at $x = 0.61\lambda/\alpha$.

For the interferometer, $\alpha = 1.3 \times 10^{-6}$, thus the first zero of $\gamma(x)$ should be at $x = 0.30$ m, i.e. the wavefront exhibits at least some spatial coherence up to a diameter of 30 cm. This confirms the order of magnitude estimate of 12 cm, to within a factor of 2.5.

The above derivations have assumed an incoherent source. A variant on the van Cittert-Zernike theorem will now be given which does not make this assumption and so will be valid for the coherent source used in the interferometer.

Consider a plane, (ξ, η) containing a source Σ (see figure 4.17). The intensity of an element $d\sigma$ of the source at a point A with co-ordinates (ξ, η) is given by $I(\xi, \eta)$. A second plane (x, y) is separated from the first by a distance R , and contains two points P_1 and P_2 .

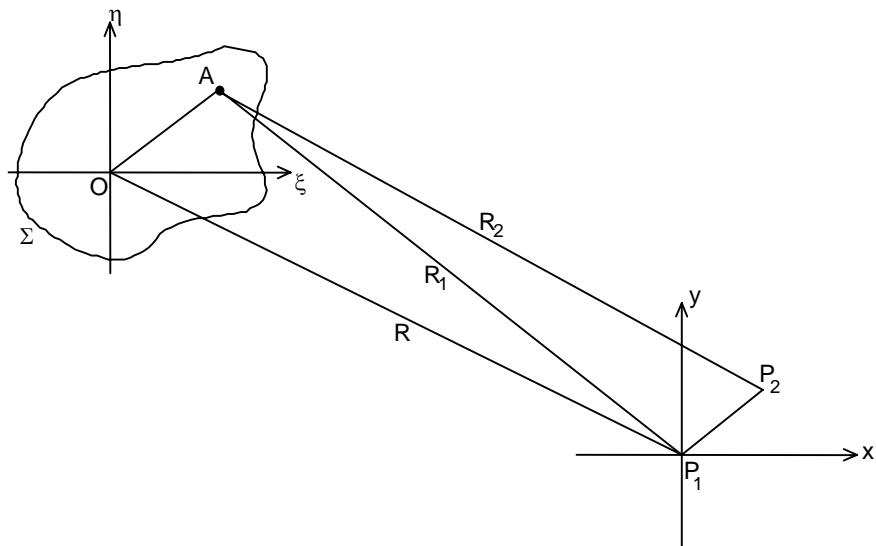


Figure 4.17 - Coherence between source and image planes separated by R

Assume that the radiation from Σ is uniform over all angles. The complex amplitudes at P_1 and P_2 produced by an element $d\sigma$ of Σ are u_1 and u_2 respectively.

$$u_1 = \frac{\sqrt{I(\xi, \eta)}}{R_1} e^{-ikR_1} \quad u_2 = \frac{\sqrt{I(\xi, \eta)}}{R_2} e^{-ikR_2} \quad k = \frac{2\pi}{\lambda}$$

γ_{12} defined before as $\gamma(x)$ is the complex degree of coherence between P_1 and P_2 and is given by

$$\gamma_{12} = \frac{1}{\sqrt{I_1 I_2}} \int_{\Sigma} \frac{I(\xi, \eta)}{R_1 R_2} e^{ik(R_2 - R_1)d\sigma} \quad (4.27)$$

i.e. the coherence factor between P_2 and P_1 is the same as the complex amplitude at P_2 in the diffraction pattern associated with the aperture Σ with the pattern centred on P_1 .

If P_1 is at the origin in the (x, y) plane and P_2 is at (x, y) and A is at (ξ, η) then

$$R_2 - R_1 = \frac{\sqrt{x^2 + y^2}}{R_1 + R_2} - \frac{2}{R_1 + R_2} (x\xi + y\eta) \quad (4.28)$$

For small values of x, y and with a sufficiently large source

$$R_2 - R_1 \approx -\frac{1}{R} (x\xi + y\eta) \quad (4.29)$$

$$\therefore \gamma_{12} = \frac{1}{R_2 \sqrt{I_1 I_2}} \int_{\Sigma} I(\xi, \eta) e^{\frac{-ik(x\xi + y\eta)}{R}} d\xi d\eta \quad (4.30)$$

i.e. γ_{12} is the Fourier transform of $I(\xi, \eta)$. This is the same result as the van Cittert-Zernike theorem, except that it holds for incoherent sources. This result is for large apertures only. It will now be developed for small apertures. This will be examined for the case of a circular source, angular radius α , centred on the origin, after Hopkins [8].

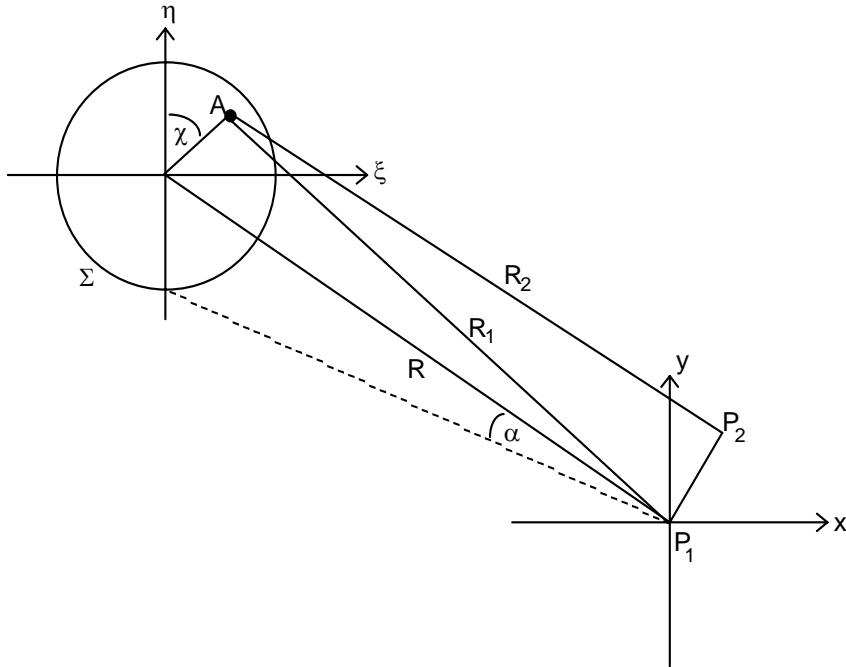


Figure 4.18 - Coherence from a uniform circular source centred at origin

Using circular co-ordinates as in figure 4.18

$$r \equiv \frac{\sin \theta}{\sin \alpha} \quad 0 \leq r \leq 1$$

From the above definition of γ_{12} , if Σ is small, then $R_2 \sim R_1$

$$\frac{1}{R_1 R_2} \approx \frac{1}{R_1^2} \quad (4.31)$$

$$\begin{aligned} \frac{d\sigma}{R_1 R_2} &= \frac{R_1 d\chi R_1 \tan \theta d\theta}{R_1 R_2} \\ &= \tan \theta d\theta d\chi \end{aligned} \quad (4.32)$$

$$\therefore \frac{d\sigma}{R_1 R_2} = \frac{\sin^2 \alpha}{\cos^2 \theta} r dr d\chi \quad (4.33)$$

Now require exponential term of γ_{12} in polar co-ordinates

Define $\rho = \sqrt{x^2 + y^2}$. It can be shown that

$$R_2 - R_1 = \frac{\rho^2}{R_1 + R_2} - \frac{2}{R_1 + R_2} (x\xi + y\eta) \quad (4.34)$$

Thus the exponential term becomes

$$ik \left\{ \frac{\rho^2 \cos \theta}{2R} - \rho \sin \theta \cos(\chi - \phi) \right\} \quad (4.35)$$

Now with $\cos \theta = 1 - \frac{1}{2} \sin^2 \theta$ this becomes

$$\begin{aligned} & ik \frac{\rho^2}{2R} \left[1 - \frac{1}{2} \sin^2 \theta \right] - ik \rho \sin \theta \cos(\chi - \phi) \\ & \frac{ik\rho^2}{2R} - \frac{ik\rho^2}{4R} \sin^2 \theta - ik \rho \sin \theta \cos(\chi - \phi) \end{aligned}$$

Now, correcting a misprint in Hopkins' paper, setting $z = k\rho \sin \alpha$ gives

$$\frac{ik\rho^2}{2R} - \frac{i\lambda}{8\pi R} (zr)^2 - izr \cos(\chi - \phi) \quad (4.36)$$

In practice, $z < 10$ and $r < 1$, also $\lambda/R \ll 1$ and therefore the term in $\frac{\lambda}{8\pi R} (zr)^2$ can be neglected. Thus the exponential term becomes

$$e^{\frac{ik\rho^2}{2R} - izr \cos(\chi - \phi)} \quad (4.37)$$

Substituting this and (4.33) in the expression for γ_{12} gives

$$\gamma_{12} = \frac{1}{\sqrt{I_1 I_2}} \int_0^{2\pi} \int_0^1 \frac{I(\xi, \eta) \sin^2 \alpha}{\cos^2 \theta} e^{\frac{ik\rho^2}{2R} - izr \cos(\chi - \phi)} r dr d\chi \quad (4.38)$$

removing terms independent of r and χ gives

$$\gamma_{12} = \frac{\sin^2 \alpha}{\sqrt{I_1 I_2}} e^{\frac{i k \rho^2}{2R}} \int_0^1 \int_0^{2\pi} \frac{I(\xi, \eta)}{\cos^2 \theta} e^{-izr \cos(\chi - \phi)} r dr d\chi \quad (4.39)$$

If the source is of uniform brightness, $I(\xi, \eta)$ is constant and equal to I , but will decrease by a factor of $\cos^2 \theta$ according to the angle θ between the normal to the plane of the source and the propagation direction

$$\begin{aligned} \gamma_{12} &= \frac{I \sin^2 \alpha}{\sqrt{I_1 I_2}} e^{\frac{i k \rho^2}{2R}} \int_0^1 e^{-izr \cos(\chi - \phi)} r dr d\chi \\ &= \frac{I \pi \sin^2 \alpha}{\sqrt{I_1 I_2}} e^{\frac{i k \rho^2}{2R}} \frac{2 J_1(z)}{z} \end{aligned} \quad (4.40)$$

Approximating

$$I_1 \sim I_2 = I \pi \sin^2 \alpha \quad \text{gives}$$

$$\gamma_{12} = \frac{2 J_1(z)}{z} e^{\frac{i k \rho^2}{2R}} \quad (4.41)$$

This is the same expression as the van Cittert-Zernike theorem except for the factor

$$\frac{i k \rho^2}{e^{2R}}$$

which represents the phase difference of P_2 relative to P_1 due to different optical path lengths from the source to the two points. This is obvious from the limiting case where the source size vanishes

$$\begin{aligned} \alpha \rightarrow 0 \quad , \quad z \rightarrow 0 \quad , \quad \frac{2 J_1(z)}{z} \rightarrow 1 \\ \gamma_{12} \rightarrow e^{\frac{i k \rho^2}{2R}} \end{aligned} \quad (4.42)$$

The magnitude of γ_{12} becomes unity and represents a simple phase difference between the two points. Hopkins stated that the modulus of this phase factor was approximately unity except for small α whereas in fact it is always unity and does not depend on α .

After passage through the collimating lens, all such points in the beam should have the same phase, though in the case where α is non-zero, the coherence will vary as the separation of the points, as dictated by γ_{12} .

Figure 4.19 shows the variation of γ_{12} over the two dimensional plane (x, y) for the case of the interferometer. Figure 4.20 is a section through this function at $y = 0$ showing the detail. It is common to take a value of γ_{12} of 0.88 as being the cut-off point for coherence (similar to Strehl criterion). From figure 4.21, this occurs at $x = 0.075$ m, for $y = 0$, i.e. the diameter of the coherent disc at the entrance pupil is 7.5 cm. Any two points in the double ended system which are sheared by less than 7.5 cm when imaged onto each other will be coherent and produce fringes of suitable contrast.

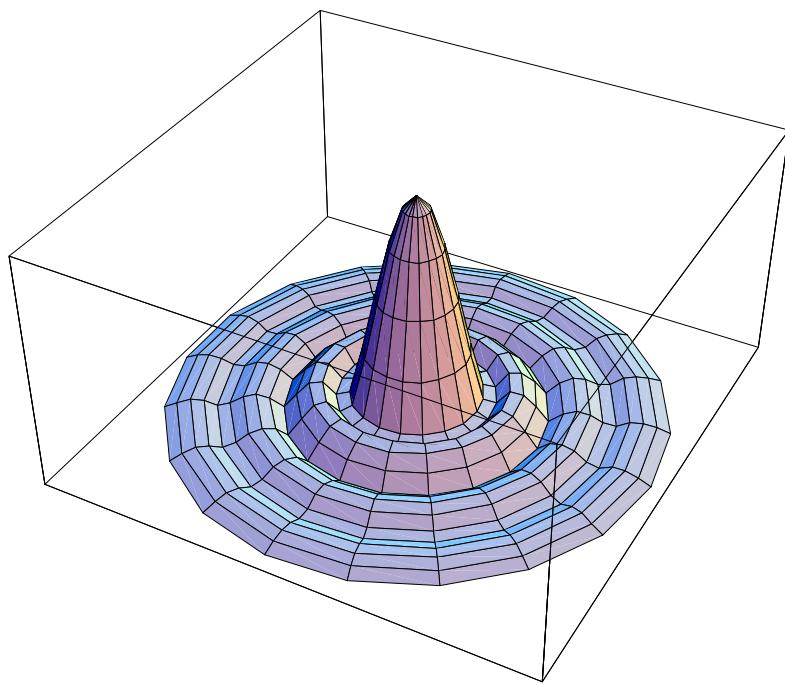


Figure 4.19 - Variation of coherence over area of image for the Primary Length Bar Interferometer

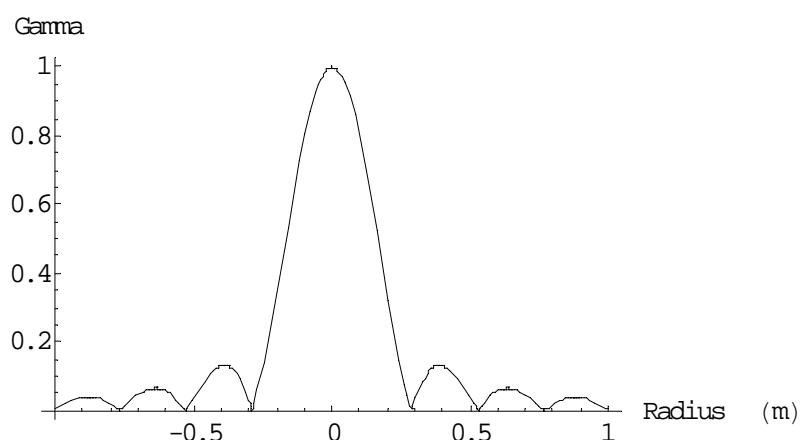


Figure 4.20 - Section through figure 4.19 showing detail

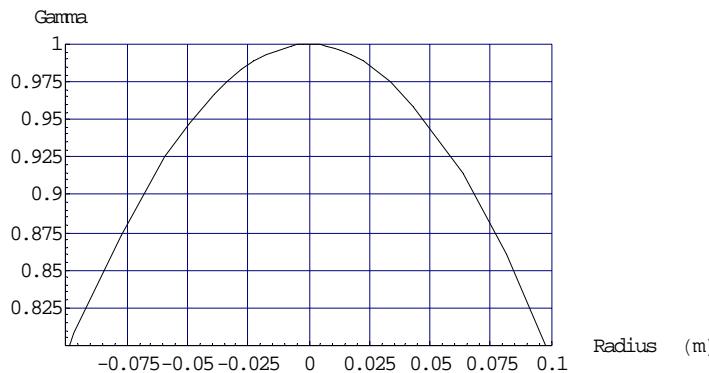


Figure 4.21 - Region of figure 4.19 about coherence limit of 0.88

The coherence between points in the beam after the collimating lens may be found using the propagation formulae of Zernike [14] or Hopkins [8]. However if it is assumed that there are no wavefront aberrations due to passage through the collimating lens then, according to Zernike:

"The degree of coherence in a plane illuminated though a lens is the same whether a source of uniform brightness be imaged on the plane or placed directly behind the lens - the phase-changing properties have no influence on the coherence."

And from Hopkins:

$$\Gamma'_{21} = e^{ik[W(x_2, y_2) - W(x_1, y_1)]} \Gamma_{21} \quad (4.43)$$

where (x_1, y_1) and (x_2, y_2) are points in the exit pupil, $W(x, y)$ is the wavefront aberration at point (x, y) in the pupil, Γ_{21} is the coherence in the entrance pupil and Γ'_{21} is the coherence in the exit pupil at the corresponding point.

These calculations of the coherence expected in the interferometer are supported by the fact that fringes are observed when the interferometer is operating in double-ended mode, although the fringes corresponding to the sheared beams do have lower contrast compared to the un-sheared fringes. The maximum shearing distance is approximately 8 cm, for which the above calculation predicts a coherence of approximately 0.87, just outside the conventional limit of 0.88. The reduction in fringe contrast can be seen in figure 4.22. The right image of the bar is of the front face, where the image is formed in the same way as for single-ended interferometry. The left image is that of the rear face of the bar, which requires a longer path difference and image shearing. The worst contrast fringes are those in the background which are formed by light which has

travelled twice along the length of the bar and also been sheared. Extra tilt has been added to the fringes for easier viewing.

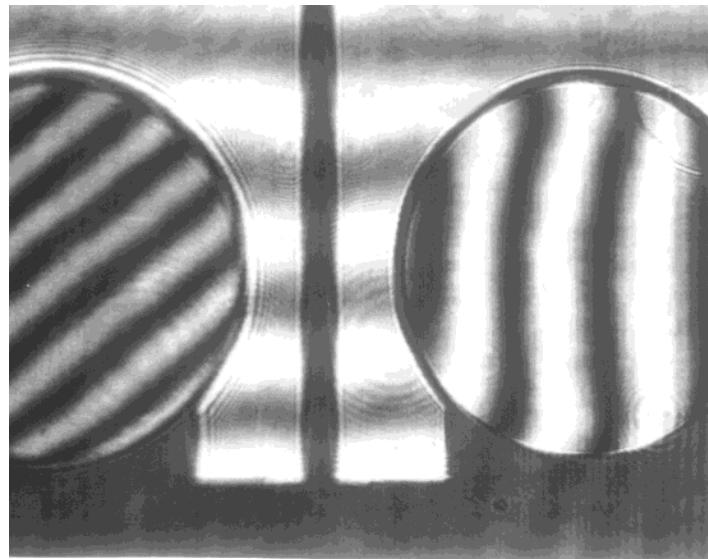


Figure 4.22 - Double-ended interferogram showing different fringe contrasts

Note that the effect of tilt due to the wedged beamsplitter is still present in the double-ended interferograms (see figure 4.23). The change in tilt with wavelength is the same as before on the front and rear faces of the bar. The change in tilt of the background fringes is much greater, *e.g.* with the optics adjusted for zero background fringes for the red wavelength, at the orange wavelength there are approximately 8 fringes of tilt across one roof mirror and at the green wavelength there are approximately 11 fringes of tilt (see figure 4.23).

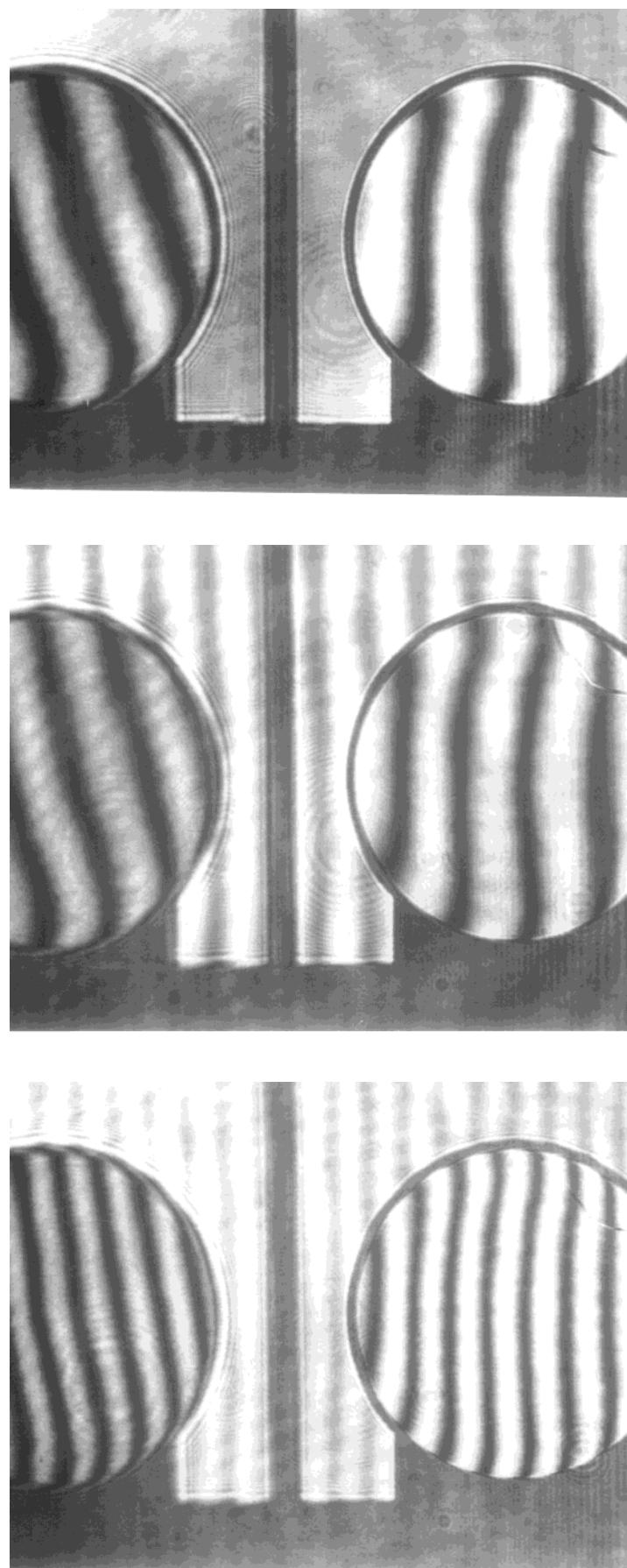


Figure 4.23 - Double-ended images for three wavelengths, same alignment of optics in all three images

REFERENCES FOR CHAPTER 4

- [1] Hecht E *Optics* 2nd Edition (Reading, MA: Addison-Wesley) (1987) 416-421
 - [2] Bruce C F The effects of collimation and oblique incidence in length interferometers *Aust. J. Phys.* **8** (1955) 224-240
 - [3] Thornton B S The effects of collimation and oblique incidence in length interferometers *Aust. J. Phys.* **8** (1955) 241-247
 - [4] 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** (1964) 531-534
 - [5] Spindler & Hoyer *Precision Optics* (1989) I1-I2
 - [6] Liepmann H W Wedged plate beamsplitter without ghost reflections *Appl. Opt.* **31** (1992) 5905 - 5906
 - [7] Welford W T *Aberrations of Optical Systems* (Bristol: Adam Hilger) (1991) 199
 - [8] Hopkins H H The concept of partial coherence in optics *Proc. Roy. Soc. A* **208** (1951) 263-277
 - [9] Hopkins H H On the diffraction theory of optical images *Proc. Roy. Soc. A* **217** (1953) 408-432
 - [10] Hopkins H H Interferometric methods for the study of diffraction images *Opt. Acta* **2** (1955) 23-29
 - [11] Hopkins H H Applications of coherence theory in microscopy and interferometry *J. Opt. Soc. Am.* **47** (1957) 508-526
 - [12] Lipson & Lipson *Optical Physics* 2nd Edition (Cambridge: CUP) (1981)
 - [13] van Cittert P H Die Wahrscheinliche Schwingungsverteilung in einer von einer Lichtquelle direkt oder mittels einer Linse Beleuchteten *Physica* **1** (1934) 201-210
 - [14] Zernike F The concept of degree of coherence and its application to optical problems *Physica* **5** (1938) 785-795
 - [15] Mandel L & Wolf E Coherence properties of optical fields *Rev. Mod. Phys.* **37** (1965) 231-287
-

CHAPTER 5

FRINGE ANALYSIS & PHASE STEPPING INTERFEROMETRY

“Wo viel Licht ist, ist starker Schatten.”
("Where there is much light, the shadows are deepest")
Goethe

5.1 ANALYSIS OF INTERFERENCE FRINGES

5.1.1 Introduction to interference fringe analysis

Using interferometry it is possible to compare measured and reference wavefronts to a high degree of accuracy. Interferometric measurement techniques such as holographic interferometry, speckle interferometry, moiré *etc.* have found many applications from the measurement of engine blocks [1], hip joint prosthesis design [2] to high accuracy measurement of optical components such as mirrors, flats and lenses.

To achieve these accuracies it is necessary to use computer evaluation of the interference fringes. Image processing can be used to enhance the fringe patterns and remove noise before the phase is evaluated. Post-processing of the data often includes finite element analysis or boundary element analysis techniques to solve specific application problems.

Each technique requires determination of the interference phase at a number of points in the field to generate a phase distribution or phase map. There are many techniques for phase extraction, each suited to a particular experimental design. These techniques are: fringe skeletonisation, phase-shifting, phase-stepping, Fourier transform, temporal heterodyning, spatial heterodyning (carrier frequency) and phase locking. There are excellent review articles which deal with the variety of techniques used [3,4,5,6,7,8].

At its most simple level, fringe analysis can be performed by eye. In the Twyman-Green interferometer shown in figure 5.1, interference fringes are produced on a surface or optical component to be tested. The fringes are detected by a CCD array and displayed on a monitor. The equation for the intensity of the fringes in the plane of the detector array is

$$I(x, y, t) = a(x, y) + b(x, y) \cos[\phi(x, y) + \Phi_R(x, y, t)] \quad (5.1)$$

Here, $a(x, y)$ represents the variation of the background illumination, $b(x, y)$ describes the noise and contrast variations, $\phi(x, y)$ is related to the surface to be measured and $\Phi_R(x, y, t)$ is the reference phase, or wavefront at time t . The co-sinusoidal variation of I leads to a set of interference fringes, of co-sinusoidal intensity.

The fringes can be regarded as contour lines of surface height of the test object, spaced at intervals of $\lambda/2$ where λ is the wavelength of the light source. Using these fringes as contours, one can determine by eye, wavefront aberrations or surface defects to about $\lambda/5$ or $\lambda/10$, simply by observing the positions of the fringes in the interferogram. It is easy to detect defects such as spherical aberration, coma, or to spot inhomogeneities in refractive components, or flatness errors in mirrors.

For more accurate measurement, it is necessary to be able to sub-divide the fringes, *i.e.* to measure the phase at all points in the display. This requires some computer processing of the intensity distribution in the image.

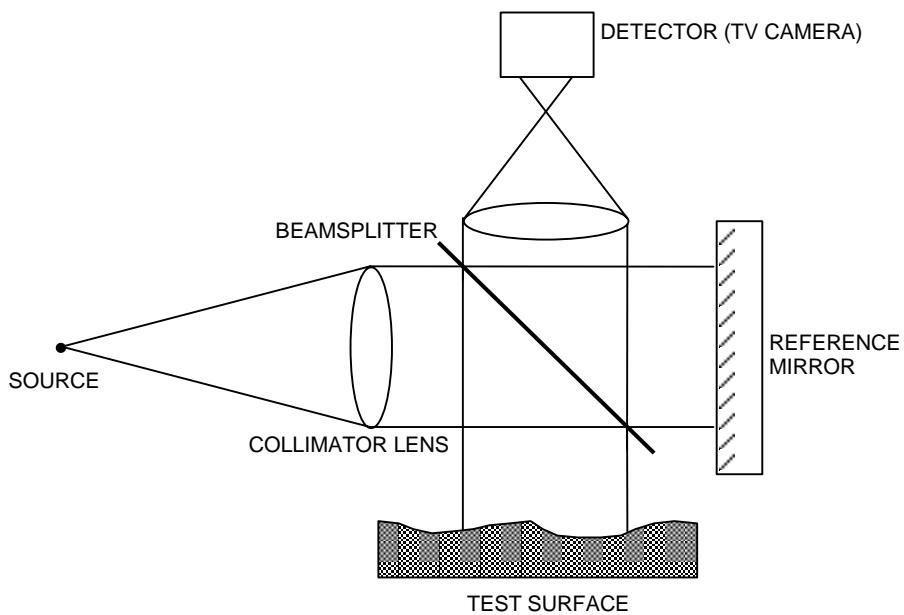


Figure 5.1 - Example Twyman-Green interferometer for optical testing

5.1.2 Fringe skeletonisation methods

Fringe skeletonisation is an extension of the fringe analysis performed by eye based on tracking fringe maxima or minima across the field. The computer algorithm searches for maxima and minima in the digitised interference pattern [9]. The phase at these points corresponds to multiples of π . Many algorithms exist for tracking along a fringe extremum, usually based on finding the normal to the maximum gradient of the intensity, or by following a path of minimum change of intensity. The result is a set of lines, one pixel wide, which correspond to the extrema, and are often overlaid on the original image for comparison. The analysis then requires the joining together of lines which are disconnected (such as near a defect) followed by numbering of the lines. This last step must usually be performed with user input, especially where lines are discontinuous [10]. The phase at points lying between fringe extrema is calculated by linear, polynomial or spline interpolation along a suitable direction in the phase map.

The main advantage of fringe skeletonisation is that it requires only one digitised interferogram and so temporal drifts of the experimental arrangement have little effect on the phase measurement. However, the accuracy is approximately $\lambda/10$, the computation time is long, there is no averaging between many frames to suppress noise and it is sometimes difficult to assign the correct sign to phase gradients, since the intensity change can appear the same for both positive and negative gradients.

5.1.3 Fourier transform methods

The Fourier transform technique requires only one interference pattern, for which the reference phase, $\Phi_R(x, y, t)$ can be arbitrarily set to zero. Expanding the cosine function in (5.1) using Euler's formula, and the definition

$$c(x, y) = \frac{1}{2} b(x, y) e^{i\phi(x, y)} \quad (5.2)$$

gives

$$I(x, y) = a(x, y) + c(x, y) + c^*(x, y) \quad (5.3)$$

Applying the two-dimensional Fourier transform to this gives

$$I(u, v) = A(u, v) + C(u, v) + C^*(u, v) \quad (5.4)$$

Since $I(x,y)$ is real in the spatial domain, it follows that $\mathbf{I}(u,v)$ is Hermitian in the spatial frequency domain, *i.e.* $\Re\{\mathbf{I}(u,v)\}$ is even and $\Im\{\mathbf{I}(u,v)\}$ is odd. The amplitude spectrum is symmetric about the zero-frequency position, and so $\mathbf{C}(u,v)$ and $\mathbf{C}^*(u,v)$ contain the same information. By bandpass filtering in the spatial frequency domain, $A(u,v)$ and $\mathbf{C}^*(u,v)$ can be removed to leave $\mathbf{C}(u,v)$, which when the inverse Fourier transform is applied gives $c(x,y)$ which is now complex. The phase can then be measured from

$$\phi(x,y) = \arctan \frac{\Im\{c(x,y)\}}{\Re\{c(x,y)\}} \quad (5.5)$$

In effect, the Fourier transform method is a least squares fit of a linear combination of harmonic functions to the interference pattern.

If only one interferogram is used in the evaluation of the phase, then there is an ambiguity in the sign of the phase, due to loss of information during the filtering stage. This can be resolved by using a second interferogram with the reference phase shifted by up to π .

5.1.4 Temporal heterodyning methods

In temporal heterodyning, the two interfering wavefronts are formed from sources which have different frequencies, approximately a few kHz apart [11]. A common technique for generating these frequencies is to split a laser output into two modes by magnetic (Zeeman) splitting. The interferogram oscillates at the frequency of the beat between the two waves. A photodetector is used to sample the signal at points in the interferogram (there are no CCD detector arrays with high enough bandwidths). The phase can be measured either as the difference in phase between two detector points or between a single point and a reference phase signal. Phase distributions can only be measured by scanning the detector in the image.

5.1.5 Spatial heterodyning methods

To perform spatial heterodyne interferometry with frequency domain processing, a system used for Fourier transform interferometry has an additional set of carrier fringes introduced by tilting a mirror. These heterodyne carrier fringes have spatial frequency f_0 . The carrier frequency will cause a phase gradient across the image of size $2\pi f_0 x$. This then takes the place of $\Phi_R(x,y,t)$ in equation (5.1). This carrier frequency is removed by shifting the filtered spectral components in the frequency domain. This

allows the two components $C(u,v)$ and $C^*(u,v)$ to be effectively separated, making the filtering operation easier to perform. In effect, the single interferogram can be regarded as a multi-channel interferogram, where the different channels are separated spatially, *i.e.* in different pixels in the image, rather than at the same pixels, but separated in time, as in temporal techniques.

Another version of spatial heterodyning uses pointwise multiplication of the digitised intensity data by $\cos(2\pi f_0 x)$ and $\sin(2\pi f_0 x)$ to analyse the data in the spatial domain. If the spatial frequency of the fringes is similar to the frequency of these additional quadrature terms, then low frequency difference components can be separated by a low pass filter. These components are in phase quadrature in terms of the phase to be measured.

The technique of spatial heterodyning requires superposition of fringes (either real or in software) of a similar spatial frequency to the original interferometric fringes. This may not be possible where the original fringes are not of equal inclination and spacing. The technique using extra tilting of the mirror requires that the phase and amplitude of the wave to be measured must not change appreciably within the period of the spatial-carrier-frequency [12], *i.e.* the surface to be measured must be flat or a large tilt must be given to produce many fringes across the surface.

5.1.6 Phase locking methods

In the phase locking technique, the phase of the reference beam is modulated sinusoidally by less than $\lambda/2$, at a frequency ω . A bandpass filter centred at ω is used to sample the intensity at each point in the interferogram. At points where $\phi(x,y) = N\pi$ the detected intensity averages to zero. Thus the phase lock technique is a dynamic method of fringe skeletonisation, in real time. The technique has the same disadvantages as that of fringe skeletonisation.

5.1.7 Summary of phase measurement methods

Table 5.1 shows a summary of the phase measurement methods examined so far as well as the technique of phase-stepping interferometry, which will be examined shortly. Whether one technique is better than another depends on the application.

	Fringe skeleton	Phase stepping	Fourier transform	Temporal heterodyne	Spatial heterodyne	Phase locking
No. of images	1	3,4,5	1(2)	1 per pixel	1	1
Resolution (λ)	1 - 0.1	0.1 - 0.001	0.1 - 0.03	0.01 - 0.001	0.1 - 0.03	1 - 0.1
Measurement at all points	no	yes	yes	yes	yes	no
Noise suppression	partial	yes	yes	partial	yes	partial
Sign detection	no	yes	no(yes)	yes	(yes)	no
Difficulty	low	high	low	very high	low	medium
Computation time	long	short	long	very long	very long	long
Real time	some	some	no	some	no	yes

Table 5.1 - Summary of phase measurement techniques

5.2 PHASE-STEPPING INTERFEROMETRY (PSI)

5.2.1 History of PSI

Phase-measuring, Phase-Shifting or Phase-Stepping Interferometry is a technique used in the analysis of interference patterns generated by multiple beam interferometry. PSI has existed in its basic form for less than 3 decades, and may be regarded as temporal multiplexing of the interferograms to be analysed (temporal-carrier) where the use of an extra time variable reduces the problem of phase extraction to reading the phase of a sinusoidal signal, with varying time co-ordinate, but fixed spatial co-ordinates [13,14,15,16,17,18,19].

There has been a resurgence of interest in PSI as a measurement technique since the mid 1980s. This has been due to recent advances in the equipment required in PSI for image detection and processing, together with a general reduction in cost of computer power. Also, there has been a trend for measurement instruments and systems to become more automated and objective in their analyses - PSI is ideal for this, as it is easily implemented on standard ranges of computers.

5.2.2 Basic theory of PSI

PSI, as its name suggests, involves the variation of phase within the interferometer, by a controlled amount. All designs of phase-stepping interferometers have a number of features in common:

- (i) The interferogram is imaged onto a detector *e.g.* CCD TV camera, photodiode array, holographic plate
- (ii) The interferogram is a comparison between the wavefronts generated by test and reference surfaces
- (iii) The relative phase of one of the interferometer arms (reference or test) is varied with respect to the other by a fixed and known amount, either continuously (phase-shifting) or in discrete steps (phase-stepping)
- (iv) The intensity of the interferogram is either summed continuously or stored at each step, depending on which method is used in (iii)
- (v) After the phase-stepping or shifting is complete, the analysis of the stored data is undertaken. A major advantage of phase-stepping interferometry is that a detector array such as a CCD camera can be used to make measurements simultaneously at a very large number of points covering the interference pattern, the resolution being limited by the optical magnification and the detector pixel size.

5.2.3 Derivation of generic PSI equations

The phase calculation of PSI is based on the fact that the intensity $I(x,y)$ at a point (x,y) in the interferogram is the result of interference of two wavefronts. Considering the Twyman-Green interferometer shown in figure 5.2, which is assumed to be made from optically perfect components, *i.e.* there are no aberrations.

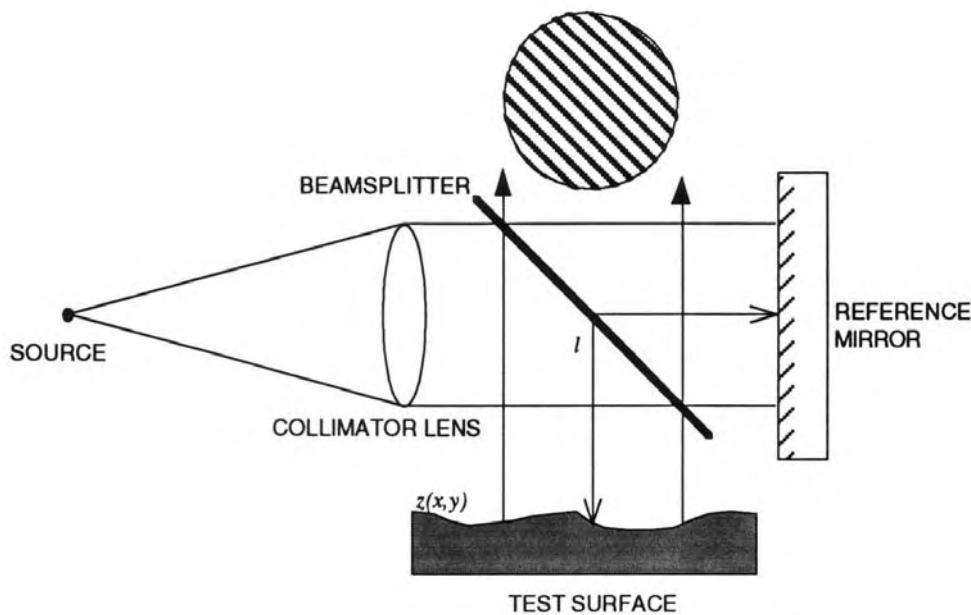


Figure 5.2 - Idealised interferometer for testing surfaces

In figure 5.2 and the following derivation, $z(x,y)$ is the surface profile of the object under test, λ is the wavelength of the monochromatic light, l is representative of the total (average) optical path difference between the two surfaces. Wavefronts from the reference and test arms are given respectively by

$$w_r = ae^{2ikl} \quad (5.5)$$

$$w_t = be^{2ikz(x,y)} \quad (5.6)$$

with $k = 2\pi/\lambda$ and a and b are the amplitudes of the interfering wavefronts, due to different reflectivities of the surfaces. In the interference pattern,

$$I(x,y,l) = (w_r + w_t)(w_r + w_t)^* \quad (5.7)$$

$$= a^2 + b^2 + 2ab \cos(2k(z(x,y) - l)) \quad (5.8)$$

The term $(a^2 + b^2)$ represents the background intensity, or DC level, and the variation of $2ab \cos(2k(z(x,y) - l))$ represents the interference fringes, observed as co-sinusoidal variations in intensity, $I(x,y,l)$. One can vary the intensity $I(x,y,l)$ either by keeping l constant, and changing x or y - i.e. moving across a non-flat surface, or by keeping x and y constant and varying l , the optical path difference. In PSI, l is varied by varying the phase of one of the beams, e.g. by moving the reference mirror longitudinally along the beam axis. As the value of l is varied, the intensity $I(x,y,l)$ at each point (x,y) in the interferogram varies in a co-sinusoidal manner (assuming linear detection, non-

aberrated optics *etc*). The premise of PSI is that by knowing the variation in l , and observing the variation in $I(x,y,l)$ at each point in the interferogram, one can correlate the two, and hence retrieve $z(x,y)$ in terms of k (or λ).

This can be seen by re-writing (5.8):

$$I(x, y, l) = a_0 + a_1 \cos 2kl + b_1 \sin 2kl \quad (5.9)$$

where a_1 and b_1 are functions of x and y , and hence contain information about $z(x,y)$. Now if the phase in one beam is stepped (by varying l) in n steps, each of which is $1/n$ of a fringe *i.e.* $\lambda/2n$ in size, then, using the orthogonality relationships for sin and cos

$$a_0 = \frac{1}{n} \sum_{i=1}^n I(x, y, l_i) = a^2 + b^2 \quad (5.10)$$

$$a_1 = \frac{2}{n} \sum_{i=1}^n I(x, y, l_i) \cos 2kl_i = 2ab \cos(2kz(x,y)) \quad (5.11)$$

$$b_1 = \frac{2}{n} \sum_{i=1}^n I(x, y, l_i) \sin 2kl_i = 2ab \sin(2kz(x,y)) \quad (5.12)$$

from which

$$\frac{b_1}{a_1} = \frac{\frac{2}{n} \sum_{i=1}^n I(x, y, l_i) \sin 2kl_i}{\frac{2}{n} \sum_{i=1}^n I(x, y, l_i) \cos 2kl_i} = \frac{2ab \sin(2kz(x,y))}{2ab \cos(2kz(x,y))} \quad (5.13)$$

$$= \tan(2kz(x,y)) \quad (5.14)$$

$$\therefore 2kz(x,y) = \arctan\left(\frac{b_1}{a_1}\right) \quad (5.15)$$

$$\therefore z(x,y) = \frac{1}{2k} \arctan\left(\frac{b_1}{a_1}\right) \quad (5.16)$$

i.e. $z(x,y)$ can be determined (to within modulo 2π) from a_1 and b_1 , the summed intensities at (x,y) . This is the basis of PSI. There are many variations of this basic equation, which are detailed in § 5.3.

5.2.4 Typical applications of PSI

PSI has been widely used [14,20,21] in the testing of optical components [22,23]. The range of surface variation measurable by basic PSI is limited to a few microns. This is due to the compromise between having lots of fringes across the interferogram either as the result of tilt or due to large test surface deviations, and the requirement that each fringe must be large enough in width to be imaged onto at least 2 detector pixels. If the fringe is smaller than 2 pixels, each pixel will integrate the intensity of the whole fringe, and no fringe modulation will be observed during phase stepping. Another requirement of these techniques is that the surface under test is smooth and has no discontinuities present of magnitude greater than the measurement wavelength, as these discontinuities cannot be distinguished from the 2π discontinuities present in the wrapped phase data.

5.2.5 Phase variation methods for PSI

In general, any technique which varies the phase in one or more of the interferometer beams can be used in PSI. The most common techniques include: moving diffraction gratings [24], moving the reference mirror by use of a PZT, the Bragg effect in an acousto-optic modulator [25] and rotating a half-wave plate in a polarised interferometer [24,26].

5.3 PHASE STEPPING TECHNIQUES

5.3.1 Basic phase-stepping techniques

Let the system of fringes in an interferometer have visibility V which is defined as

$$V = \frac{I_{\max} - I_{\min}}{I_{\max} + I_{\min}} \quad (5.17)$$

The mean intensity of the two beams is I_0 . When the phase of the reference beam is Φ , the intensity at a point in the interferogram will be given by

$$I_r = I_0 + I_0 V \cos(\Phi - \phi) \quad (5.18)$$

Taking the expression

$$I(x, y) = I_0(1 + V(x, y) \cos(\Phi - \phi(x, y))) \quad (5.19)$$

and using

$$\cos(A - B) = \cos A \cos B + \sin A \sin B$$

gives

$$I(x, y) = I_0 + I_0 V \cos \phi(x, y) \cos \Phi + I_0 V \sin \phi(x, y) \sin \Phi \quad (5.20)$$

Now using phase stepping, *i.e.* picking discrete values Φ_r of Φ given by

$$\Phi_r = \frac{(r-1)2\pi}{R} \quad \text{with } r = 1, 2, \dots, R \quad (5.21)$$

the intensity at a particular point (x, y) in the interferogram will be given by

$$I_r = I_0 + I_0 V \cos(\Phi_r - \phi) \quad (5.22)$$

expanding this gives

$$I_r = I_0 + I_0 V \cos \Phi_r \cos \phi + I_0 V \sin \Phi_r \sin \phi \quad (5.23)$$

Multiplying (5.23) by $\cos \Phi_r$ and $\sin \Phi_r$ separately gives

$$I_r \cos \Phi_r = I_0 \cos \Phi_r + I_0 V \cos \phi \cos^2 \Phi_r + I_0 V \sin \phi \sin \Phi_r \cos \Phi_r \quad (5.24)$$

$$I_r \sin \Phi_r = I_0 \sin \Phi_r + I_0 V \cos \phi \cos \Phi_r \sin \Phi_r + I_0 V \sin \phi \sin^2 \Phi_r \quad (5.25)$$

Now summing equations (5.23) to (5.25) over r

$$\sum_{r=1}^R I_r = \sum_{r=1}^R I_0 + \sum_{r=1}^R I_0 V \cos \Phi_r \cos \phi + \sum_{r=1}^R I_0 V \sin \Phi_r \sin \phi \quad (5.26)$$

$$\sum_{r=1}^R I_r \cos \Phi_r = \sum_{r=1}^R I_0 \cos \Phi_r + \sum_{r=1}^R I_0 V \cos \phi \cos^2 \Phi_r + \sum_{r=1}^R I_0 V \sin \phi \sin \Phi_r \cos \Phi_r \quad (5.27)$$

$$\sum_{r=1}^R I_r \sin \Phi_r = \sum_{r=1}^R I_0 \sin \Phi_r + \sum_{r=1}^R I_0 V \cos \phi \cos \Phi_r \sin \Phi_r + \sum_{r=1}^R I_0 V \sin \phi \sin^2 \Phi_r \quad (5.28)$$

Now using the orthogonality relationships for sin and cos:

$$\sum_x^{2\pi} \sin(mx) \sin(nx) = \begin{cases} 0 & \forall m \neq n \\ \pi & \forall m = n \neq 0 \end{cases}$$

$$\sum_x^{2\pi} \cos(mx) \cos(nx) = \begin{cases} 0 & \forall m \neq n \\ \pi & \forall m = n \neq 0 \end{cases}$$

$$\sum_x^{2\pi} \cos(mx) \sin(nx) = 0 \quad \forall m, n$$

then equations (5.26) to (5.28) reduce to

$$\sum_{r=1}^R I_r = RI_0 \quad (5.29)$$

$$\sum_{r=1}^R I_r \cos \Phi_r = \frac{1}{2} RI_0 V \cos \phi \quad (5.30)$$

$$\sum_{r=1}^R I_r \sin \Phi_r = \frac{1}{2} RI_0 V \sin \phi \quad (5.31)$$

from which it follows that

$$\frac{2 \sum_{r=1}^R I_r \sin \Phi_r}{2 \sum_{r=1}^R I_r \cos \Phi_r} = \frac{RI_0 \sin \phi}{RI_0 \cos \phi} \quad (5.32)$$

and hence

$$\tan \phi = \frac{\sum_{r=1}^R I_r \sin \Phi_r}{\sum_{r=1}^R I_r \cos \Phi_r} \quad (5.33)$$

This is the basic equation for all multi-step phase-stepping techniques.

5.3.2 Phase-shifting interferometry

There is another form of phase-measuring interferometry, phase-shifting interferometry. Here the phase is continuously varied and the detector integrates the intensity at each point over a range of phases. Grievenkamp [27] shows the integrated intensity to be

$$I_i(x, y) = \frac{1}{\Delta} \int_{x_i - \Delta/2}^{x_i + \Delta/2} I_0(x, y) \{1 + \gamma_0 \cos[\phi(x, y) + \alpha(t)]\} d\alpha(t) \quad (5.34)$$

$I_0(x, y)$ is the average intensity at detector point (x, y) , γ_0 is the modulation of the fringe pattern (corresponds to V used above), α_i is the average value of the relative phase shift for the i^{th} exposure, $\phi(x, y)$ is the test wavefront phase to be determined, and Δ is the phase shift over which the intensities are summed.

Thus

$$I_i(x, y) = I_0(x, y) \{1 + \gamma_0 \operatorname{sinc}(\Delta / 2) \cos[\phi(x, y) + \alpha_i]\}$$

Substituting $\Delta = 0$ (integrating over zero phase range), the above equation reduces to the phase-stepping case. The phase-shifting technique is often referred to as the ‘Integrating Bucket’ approach.

5.3.3 Four quadrant arctangent routine

The basic phase-stepping or phase-shifting equations have an initial limitation. Simply applying an equation of the form

$$\phi = \arctan\left(\frac{a}{b}\right)$$

returns values of ϕ in the range $-\pi/2$ to $\pi/2$ i.e. a range of π . This is unsatisfactory as each interference fringe corresponds to a range of phase values over the range 0 to 2π . This is easily resolved by noting that a corresponds to a sinusoid and b to a co-sinusoid, and thus the signs (+ or -) of these quantities can be used to uniquely define a quadrant for each calculation of ϕ , based on the four possible combinations.

<i>a</i>	<i>b</i>	ϕ
+	+	ϕ
+	-	$\pi - \phi$
-	+	$2\pi - \phi$
-	-	$\pi + \phi$

Table 5.2 - Four-quadrant lookup table

Suitable adjustments are made when either *a* or *b* or both are zero. Thus by use of a PSI technique based around equation (5.33), followed by application of a 4-quadrant arctangent, the relative phase at each point in the interferogram can be determined modulo 2π .

5.3.4 Two position phase-stepping technique

A two position technique has been used by Santoyo *et al* [28] in the analysis of Electronic Speckle Pattern Interference, where the fringes are defined by a different equation to that of conventional interferometry. It is not suitable for general PSI, as with only two measurements, I_1 and I_2 , it is not possible to solve for all three variables of the general PSI equation. However the technique is suited to the analysis of speckle pattern interferograms, as these are of the form

$$I(x, y) \propto \sin(\theta + \Delta\phi / 2) \quad (5.35)$$

where $\Delta\phi$ is the phase change due to surface deformation, and θ is the relative phase between the 2 beams. Hence with 2 values of θ , separated by $\pi/2$,

$$\frac{I_2}{I_1} \propto \frac{\sin(\theta + \Delta\phi / 2) \sin(\Delta\phi / 2)}{(\theta + \Delta\phi / 2 + \pi / 2) \sin(\Delta\phi / 2)} \quad (5.36)$$

$$\frac{\sin(\theta + \Delta\phi / 2)}{\cos(\theta + \Delta\phi / 2)}$$

$$\tan(\theta + \Delta\phi / 2)$$

Hence knowing θ to be constant, one can determine $\Delta\phi$ for a deformation.

The initial frames are first processed to improve contrast, and then the phase calculation is performed using

$$\phi(x, y) = \arctan\left(\frac{I_2(x, y)}{I_1(x, y)}\right) \quad (5.37)$$

Processing is carried out using a four-quadrant arctan lookup table using sign information about I_1 and I_2 to resolve quadrant ambiguities. However the method requires that the phase step θ be exactly $\pi/2$, otherwise the calculated value of ϕ will be incorrect.

5.3.5 Three position phase-stepping technique

As mentioned above, a minimum of three sets of recorded intensity data are required to solve the PSI equation (5.8). A common 3 position technique uses phase steps of $\pi/2$, using relative phases of $\pi/4$, $3\pi/4$, $5\pi/4$. Under these conditions the following analysis can be applied for all points (x, y) .

$$\begin{aligned} I_1 &= I_0 + I_0 \gamma \cos(\phi + \pi/4) \\ I_2 &= I_0 + I_0 \gamma \cos(\phi + 3\pi/4) \\ I_3 &= I_0 + I_0 \gamma \cos(\phi + 5\pi/4) \end{aligned} \quad (5.38)$$

from which it can be shown that

$$\phi = \arctan\left(\frac{I_3 - I_2}{I_1 - I_2}\right) \quad (5.39)$$

Other variations in the three position technique use a phase shift of $2\pi/3$ between each image, for which the phase is calculated from

$$\phi = \arctan\left(\sqrt{3} \frac{I_3 - I_2}{2I_1 - I_2 - I_3}\right) \quad (5.40)$$

However this technique takes longer to perform the phase calculation as there are more terms.

The three-position technique is subject to the same basic error sources as the four-position technique, and analysis of the errors will be dealt with simultaneously in § 5.3.7 for comparison.

5.3.6 Four position phase-stepping technique

Although only 3 images are required to solve for the three variables of the PSI equation, in practice more than 3 images are often digitised for ease of computation, noise suppression and reduction in sensitivity to phase stepper errors. In the four position technique, the nominal phase-step is $\pi/2$, and the reference phase takes values of 0, $\pi/2$, π , and $3\pi/2$. Using these values, the intensities at each point in images 1 to 4 are

$$\begin{aligned} I_1 &= I_0 + I_0 \gamma \cos(\phi) \\ I_2 &= I_0 + I_0 \gamma \cos(\phi + \pi/2) = I_0 - I_0 \gamma \sin(\phi) \\ I_3 &= I_0 + I_0 \gamma \cos(\phi + \pi) = I_0 - I_0 \gamma \cos(\phi) \\ I_4 &= I_0 + I_0 \gamma \cos(\phi + 3\pi/2) = I_0 + I_0 \gamma \sin(\phi) \end{aligned} \quad (5.41)$$

from which

$$\phi = \arctan\left(\frac{I_4 - I_2}{I_1 - I_3}\right)$$

As expected, due to the averaging over more images in the four position technique, it has a lower error than the three position technique, although it requires more storage and takes longer to process the images to extract the phase.

5.3.7 Errors for three and four position techniques

5.3.7.1 Error due to phase stepper error

The principal error which affects most PSI technique is that of phase-stepper error. The techniques of PSI assume a fixed and known phase step size, which for the 3 and 4 position techniques is $\pi/2$. However non-linearities in the movement of a PZT performing the phase-stepping, or a mis-calibration of phase step size can cause the calculated phase to be in error. A general equation for the error in the phase map due to the phase step error can be derived as follows.

The following analysis is assumed to apply to every point (x,y) in the interferogram.

Assuming an error ε_r in the size of the phase step, *i.e.*

$$\Phi'_r = \Phi_r + \varepsilon_r \quad (5.42)$$

where Φ'_r is the achieved phase step, and Φ_r is the correct phase step. Using

$$I_r = I_0 + I_0 \gamma \cos(\phi - \Phi_r) \quad (5.43)$$

as an equation for the intensity at a point for phase step angle Φ_r ,

$$I'_r = I_0 + I_0 \gamma \cos(\phi - \Phi_r + \varepsilon_r) \quad (5.44)$$

Substituting this into the general equation of phase stepping,

$$\tan \phi = \frac{\sum_{r=1}^R I_r \sin \Phi_r}{\sum_{r=1}^R I_r \cos \Phi_r} \quad (5.45)$$

gives

$$\tan \phi' = \frac{\sum_{r=1}^R I'_r \sin \Phi_r}{\sum_{r=1}^R I'_r \cos \Phi_r} \quad (5.46)$$

The error in the calculated value of ϕ will be

$$\Delta \phi = \arctan \frac{\sum_{r=1}^R I'_r \sin \Phi_r}{\sum_{r=1}^R I'_r \cos \Phi_r} - \arctan(\tan \phi) \quad (5.47)$$

Assuming that ε_r is small, it can be shown [17] that

$$\Delta \phi = \arctan \left\{ \frac{\sum_{r=1}^R \varepsilon_r - \sum_{r=1}^R \varepsilon_r \cos 2\Phi_r \cos 2\phi - \sum_{r=1}^R \varepsilon_r \sin 2\Phi_r \sin 2\phi}{R - \sum_{r=1}^R \varepsilon_r \cos 2\Phi_r \sin 2\phi + \sum_{r=1}^R \varepsilon_r \sin 2\Phi_r \cos 2\phi} \right\} \quad (5.48)$$

This expression is plotted in figure 5.3 for values of R from 3 to 5. The general trend is that of an error in calculated phase at double the frequency of the phase, *i.e.* at 2Φ , centred at approximately ε_r due to the dominant term $\sum_{r=1}^R \varepsilon_r / R$ in the above expression.

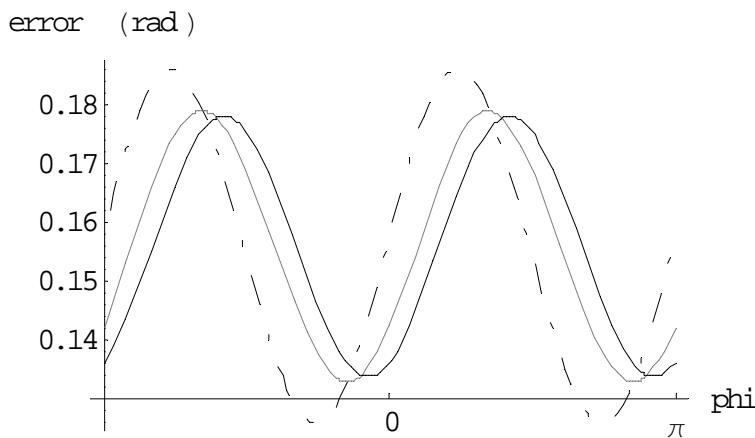


Figure 5.3 - Errors in general 3, 4 and 5-position techniques for phase step error $\varepsilon_r = \pi/20$.
The 3, 4 and 5 position techniques are represented by the dashed, dotted and solid lines, respectively

However for a given value of R , this represents a constant offset of the calculated phase which can be removed from the resulting phase map as a constant term. The error at twice the phase frequency is visible as an apparent surface undulation in the phase map. One method of minimising its effect is to increase the number of fringes across the image, and then use smoothing or filtering to remove the high frequency noise from the low frequency surface undulations.

However when measuring long objects, the increase in the number of fringes across the surface amounts to extra tilt of one of the wavefronts. This is an obliquity effect (see § 4.1.2.1) and causes an error in the measured length which is dependent on the length being measured and on the angle of the obliquity effect. Introducing too much tilt may also compromise the detection of the intensity data as the size of each fringe approaches the detector's pixel size.

The analysis of phase-shifting errors, both linear and non-linear, have been simulated by Creath [3]. The results of the simulations confirm the 2Φ nature of the error, and show that the greater the number of steps, the lower the amplitude of the error. Thus the techniques of phase-stepping and phase-shifting offer similar accuracies.

One technique to remove the errors introduced by incorrect phase-stepping is to use an additional set of interferograms to directly evaluate the size of each phase-step, using a FFT method [29]. This technique offers a repeatability in phase determination to $\lambda/500$ RMS, but requires either a nominally fixed phase-step, taken 8 times (an 8-position technique) or a 10 step technique used with random phase-steps. This technique is also prone to errors with poor fringe contrast or limited detector quantisation range.

Self calibration algorithms often experience problems when there are very few fringes across the image, as the calculations for α and ϕ can then contain numerators and denominators close to zero, leading to errors in the arctangent calculation.

5.3.7.2 Error due to detector response

A second source of error in the phase calculation is due to the response of the detector used to digitise the interferograms. For all of the techniques examined, it is assumed that the detector has a linear response, *i.e.* the digitised level of fringe intensity, I_{dig} , is linearly related to the actual intensity, I .

$$I_{dig} = \beta I \quad (5.49)$$

where β is a constant. However it is conceivable that for certain detectors, this may not be true, and non-linearities of orders 2, 3, *etc.* may be present.

$$I_{dig} = \beta I + \kappa I^2 + \rho I^3 + \dots \quad (5.50)$$

Stetson & Brohinski [30] have analysed various algorithms and non-linearities, and their results are shown in table 5.3. An asterisk indicates that the non-linearity affects the phase calculation, a blank indicates that the effect of the non-linearity is cancelled in the calculation method. The results for 2nd and 3rd order non-linearities have been confirmed by van Wingerden *et al* [31].

R	2nd order	3rd order	4th order	5th order	6th order
3	*		*	*	
4		*		*	
5			*		*

Table 5.3 - Non-linearity effects present for R -step algorithm

The effect of non-linearities on the phase calculation diminishes with the order of the non-linearity, *i.e.* the effect of a 3rd order non-linearity will be larger than that of a 4th order. The effects of orders greater than 3 are negligible, hence from the above table a minimum of 5 steps should ensure that the effects of detector non-linearities are removed from the phase calculation.

It may be argued that one could use a larger number of phase-steps and completely remove the effects of detector non-linearity, and also achieve greater averaging of the error due to the phase-step error [29], however these techniques require much more storage for the digitised images, and longer processing times. The overall resolution of the techniques is limited by vibration, air turbulence, and surface form. PSI is used to measure surface displacements of the order of nanometres, and this is approaching the dimensions of atomic spacings, approximately 0.5 nm. As the number of phase-steps, R , is increased, it is difficult to stabilise the measurement system for the longer time necessary for the extra digitisation. Thus it is rarely useful to increase the number of phase steps and complicate matters, when the technique itself is fundamentally limited to approximately $\lambda/500$ to $\lambda/1000$.

5.3.7.3 Error due to multiply-reflected beams

A third possible source of error is due to multiply reflected beams in the interferometer. This produces fringes with a profile similar to those in Fizeau interferometers. Hariharan [32] examined the effect of multiply-reflected beams by expanding the classical fringe intensity equation for a Fizeau interferometer.

$$I = I_0 \left[\frac{2R(1 - \cos \phi)}{1 + R^2 - 2R \cos \phi} \right] \quad (5.51)$$

$$= \frac{2I_0 R}{1 + R^2} \left\{ \frac{1 - \cos \phi}{1 - \frac{2R \cos \phi}{1 + R^2}} \right\} \quad (5.52)$$

assuming $\frac{2R \cos \phi}{1 + R^2}$ is small, then

$$I \approx \frac{2I_0 R}{1 + R^2} \left\{ 1 - \cos \phi + \frac{2R \cos \phi}{1 + R^2} - \frac{2R \cos^2 \phi}{1 + R^2} \right\} \quad (5.53)$$

$$I = \frac{2I_0 R}{1 + R^2} \left\{ (1 - R) - (2R - 1) \cos \phi - R \cos 2\phi \right\} \quad (5.54)$$

Inside the brackets, the first term, $(1-R)$, represents the background intensity, the second term represents the $\cos \phi$ fringes, and the third term appears as extra harmonics of $\cos 2\phi$. Note, this expression is only valid for $R \ll 1$.

For a 3-step technique, Hariharan showed this to introduce a phase error proportional to $\cos\phi \cos 2\phi$, to a first approximation. For $R = 0.05$, the maximum phase error was 6.3° . With a 4-step technique, Hariharan showed the error was reduced to 0.24° . Thus the increased number of digitised images acts as a Fourier filter, removing terms involving $\cos 2\phi$. The Fourier response of a particular 5-step technique are detailed below in § 5.4. Schwider *et al* [17] considered the effect of extraneous coherent light at a different phase to the reference and test beams. They showed the error to be periodic in the difference between ϕ and the phase of the extraneous light.

Chen & Murata [33] demonstrated a phase-stepping Fizeau interferometer, using spatial filtering to remove the effects of multiply-reflected beams, to approximate a sinusoid. Recently Bönsch & Böhme [34] have demonstrated a phase-stepping Fizeau algorithm which uses a four-position technique to solve for the 4 unknowns of the Fizeau fringe profile equation. However this technique is prone to discontinuities and errors which depend on the reflectivities of the surfaces and phase stepper accuracy [35].

5.3.7.4 Error due to quantisation noise during digitisation

The intensity of the interferogram at each point is sampled using a CCD camera and then digitised by an analogue to digital converter. The limited number of quantisation levels of the converter will introduce quantisation noise. The magnitude of the noise will be half of one digitisation level, thus the use of more levels decreases the noise. Van Wingerden *et al* [31] have derived a result for the error $\delta\phi$ in calculated phase due to quantisation noise in the digitiser for a generalised phase-stepping technique where R images are used at N bit quantisation with a fringe intensity modulation depth of m . Their result is given in equation (5.55).

$$\delta\phi = \frac{1+m}{\sqrt{3R}2^{N+1/2}m} \quad (5.55)$$

For the Primary Length Bar Interferometer which uses an 8 bit digitiser with 5 digitised images of between 0.9 and 1.0 modulation depth, the error in the measured phase is approximately 0.0015 radians (see table 5.4), equivalent to 0.00024 fringe or 0.07 nm.

N	$m = 1$	$m = 0.9$	$m = 0.5$
6 (64 levels)	0.0057	0.0060	0.0086
8 (256 levels)	0.0014	0.0015	0.0021
10 (1024 levels)	0.0004	0.0004	0.0005
12 (4096 levels)	0.0001	0.0001	0.0001

Table 5.4 - Phase measurement error (radians) due to digitisation quantisation noise for an N -bit digitiser with fringes of modulation depth m using a 5-step technique

5.4 AN ERROR-COMPENSATING FIVE POSITION TECHNIQUE

In table 5.3 above, it was seen that a 5-position technique is insensitive to low order detector non-linearities. In the appendix of their paper, Schwider *et al* [17] mention a 5-position technique, using phase step values of

$$\Phi_r = 0, \pi/2, \pi, 3\pi/2, 2\pi \quad (5.56)$$

with the phase calculated from

$$\phi = \arctan \left[\frac{2(I_2 - I_4)}{2I_3 - I_5 - I_1} \right] \quad (5.57)$$

for which they estimate an error of size $\arctan(\varepsilon/2)$, where ε is the phase-step error. However, Hariharan *et al* [36] re-calculated the error to be much smaller than this, and the conclusion of van Wingerden *et al* [31] is that the 5 position technique is always preferable to the 4 position technique as the measurement errors are the same or better, and the formula takes less computation time.

To analyse these findings it is necessary to derive the 5-position equation. For ease of derivation, assume that the phase steps have relative phases of $-2\alpha, -\alpha, 0, \alpha, 2\alpha$. The intensity at a point in the interferogram can be written as a simple function of the interference of two beams, with the two beam intensities A and B , and five intensity values I_1 to I_5 , corresponding to the above phase shifts.

$$\begin{aligned} I_1 &= A + B + 2\sqrt{AB} \cos(\phi - 2\alpha) \\ I_2 &= A + B + 2\sqrt{AB} \cos(\phi - \alpha) \\ I_3 &= A + B + 2\sqrt{AB} \cos(\phi) \\ I_4 &= A + B + 2\sqrt{AB} \cos(\phi + \alpha) \\ I_5 &= A + B + 2\sqrt{AB} \cos(\phi + 2\alpha) \end{aligned} \quad (5.58)$$

i.e. the reference phase Φ_r takes values of $-2\alpha, -\alpha, 0, \alpha, 2\alpha$ for $r = 1$ to 5 respectively.

Then it can be seen that the expression $\frac{I_2 - I_4}{2I_3 - I_5 - I_1}$ is equal to

$$\frac{A + B + 2\sqrt{AB} \cos(\phi - \alpha) - A - B - 2\sqrt{AB} \cos(\phi + \alpha)}{2A + 2B + 4\sqrt{AB} \cos(\phi) - A - B - 2\sqrt{AB} \cos(\phi + 2\alpha) - A - B - 2\sqrt{AB} \cos(\phi - 2\alpha)}$$

and then using the sum of angles relation for cosine,

$$\cos(x+y) = \cos(x)\cos(y) - \sin(x)\sin(y)$$

gives

$$\begin{aligned} \frac{I_2 - I_4}{2I_3 - I_5 - I_1} &= \frac{\sin \alpha \sin \phi}{(1 - \cos 2\alpha) \cos \phi} \\ &= \tan \phi \left(\frac{\sin \alpha}{1 - \cos 2\alpha} \right) \end{aligned} \quad (5.59)$$

The phase step factor $\left(\frac{\sin \alpha}{1 - \cos 2\alpha} \right)$ has a value of 0.5 at $\alpha = 90^\circ$ and does not depart from this value for small deviations in α from 90° . If α remains between 86° and 94° then the value of this factor does not alter by more than 0.001 and can be assumed to be constant, see figure 5.4. Assuming a value of 0.5 allows equation (5.59) to be simplified, leading to (5.57).

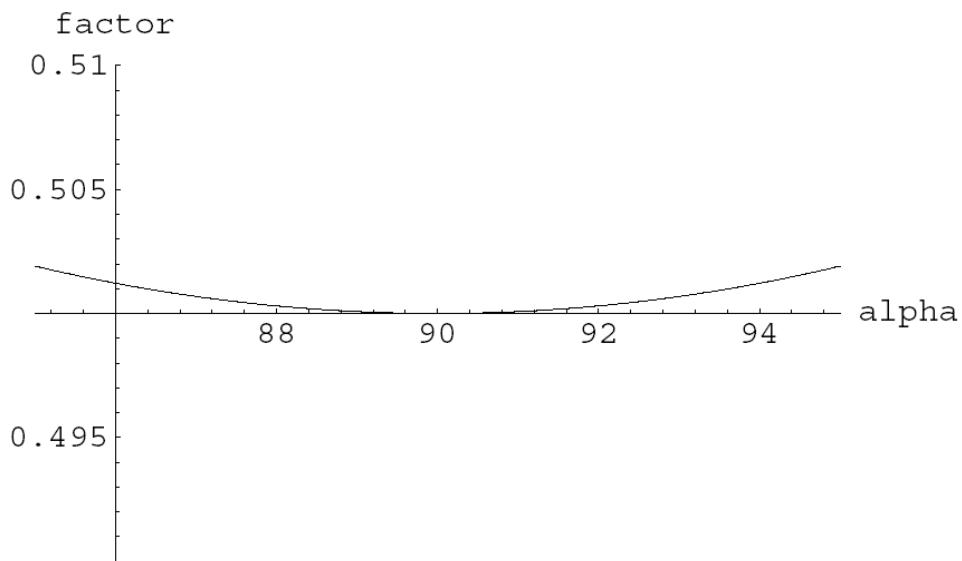


Figure 5.4 - Variation of phase step factor as α is varied

In fact, if we assume a phase step error of ε , then

$$\alpha = \pi / 2 + \varepsilon \quad (5.60)$$

and

$$\tan \phi' \approx (1 + \varepsilon^2 / 2) \tan \phi \quad (5.61)$$

$$\Delta\phi = \phi - \phi' = \frac{\varepsilon^2}{4} \sin(2\phi) \quad (5.62)$$

Thus the error in the phase calculation has the expected 2ϕ dependence, but its magnitude is one quarter of the square of the original phase-step error. As an example, if $\varepsilon = 1^\circ$, then the maximum error $\Delta\phi = 0.02^\circ$. For a double pass interferometer where each fringe corresponds to approximately 316 nm path difference, this amounts to an error in the surface or length measurement of 0.02 nm.

This approximation can be checked by a more rigorous approach.

Assuming $\alpha = \pi/2 + \varepsilon$, then

$$I_1 = A + B + 2\sqrt{AB} \cos(\phi - \pi - 2\varepsilon)$$

$$I_2 = A + B + 2\sqrt{AB} \cos(\phi - \pi/2 - \varepsilon)$$

$$I_3 = A + B + 2\sqrt{AB} \cos(\phi) \quad (5.63)$$

$$I_4 = A + B + 2\sqrt{AB} \cos(\phi + \pi/2 + \varepsilon)$$

$$I_5 = A + B + 2\sqrt{AB} \cos(\phi + \pi + \varepsilon)$$

$$\tan \phi' = \frac{2[\cos(\phi - \pi/2 - \varepsilon) - \cos(\phi + \pi/2 + \varepsilon)]}{2\cos \phi - \cos(\phi + \pi + \varepsilon) - \cos(\phi - \pi - 2\varepsilon)} \quad (5.64)$$

$$= \frac{2[\sin(\phi - \varepsilon) + \sin(\phi + \varepsilon)]}{2\cos \phi + \cos(\phi + 2\varepsilon) + \cos(\phi - 2\varepsilon)} \quad (5.64)$$

$$= \frac{2 \sin \phi \cos \varepsilon}{\cos \phi + \cos \phi \cos 2\varepsilon} \quad (5.65)$$

$$\therefore \tan \phi' = \tan \phi \left(\frac{2 \cos \varepsilon}{1 + \cos 2\varepsilon} \right) \quad (5.65)$$

For the above phase step error of 1° , this expression predicts a maximum error in the calculated value of ϕ to be 0.05° , similar to the approximate result above.

The expression for $\arctan \phi$ (5.57) is such that it is impossible for both the numerator and denominator to be simultaneously zero, with sinusoidal fringes, so no accuracy is lost due to small angle problems.

It is possible to calculate α from the intensity data, allowing a check on the performance of the phase-stepping of the reference mirror:

$$\cos \alpha = \frac{I_5 - I_1}{2(I_4 - I_2)} \quad (5.66)$$

α should have a uniform value of 90° over the measurement surface if the phase-stepping has been performed correctly. Any tilting of the reference mirror during phase-stepping can be identified, as can incorrect calibration of the phase-step size.

Recent work by Larkin and Oreb [37] has shown this 5-position technique to be one of a class of ‘N+1 symmetrical’ techniques. Using Fourier analysis of the effective sampling algorithms, *i.e.* the step positions for which the intensity is digitised, they have shown that the frequency response of the numerator and denominator of equation (5.57) have certain features which make the algorithm insensitive to certain errors:

- The numerator has stationary points at the fundamental fringe frequency, and at odd-multiples of this frequency. Thus, at these frequencies, the numerator is insensitive to phase-step errors (which produce a frequency slightly different to the fundamental frequency).
- The numerator also has zeroes at all even-multiples of the fundamental frequency, making it insensitive to even-order detector non-linearities.
- The denominator has stationary points at the fundamental frequency and all multiples. Thus the denominator is insensitive to phase-step errors.
- The denominator also has zeroes at the even-harmonics, and hence is not affected by even-order detector non-linearities.

Hence the overall technique is insensitive to even-order detector non-linearities and phase-step errors, particularly those close to the fundamental fringe frequency, *i.e.* small phase-step errors, as demonstrated above in figure 5.4. It is simple to implement and provides a self-check of attained phase-step values.

5.5 IMPLEMENTATION OF THE FIVE POSITION TECHNIQUE IN THE PRIMARY INTERFEROMETER

Due to the advantages summarised above, the 5 position technique was chosen for use in the analysis of the interference patterns in the Primary Length Bar Interferometer. A phase-step of size $\pi/2$ is provided by moving the mirror in the reference arm of the interferometer by 1/4 of a fringe (at $\lambda = 633$ nm, this is equal to 79 nm). Problems of incorrect phase-stepping have been overcome by design of the mirror mount (see § 3.2.3), and by using a commercial PZT system which uses capacitive sensing to maintain the PZT calibration. The PZT can be moved in steps of size 1.07 nm by setting the digital offset in the control electronics by computer control.

The phase-stepping is performed as follows. Firstly the PZT is positioned at the centre of its range (digital offset = 0). The red laser is selected. After a 2 second pause, the image is digitised. The PZT is moved to the next position (offset = 74), and allowed to stabilised for 0.25 sec before the 2nd image is captured. The PZT is then moved to the 3rd position (offset = 148) and stabilised before the image is captured. This is repeated until 5 images have been digitised for the red wavelength.

The red laser is de-selected, the green laser selected and the PZT positioned back at the starting position. The process of digitising the image, moving the mirror, stabilising, *etc.* is repeated for the green wavelength and then for the orange wavelength.

The size of each step is adjusted for the wavelength being used. The whole 3-wavelength phase-stepping procedure lasts approximately 7 seconds.

Equation (5.66) is used to calculate the exact phase step at each non-masked pixel, and the average of the values from all non-masked pixels provides a check on the calibration of the PZT movement and the accuracy of the phase-stepping. Any tilt of the mirror during stepping can be seen in the α map as a change in phase step angle. In practice the phase-stepping is very reliable and usually phase-step correction is unnecessary. Average phase step sizes are $90^\circ \pm 2^\circ$, leading to a maximum phase measurement error of 0.06° or 0.05 nm.

REFERENCES FOR CHAPTER 5

- [1] Brown G M Fringe analysis for automotive applications *Proceedings of conference Fringe Analysis '92*, Leeds 1992, (London: Institute of Physics)
 - [2] Ascough J Applications of three-dimensional photoelasticity in hip joint prosthesis design *Proceedings of conference Fringe Analysis '92*, Leeds 1992, (London: Institute of Physics)
 - [3] Creath K Phase measurement interferometry techniques *Progress in Optics* (Elsevier: North Holland) **26** (1988) 349-393
 - [4] Hariharan P Optical Interferometry *Rep. Prog. Phys.* **54** (1991) 339-390
 - [5] Hariharan P Interferometry with lasers *Progress in Optics* (Elsevier: North Holland) **24** (1987) 103-164
 - [6] Reid G T Automatic fringe pattern analysis: a review *Optics and Lasers in Engineering* **7** (1986) 37-68
 - [7] Robinson D W Automatic fringe analysis in optical metrology *Appl. Opt.* **22** (1983) 2169-2176
 - [8] Schwider J Advanced evaluation techniques in interferometry *Progress in Optics* (Elsevier: North Holland) **28** (1990) 217-359
 - [9] Budzinski J SNOP: a method for fringe skeletonization of a fringe pattern along a fringe direction *Appl. Opt.* **31** (1992) 3109-3113
 - [10] Parthiban V & Sirohi R S Interactive fringe processing algorithm for interferogram analysis *Optics and Lasers in Engineering* **11** (1989) 103-113
 - [11] Hou W & Wilkening G Investigation and compensation of the nonlinearity of heterodyne interferometers *Precis. Eng.* **14** (1992) 91-98
 - [12] Takeda M Spatial-carrier fringe-pattern analysis and its applications to precision interferometry and profilometry: an overview *Industrial Metrology* **1** (1990) 79-99
 - [13] Carré P Installation et utilisation du comparateur photoélectrique et interférentiel du Bureau International des Poids et Mesures *Metrologia* **2** (1966) 13-23
 - [14] Bruning J H, Herriott D R, Gallagher J E, Rosenfeld D P, White A D & Brangaccio D J Digital wavefront measuring interferometers for testing optical surfaces and lenses *Appl. Opt.* **13** (1974) 2693-2703
 - [15] Sommargren G E Optical heterodyne profilometry *Appl. Opt.* **20** (1981) 610-618
-

- [16] Wyant J C Use of an ac heterodyne lateral shear interferometer with real-time wavefront corection systems *Appl. Opt.* **14** (1975) 2622-2626
- [17] Schwider J, Burow R, Elssner K E, Grzanna R, Spolaczyk R & Merkel K Digital wavefront measuring interferometry: some systematic error sources *Appl. Opt.* **22** (1983) 3421-3432
- [18] Grievenkamp J E Generalized data reduction for heterodyne interferometry *Opt. Eng.* **23** (1984) 350-352
- [19] Cheng YY & Wyant J C Two wavelength phase shifting interferometry *Appl. Opt.* **23** (1984) 4539-4543
- [20] Pettyjohns K N, De Vore S, Deraniak E & Wyant J C Direct phase measurement interferometer working at $3.8\text{ }\mu\text{m}$ *Appl. Opt.* **24** (1985) 2211-2216
- [21] Creath K Step height measurement using two-wavelength phase-shifting interferometry *Appl. Opt.* **26** (1987) 2810-2816
- [22] Barnes T H Heterodyne Fizeau interferometer for testing optical flats *Appl. Opt.* **26** (1987) 2804-2809
- [23] Srinivasan V, Liu H C & Halioua M Automated phase measuring profilometry of 3D diffuse objects *Appl. Opt.* **23** (1984) 3105-3108
- [24] Shagam R N & Wyant J C Optical frequency shifter for heterodyne interferometers using multiple rotating polarisation retarders *Appl. Opt.* **17** (1978) 3034-3035
- [25] Colucci D & Wizinowich P Millisecond phase acquisition at video rates *Appl. Opt.* **31** (1992) 5919-5925
- [26] Kothiyal M P & Delisle C Rotating analyzer heterodyne interferometer: error sources *Appl. Opt.* **24** (1985) 2288-2290
- [27] Grievenkamp J E Generalized data reduction for heterodyne interferometry *Opt. Eng.* **23** (1984) 350-352
- [28] Santoyo F M, Kerr D & Tyrer J R Interferometric fringe analysis using a single phase step technique *Appl. Opt.* **27** (1988) 4362-4364
- [29] Lai G & Yatagai T Generalized phase-shifting interferometry *J. Opt. Soc. Am. A* **8** (1991) 822-827
- [30] Stetson K A & Brohinski W R Electrooptic holography and its application to hologram interferometry *Appl. Opt.* **24** (1985) 3631-3637
- [31] van Wingerden T, Franken H J, Smorenburg C Linear approximation for measurement errors in phase shifting interferometry *Appl. Opt.* **30** (1991) 2718-2729

- [32] Hariharan P Digital phase-stepping interferometry: effects of multiply reflected beams *Appl. Opt.* **26** (1987) 2506-2507
- [33] Chen J & Murata K Digital phase measuring Fizeau interferometer for testing of flat and spherical surfaces *Optik* **81** (1988) 28-32
- [34] Bönsch G & Böhme H Phase-determination of Fizeau interferences by phase-shifting interferometry *Optik* **82** (1989) 161-164
- [35] Nicolaus R A Evaluation of Fizeau interferences: a comparison of phase-stepping algorithms *Proc. SPIE* **1319** (1990) 237-238
- [36] Hariharan P, Oreb B F & Eiju T Digital phase-shifting interferometry: a simple error-compensating phase calculation algorithm *Appl. Opt.* **26** (1987) 2504-2505
- [37] Larkin K G & Oreb B F Design and assessment of symmetrical phase-shifting algorithms *J. Opt. Soc. Am. A* **9** (1992) 1740-1748

CHAPTER 6

DATA PROCESSING

“Science cannot solve the ultimate mystery of nature. And it is because in the last analysis we ourselves are part of the mystery we are trying to solve.”

M Planck

6.1 OVERVIEW OF THE DATA PROCESSING

Once the phase-stepping has been performed as detailed in Chapter 5 and parameters such as bar temperature, air temperature, pressure, humidity and CO₂ content have been measured, the data must be processed to produce results for: the central length of the bar (corrected to 20 °C), the form of the exposed measurement face, the flatness of the face and its parallelism with respect to the wrung face. This processing forms the major part of the computer program which controls the interferometer. The majority of the computer processing is devoted to extracting the phase data into a form in which it can be used in the multiple-wavelength analysis, to calculate the length of the bar.

The stages of the data processing are illustrated in figure 6.1. The raw data is stored in the framestore in the form of 15 images and a mask. The images are stored as 256 x 256 pixel arrays at a resolution of 8 bits (256 levels). The phase data is calculated from the images, one wavelength at a time, by applying the phase stepping equation (5.57) to each set of 5 images. The 3 phase maps then contain phase data in the range - π to + π , including discontinuities of magnitude 2π at the boundaries between fringes and a discontinuity at the edge of the bar. The discontinuities are then removed by a 3-pass routine developed specifically for the interferometer. A surface is fitted to the data of the platen to account for any deviation from flatness so that the phase of the fitted surface can be subtracted from the phase data over the whole image to remove tilt. The resulting phase maps are scaled to fringe fractions by dividing by 2π . The phase maps are then representations of the difference in phase between the measured phase and the phase fitted to the surface of the platen, *i.e.* the phase data on the surface of the bar is now directly related to the fringe fractions required for length calculation.

The red phase map is then used to calculate the flatness of the exposed face of the bar and its parallelism to the wrung face by fitting a plane to the phase data. Fringe fractions at the centre of the bar are averaged for each of the 3 wavelengths and the

resulting fractions used in the multiple-wavelength calculations. The results are displayed on the screen with the option for hardcopy.

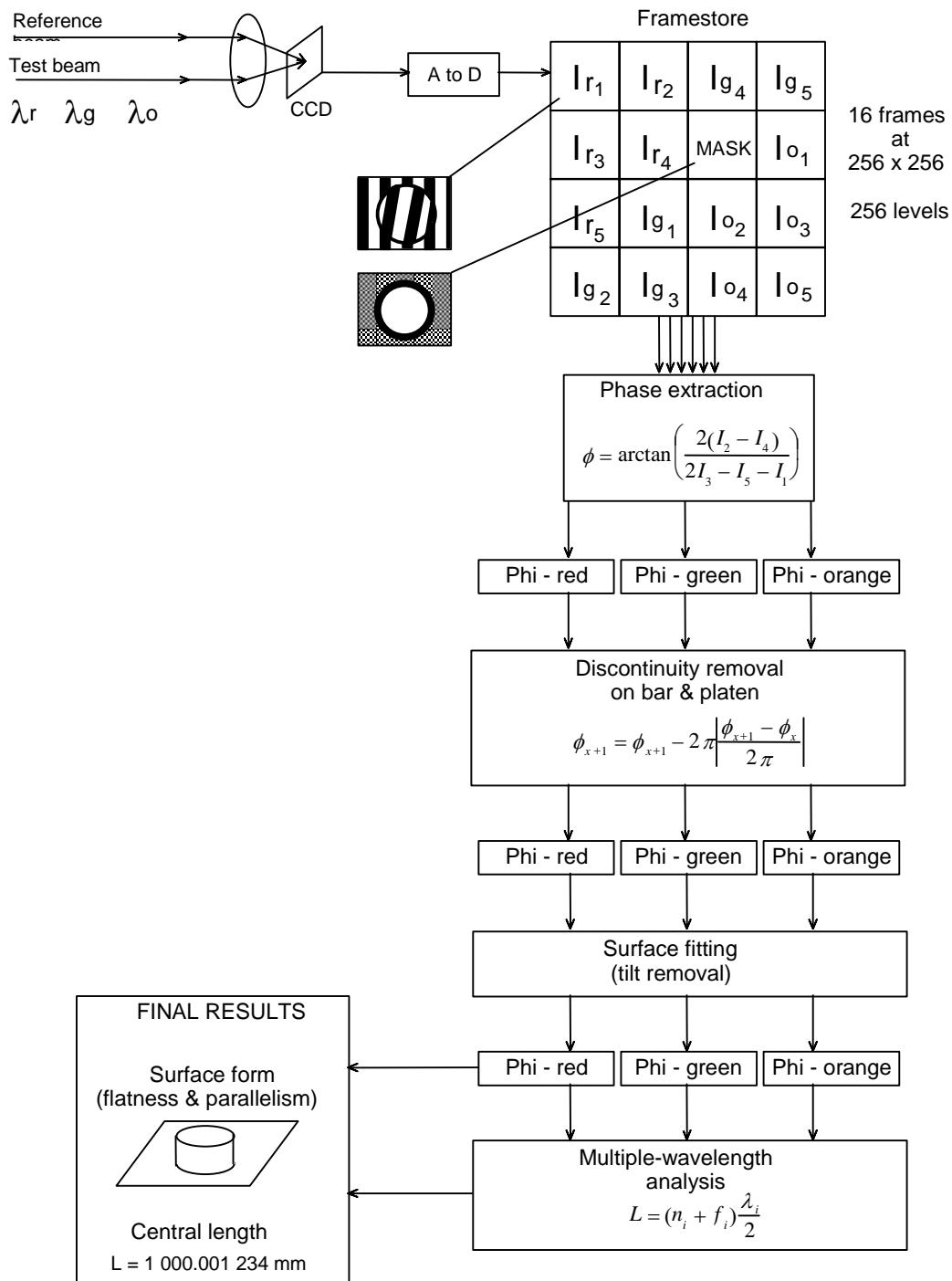


Figure 6.1 - Flow diagram of data processing

6.2 COMPUTING SYSTEM

The original computer used for the data analysis and instrument control was a Hewlett-Packard Vectra ES12, an 80286 PC compatible with a maths co-processor, VGA compatible display and a clock speed of 12 MHz. A dot-matrix printer was used for hard-copy. The following cards were installed inside the PC: A Matrox PIP1024B framestore, a CEC IEEE interface card, an Amplicon PC14A digital I-O card. Eventually the speed of the computer proved to be a limiting factor: a compilation of the program took up to 8 minutes and a single measurement (including set-up and analysis) took 6 minutes to perform.

The computer was replaced with an Elonex PC450, an 80486 DX2 system, operating at a clock-doubled speed of 50 MHz, with an on-board co-processor and 64 K cache. An S-VGA compatible monitor and a Hewlett-Packard DeskJet 550C colour inkjet printer were added.

With the PC450, it was possible to place DOS in High Memory and to set up a RAMDRIVE and SMARTDRIVE disc cache, enabling much faster compilation and program execution. The compilation time was reduced to 31 seconds and the measurement time to 2 minutes 17 seconds. The actual time taken for the calculations is 45 seconds, with the remainder of the time required for the setting up, phase-stepping and temperature measurements.

The program is approximately 4500 lines of Pascal code which is compiled by a Microsoft Pascal 4.0 compiler into an executable file of size 140 K. Microsoft Pascal was chosen as the programming language as it is well structured and could interface with the libraries of routines provided with the interface boards used inside the PC which were supplied as compiled Microsoft Pascal and C object modules.

The major limitation of Microsoft Pascal is that there is a maximum limit of 64 K allowed for data, with no single data structure allowed to be greater than this limit. Thus there were two problems for the data processing. Firstly, a 256 x 256 array of REAL numbers (which would be required for each 256 x 256 phase map) would be stored as 4 bytes per number, thus resulting in an array size of 256 K. Thus the images had to be sampled at 128 x 128 resolution to result in 64 K phase arrays. Secondly, since a maximum of 64 K was allowed for ALL variables, some alternative method had to be used to store the 3 phase maps and the phase-step map. These arrays were placed high in memory, *i.e.* outside the default data segment. Unfortunately this meant that they were not protected from being violated by other programs and some further programming was necessary to avoid clashes. Microsoft Pascal is not able to use extended or expanded memory, so only a maximum of 640 K was available to the program.

6.3 IMAGE PROCESSING

6.3.1 Interferogram digitisation

Each interferogram is digitised by being imaged onto the CCD array of a Sony AVC-D5CE monochrome video camera. The array size is 8.8 mm x 6.6 mm at 500 x 582 pixels. The image of the interferogram over-fills the array and only the central region is digitised - this avoids the inclusion of diffracted beams at the edge of the image. The camera is synchronised to the monitor signal derived from the Matrox framestore board ensuring that the image position is fixed with respect to the framestore pixels. The camera and framestore are connected with standard $75\ \Omega$ BNC cable.

The Matrox framestore is configured as a single store of size 1024 x 1024 pixels, with the zoom option enabled allowing digitisation to a 256 x 256 image (total 16 images) at 8 bit resolution. The organisation of the 16 image areas in the framestore is shown in figure 6.2. Fifteen images are used for storing the interferograms for the 3 wavelengths and the remaining image is used to store the 3-level mask. Access to the intensity data stored on the framestore is via library routines with speed of access below that of direct memory access of main computer RAM hence all calculations are performed on arrays stored in conventional memory.

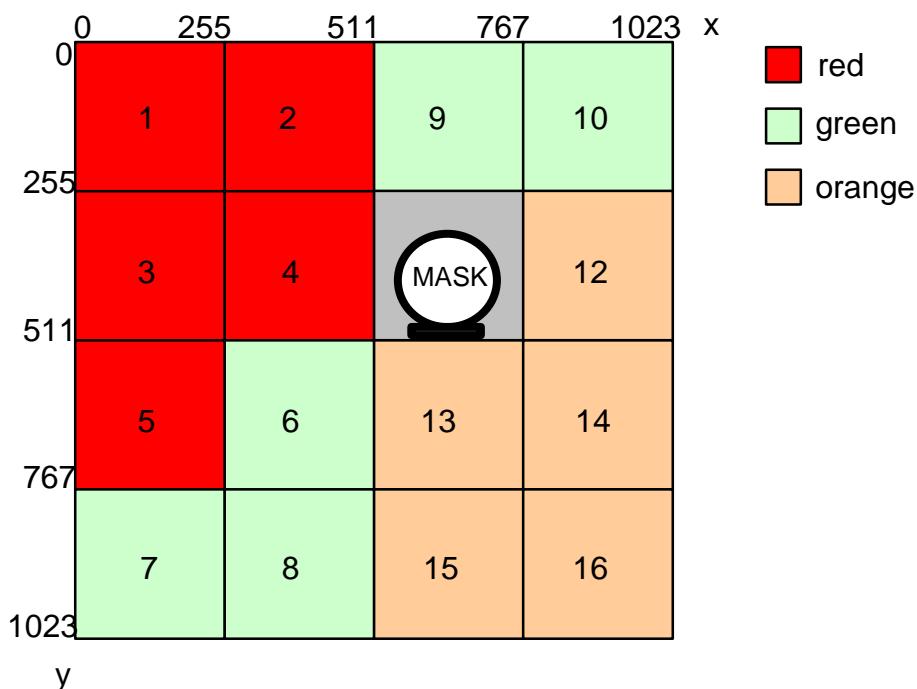


Figure 6.2 - Organisation of framestore memory

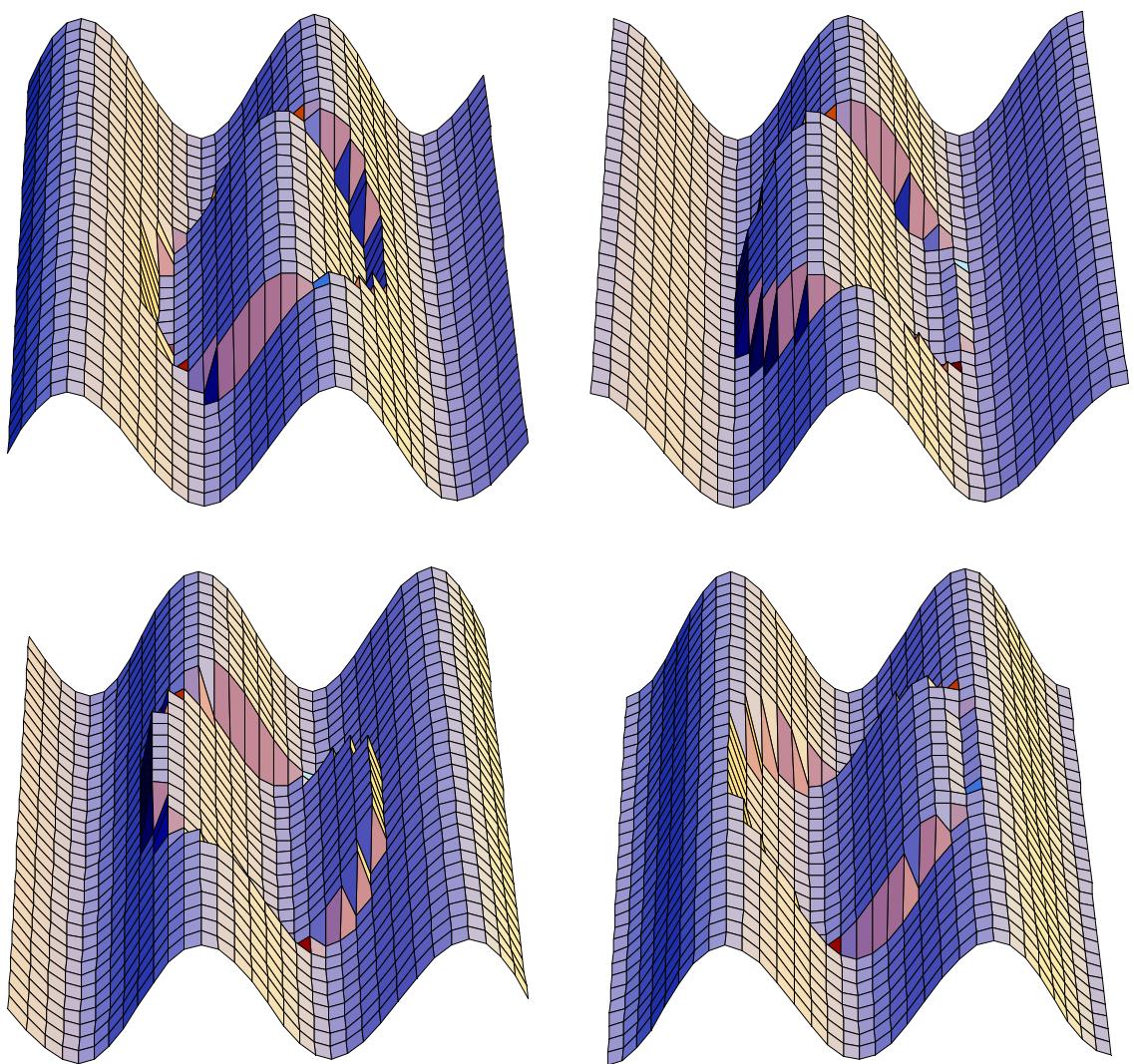


Figure 6.3 - Simulated intensity arrays for the first 4 digitised images

The gain and offset of the analogue to digital converter on the framestore card can be manually adjusted using the computer program so that most of the 256 level range of the framestore is used. Typical histograms of digitised interferograms are shown in figures 6.4 & 6.5. The extra peak at level 128 is due to the areas in the image where the surfaces are not aligned or smooth enough for interference to be visible, such as the supports, or the edge of the bar, and hence have an average intensity of half the digitised range. The widths of the low and high level peaks are dependent on the number of fringes in view: by adjusting the numbers of fringes it is possible to alter the relative widths of these peaks. When the fringes are adjusted such that a bright fringe is at both the left and right edges of the screen, the high intensity peak is widest. The most important feature is that almost the whole range of the digitiser is used, decreasing the

noise in the phase measurements due to quantisation noise in the digitiser (see § 5.3.7.4), but without the peaks being truncated at the extremes of the digitisation range.

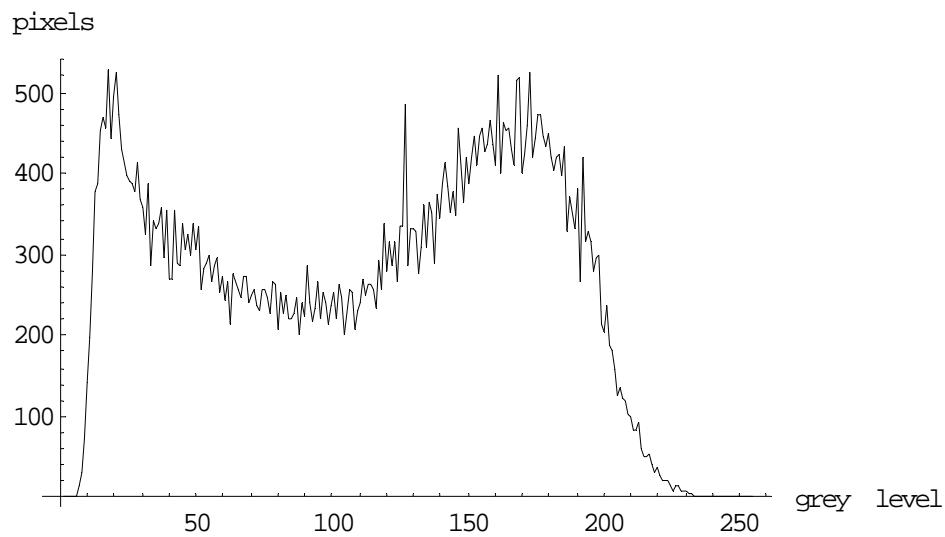


Figure 6.4 - Typical digitisation histogram showing number of pixels within given intensity levels

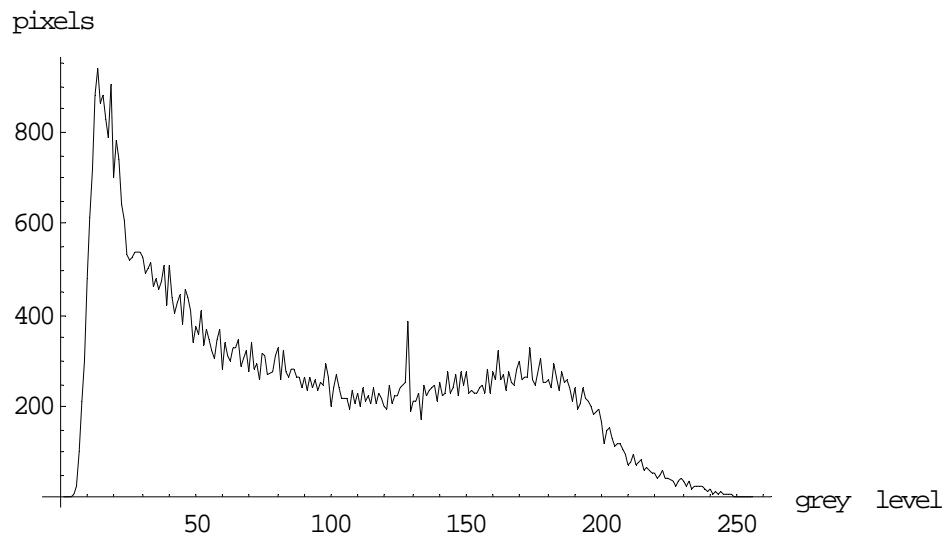


Figure 6.5 - Typical digitisation histogram showing number of pixels within given intensity levels, but with more dark fringes in the image

6.3.2 Phase extraction

After image digitisation and measurement of temperature, pressure *etc*, the phase is extracted from the 15 digitised images, one wavelength at a time by applying equation

6.1 to the intensity data stored in the framestore. The phase is extracted on a pixel-by-pixel basis, starting at the top left pixel, and progressing along to the end of the row, before extracting the next row. The results are stored in 3 arrays, one for each wavelength (`phi_red`, `phi_green`, `phi_orange`). Next, use is made of a 4-quadrant arctangent routine (described in § 5.3.3) which returns a value in the range $-\pi$ to $+\pi$ depending on the sign of the numerator and denominator in equation 6.1. At each pixel, the actual phase-step value (nominally 90° or $\pi/2$) is also calculated and stored in an array (`alpha`).

$$\phi(x, y) = \arctan \left[\frac{2(I_2(x, y) - I_4(x, y))}{2I_3(x, y) - 2I_5(x, y) - I_1(x, y)} \right] \quad (6.1)$$

After phase extraction, the three arrays `phi_red`, `phi_green` and `phi_orange` contain phase information including tilt and 2π discontinuities at the boundary between fringes. A simulated example of one of these phase maps is shown in figure 6.6.

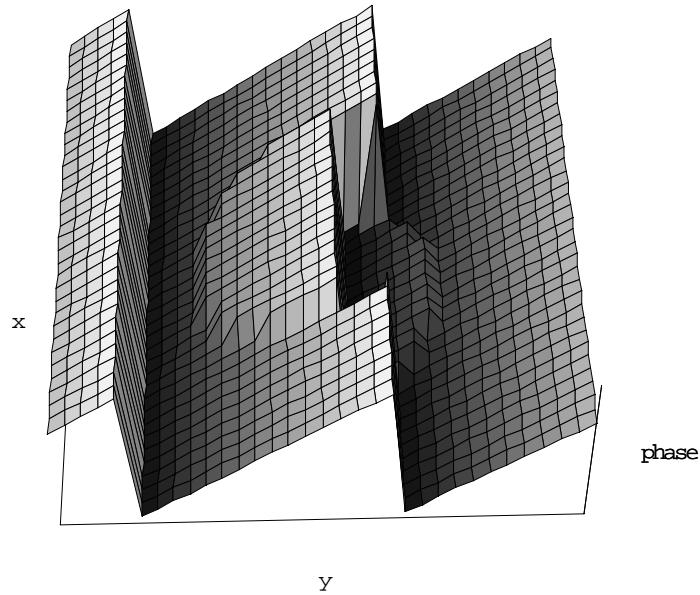


Figure 6.6 - Simulated phase map containing 2π discontinuities, tilt, and phase difference due to the bar

6.3.3 Discontinuity removal

The modulo 2π discontinuities in each phase-map are removed on a line-by-line basis using a three-pass routine. Initially, a three-level mask is generated by the user and stored in the framestore before any measurements are made. This mask is used in the phase-unwrapping to distinguish between (i) data on the platen, (ii) data on the end of the bar and (iii) invalid data such as at the edges of the bar which are radiusied.

During the phase-unwrapping, the phase is unwrapped first for data corresponding to the surface of the platen. The first line to be unwrapped is the top horizontal line of the image. The unwrapping algorithm scans across this line on a pixel-by-pixel basis, removing discontinuities of magnitude $\sim 2\pi$ by adding or subtracting multiples of 2π to the pixel following the discontinuity.

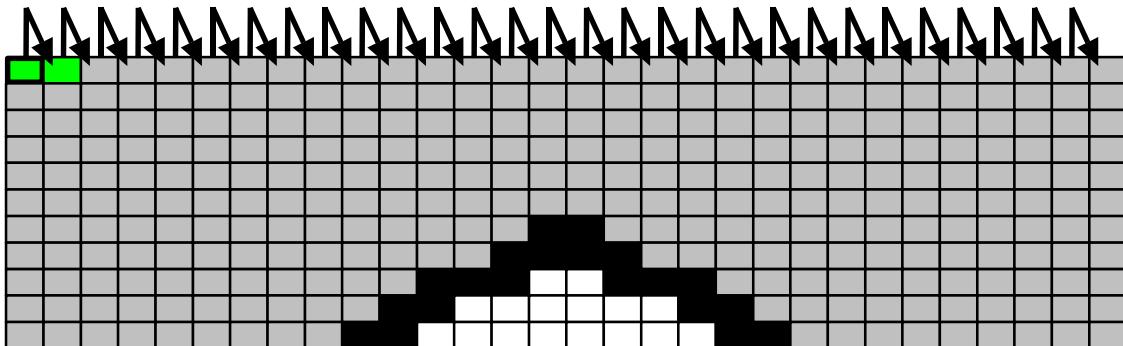


Figure 6.7(a) - Discontinuity removal 1st pass - removal of discontinuities across top line of image using pixel-by-pixel comparison

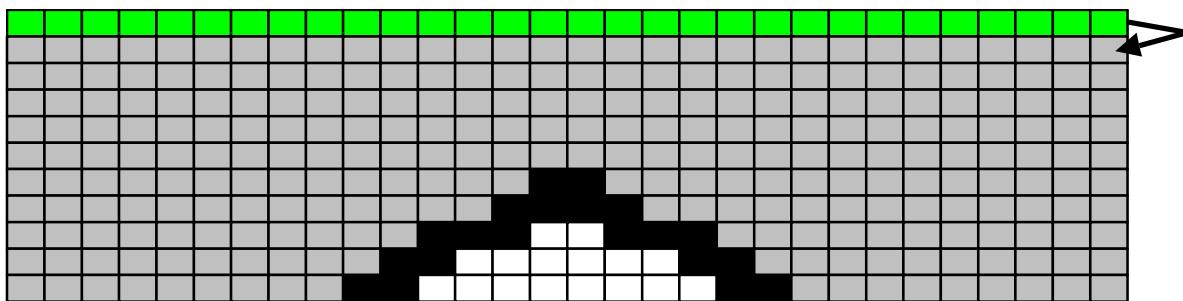


Figure 6.7(b) - Discontinuity removal 1st pass - comparison of phase of top line with next line below

$$\phi_{x+1} = \phi_{x+1} - 2\pi(\phi_{x-1} - \phi_x) \bmod 2\pi \quad (6.2)$$

This unwrapped line is used as a reference for the rest of the data on the platen. The unwrapping algorithm is thus dependent on this line being a section through smoothly-varying phase values corresponding to the surface of the platen. More generalised algorithms exist which can cope with discontinuous phase [1,2,3] though they are generally more computationally intensive.

The phase values of the next line below are compared to those of the unwrapped line, and adjustments made as necessary. This procedure continues down the image. Each pixel is checked before being unwrapped to verify that it corresponds to valid data from the platen. This first pass halts at the centre of the image.

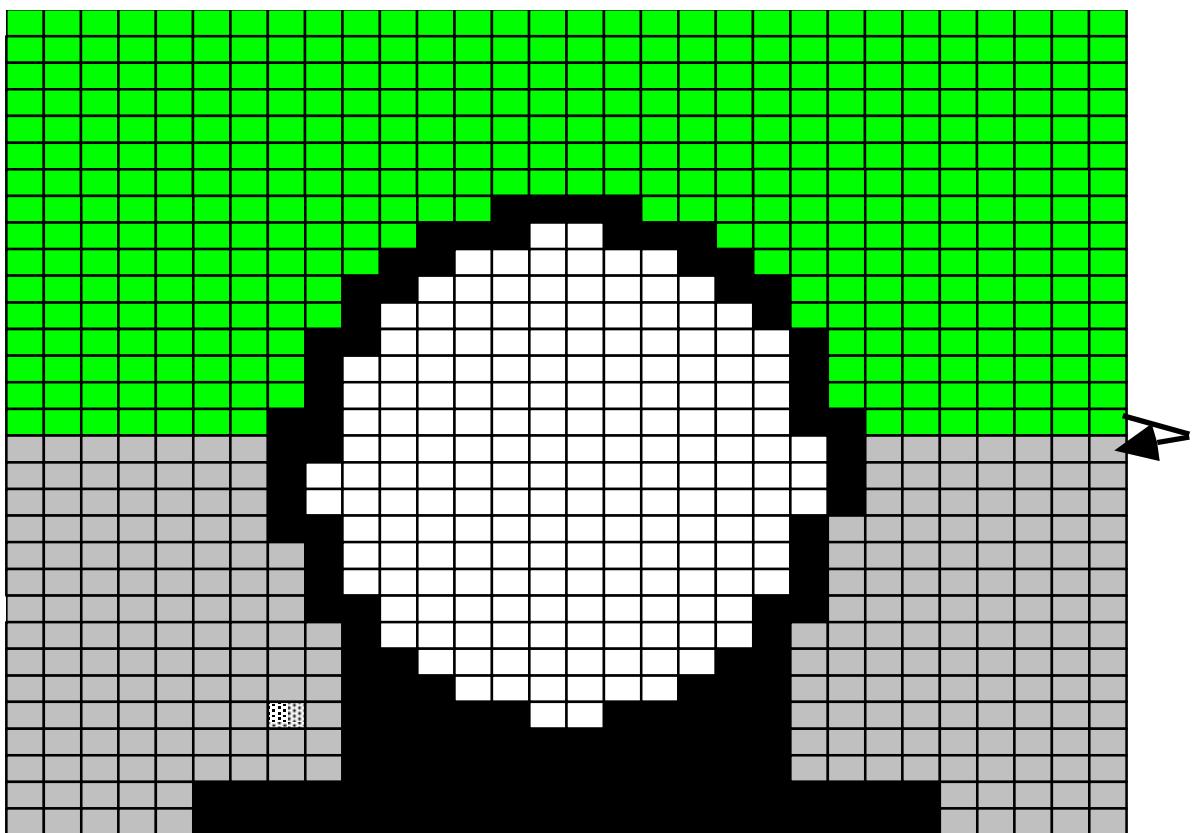


Figure 6.7(c) - Discontinuity removal 1st pass - continuation until middle of image

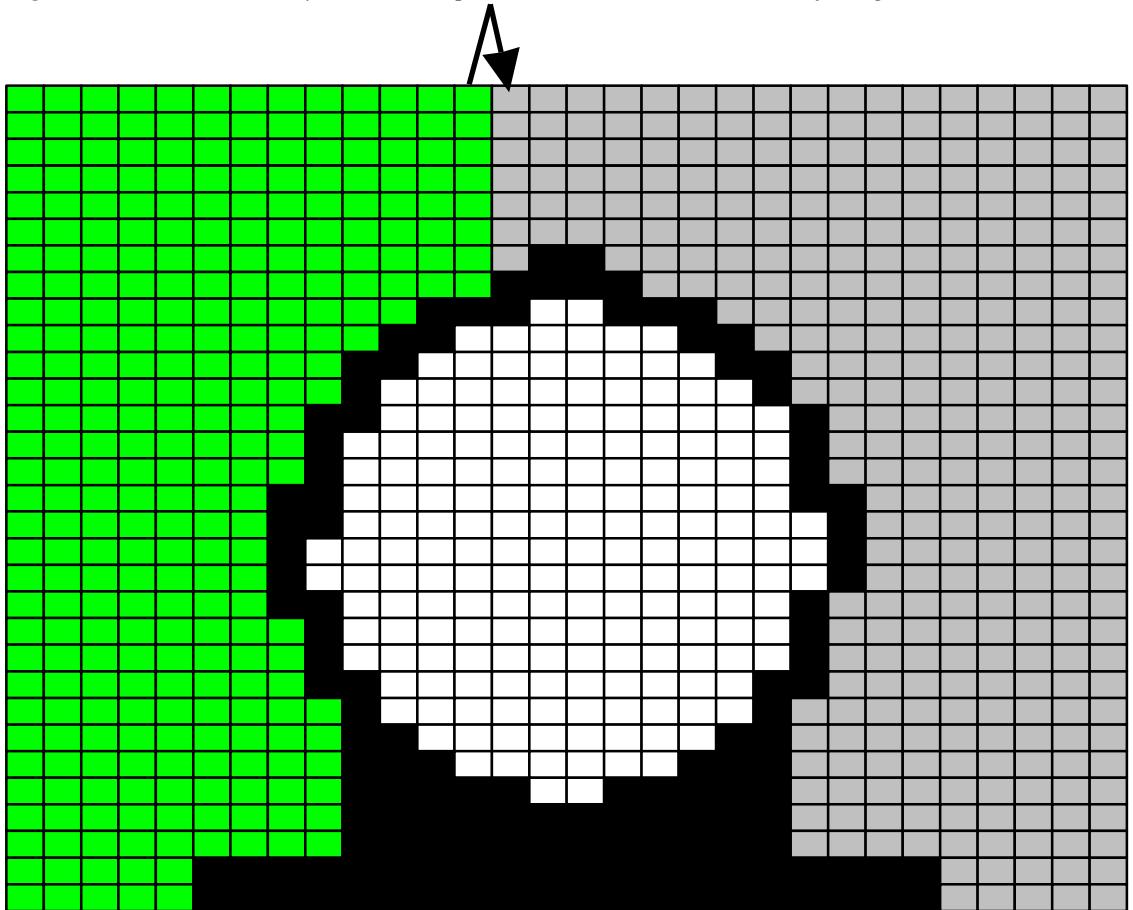


Figure 6.8 - Discontinuity removal 2nd pass from left to right

The second pass starts by unwrapping the vertical line along the left edge of the image. This line is then used as a reference for subsequent lines using the same procedure as the first pass. The second pass stops at the centre of the image.

The third pass is the same as the second, though starting from the right edge of the image. The three pass algorithm thus fits together the phase data around the edge of the image, then moves inwards towards the bar.

The phase data of the bar are unwrapped using a similar two pass algorithm, starting at the centre of the bar and then moving upwards and downwards.

After discontinuity removal the three phase maps are smoothly-varying but contain tilt due to the presence of tilt fringes in the original images, as shown in figure 6.9.

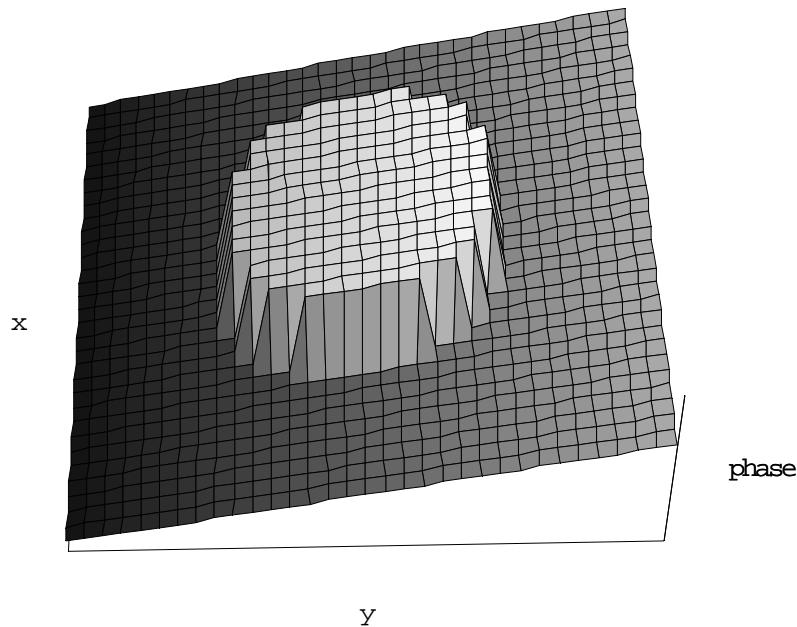


Figure 6.9 - Simulated phase map after discontinuity removal, showing tilt and phase difference due to bar

The tilt in each phase map is removed by fitting a suitable polynomial surface to the phase data of the platen and then subtracting this fitted surface from the measured data. The results of the image processing are three phase maps, at the three measurement wavelengths corresponding to the difference between the measured phase and the fitted surface, at each point in the image. Thus the phase maps represent (1) the deviation of the platen from the fitted surface and (2) the phase of the end of the bar with respect to

the fitted surface, *i.e.* the length of the bar, as defined in BS 5317. Typical phase maps after discontinuity and tilt removal are shown in figure 6.10.

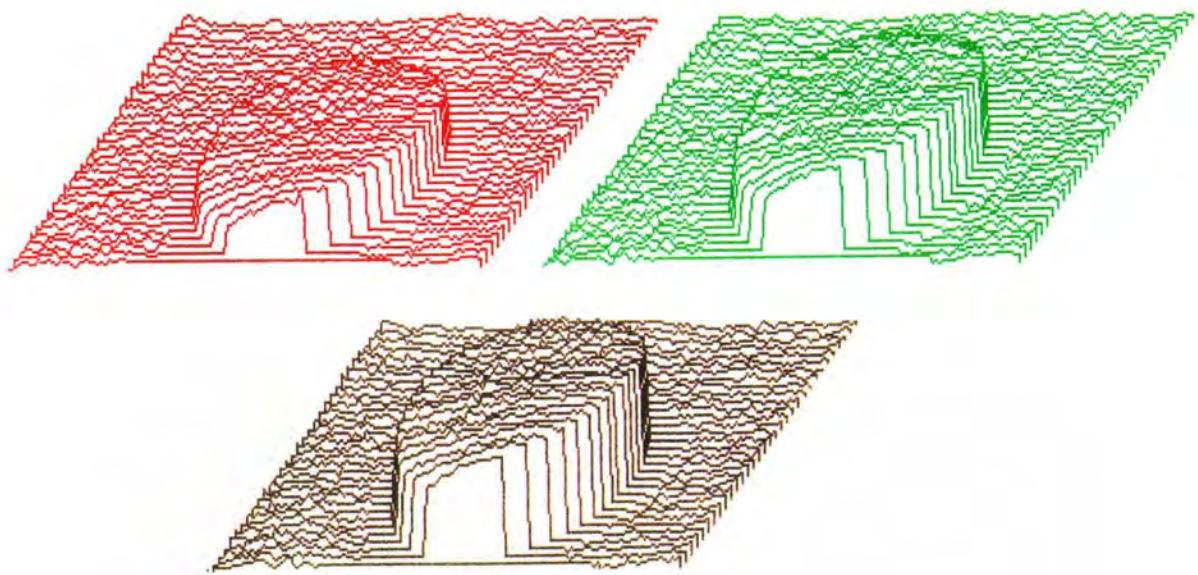


Figure 6.10 - Phase maps after discontinuity and tilt removal,
clockwise from top left: 633 nm, 543 nm, 612 nm

The values in the three phase maps are converted to fringe fractions by dividing by 2π . Statistics of the surface measurements such as peak-valley variation, flatness and parallelism of the faces are easily calculated from the phase maps. In practice, these parameters are calculated from only the 633 nm phase map, `phi_red`.

For the calculation of the length of the bar, the fringe fractions of 9x9 pixels at the centre of the bar are averaged for each of the three wavelengths. These three fractions, f_1 , f_2 and f_3 corresponding to the red, green and orange wavelengths respectively, can be combined in the technique of multiple-wavelength interferometry to calculate the length of the bar, based on solution of equation (6.3).

$$L = (n + f) \lambda' / 2 \quad (6.3)$$

6.4 MULTIPLE-WAVELENGTH INTERFEROMETRY

6.4.1 Multiple wavelength analysis

The reason for using three wavelengths rather than just one will now be explained. With one wavelength the corresponding fringe fraction, f , can be measured by the

interferometer. In order to solve equation 6.3, it is then necessary to know the value of n , *i.e.* it requires prior knowledge of L to within $\pm 1/4$ of a fringe or approximately 150 nm. This accuracy cannot be achieved with conventional techniques.

To overcome this, use can be made of a second wavelength using the method of exact fractions [4]. If the effective range S_1 of the single wavelength technique is $\lambda_1 / 2$, then the range of the two wavelength system, $S_{1,2}$, is given by

$$S_{1,2} = \frac{\lambda_1 \lambda_2}{2(\lambda_1 - \lambda_2)} \quad (6.4)$$

where λ_1 and λ_2 are the two wavelengths used [5,6,7]. The range $S_{1,2}$ of the two-wavelength technique can thus be increased by making $\lambda_1 - \lambda_2$ small, *i.e.* by using two similar wavelengths. For example, with $\lambda_1 = 633$ nm and $\lambda_2 = 543$ nm, as used in previous interferometers [8], $S_{1,2} = 1.9$ μm , *i.e.* an estimate of the length of the bar within ± 0.9 μm will allow unambiguous calculation of the accurate length of the bar. For long bars in particular, such as those over 1 m in length, this accuracy is difficult to achieve, especially without accurate knowledge of the thermal expansion coefficient of the length bar (see Chapter 8). It is possible to increase the range $S_{1,2}$ by using other wavelengths, *e.g.* $\lambda_1 = 633$ nm and $\lambda_2 = 612$ nm, for which $S_{1,2} = 9.2$ μm , however the effective range is actually smaller than $S_{1,2}$ because it is limited by the accuracy of the measurement of the fringe fractions f_1 and f_2 . This can be overcome by using a third wavelength, λ_3 . To see why a third wavelength is necessary, the system of solutions to equation (6.3) for two wavelengths will be examined.

Rewriting equation (6.3) for the two ambient wavelengths λ'_1 and λ'_2 gives

$$L_1 = (n_1 + f_1)\lambda'_1 / 2 \quad (6.5a)$$

$$L_2 = (n_2 + f_2)\lambda'_2 / 2 \quad (6.5b)$$

Values of f_1 and f_2 are measured in the interferometer. With no *a priori* knowledge of L , values of n_1 and n_2 are undetermined and solutions of (6.5a) and (6.5b) are periodic in $\lambda'_1 / 2$ & $\lambda'_2 / 2$, respectively (see figure 6.11). For certain values of n_1 and n_2 the solutions of (6.5a) and (6.5b) are equal to within a small margin of error. In figure 6.11 this occurs for $(n_1 = 1, n_2 = 1)$, $(n_1 = 6, n_2 = 7)$, $(n_1 = 12, n_2 = 14)$ and $(n_1 = 17, n_2 = 20)$. Only one of these solutions corresponds to the length of the bar. The correct solution is deemed to be the one for which the two individual solutions agree most closely, in this case either $(n_1 = 1, n_2 = 1)$ or $(n_1 = 17, n_2 = 20)$ would be selected, *i.e.* the effective range of the technique is limited to 17 orders of $\lambda'_1 / 2$ (in this example), unless these two close matches can be resolved.

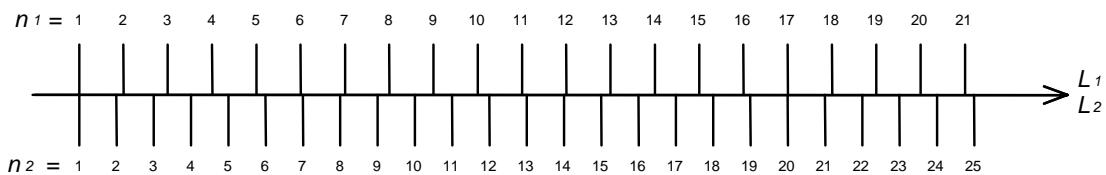


Figure 6.11 - Coincidences for two wavelengths 633 nm & 543 nm

To distinguish between the close matches requires a measurement resolution which depends on the accuracy with which the values of $(n + f)\lambda'/2$ can be measured. In the interferometer, the fundamental limit on the accuracy of measurement is the knowledge of the ambient wavelength λ' caused by uncertainty in the determination of the refractive index of the ambient air at different wavelengths (see Chapter 7). The uncertainty in this dispersion correction is approximately $\pm 2.48 \times 10^{-8}$. This is equal to a measurement uncertainty of ± 25 nm for a 1000 mm bar, or 0.08 fringes at wavelength 633 nm. Thus any solutions which agree to closer than $0.08\lambda_1'/2$ will not be resolved. In the case of $\lambda_1 = 633$ nm and $\lambda_2 = 612$ nm, the corresponding solutions occur for $(n_1 = 1, n_2 = 1)$ and $(n_1 = 2, n_2 = 2)$ because the two wavelengths are so similar. Thus using these 2 wavelengths the technique is limited to only 1 order of $\lambda_1'/2$, or 306 nm which is no better than using single wavelength interferometry.

To distinguish between these solutions, a third 543 nm wavelength is used. This leads to a set of 3 equations of the form of (6.5a) and (6.5b) with three periodic sets of solutions. The correct solution is identified by close matches at all three wavelengths with the coincidence at $n_1 = 17$ resolved, as shown in figure 6.12. In the interferometer, with $\lambda_1 = 633$ nm, $\lambda_2 = 543$ nm and $\lambda_3 = 612$ nm, the effective range is extended to 30 orders of $\lambda_1'/2$, or approximately 9.5 μm , which is where the next close match between all three solutions occurs. It is relatively easy to measure the length of a bar to within ± 9 μm by comparison with other measured bars e.g. using a CMM. Also, the length tolerances for length bars according to BS 5317 mean that any bar should be well within ± 9 μm of its nominal length, e.g. 1000 mm ± 1.55 μm is a typical tolerance.

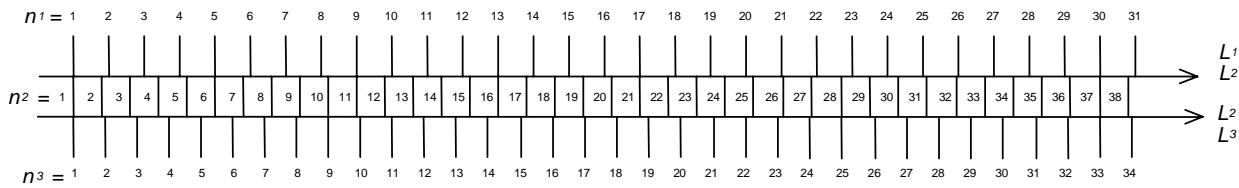


Figure 6.12 - Coincidences for three wavelengths 633 nm, 543 nm, 612 nm

Thus to find the length of the bar, based on a nominal value and three measured fringe fractions, the system of coincidences between the three wavelengths is examined for the closest match over a range of ± 15 orders of the red wavelength, centred on the order corresponding to the nominal size of the bar input by the user. This produces one solution which corresponds to the length of the bar.

This system of matching solutions for functions of different periods is similar to the integer analysis proposed by Guzhov & Solodkin [9], however the technique of integer decoding of interferograms is not valid in this particular situation as the uncertainty associated with the measurement of the fringe fractions precludes the use of integers to describe them, *i.e.* the values are real numbers with associated uncertainties.

The accuracy of the matching of the 3 solutions is measured by examining the ‘residuals’ - these are calculated as follows.

Let L_r , L_g , L_o be the lengths of the bar calculated using the 3 measured fringe fractions (red, green and orange respectively). The length of the bar can be calculated as the mean of these 3 results:

$$L = \frac{L_r, L_g, L_o}{3} \quad (6.6)$$

The red, green and orange residuals r_r , r_g , r_o are calculated as follows

$$r_r = \frac{2(L - L_r)}{\lambda'_r} \quad r_g = \frac{2(L - L_g)}{\lambda'_g} \quad r_o = \frac{2(L - L_o)}{\lambda'_o} \quad (6.7)$$

i.e. they represent the departure from perfect agreement, in units of one fringe (at each wavelength). Typical values of the residuals are dependent on the length being measured (showing the uncertainty in measurement to be length dependent and hence

due to factors such as alignment, refractive index, *etc*): for a short bar (up to 300 mm) they are typically less than about ± 0.01 fringe (± 3 nm) and for long bars (approximately 1000 mm), they are ± 0.03 fringe (± 9 nm). Note that some of the residuals will be positive, some negative. When two of the residuals have one sign for all measurements, and the third has the opposite sign, then that indicates that the single laser has drifted out of calibration more than the mean of the two other lasers. If the values of all the residuals increases with time, but with randomly varying signs, then this indicates that the equipment used for refractive index determination or temperature measurement is due for recalibrating (*e.g.* Druck, Michell, Tinsley, PRTs). In the interferometer, the frequency stability of the red laser is better than the green and orange lasers, so the measured length is the red measured length, rather than the average of the three measured lengths.

6.4.2 Limit to multiple-wavelength technique due to source instability

Walsh [10] has proposed a technique for calculating the limit to the length which can be measured using multiple-wavelength interferometry due to frequency instability of the light source(s). This will be examined for the case of the interferometer presented here.

Firstly, the equations for the length measurement are re-written as follows:

The equation

$$2L = \lambda_i(N_i + f_i) \quad (6.8)$$

is written as

$$2L = \lambda_s(N_s + f_s) \quad (6.9)$$

where

$$\lambda_s = \frac{1}{\sum A_i \sigma_i} \quad \sigma_i = \frac{1}{\lambda_i} \quad N_s = \sum A_i N_i \quad f_s = \sum A_i f_i \quad (6.10)$$

the f_i are fringe fractions, the N_i are the integer interference orders and the A_i are coefficients, chosen such that

$$\lambda_s \gg \lambda_i \quad \text{and} \quad \sum A_i = 0 \quad (6.11)$$

(Note that the effective wavelength is given by λ_s , which for the two-wavelength case, is the same as that given by (6.4)). The length of the object can then be calculated from

$$2L = \frac{1}{\sum A_i \sigma_i} \left(\sum A_i N_i + \sum A_i f_i \right) \quad (6.12)$$

provided that the length of the object is already known to within $\pm \lambda_s/2$.

Walsh shows that the maximum length that can be measured using this technique is given by

$$L \leq \frac{1}{2} \frac{c}{\sum |A_i| \Delta v} \quad (6.13)$$

where Δv is the uncertainty in the frequency of the light source. For the interferometer, the largest value of Δv is approximately 5 MHz (for the green laser). The choice of values for the A_i coefficients depends on the operational mode. For two-wavelength interferometry, $A_1 = 1$ and $A_2 = -1$, leading to a value for $\sum |A_i|$ of 2. The maximum length measurable with this technique using the green and red wavelengths is 15 m, provided that the length is already known to within $\pm 1.9 \mu\text{m}$ (as before). For three-wavelength interferometry, $A_1 = -1$, $A_2 = 2$, $A_3 = -1$, giving $\sum |A_i| = 4$. The maximum length measurable with the three-wavelength technique using the three wavelengths of the interferometer is 7.5 m, provided the length is already known to within $\pm 3.2 \mu\text{m}$.

Using the above equations for the calculation of the length of the object shows that the frequency instability of the green laser places a tighter tolerance on the initial length estimate, unless the effect of the laser drift can be decreased, for example by more frequent calibration of the lasers. The interferometer does not use equation (6.12) to calculate the length of the bar directly because the N_i are not directly known: the method of excess fractions is used. These two techniques are fundamentally the same and the limit of 7.5 m does indeed apply to the interferometer and represents the maximum length which can be measured using this technique with these laser wavelengths.

6.4.3 Multiple wavelength algorithm - method of exact fractions

The procedure from the computer program which calculates the result is presented here. Comments on the code are in bold type.

```

procedure calc_length;

{procedure to calculate multiple-wavelength solution to measured length}
{works with two or three wavelengths}

var
  alter,range           :integer;
  i,x,y,best_solution,again   :integer;
  amb_order_red          :integer4;
  amb_order_green         :integer4;
  amb_order_orange        :integer4;
  amb_red_est             :real8;
  amb_green_est           :real8;
  amb_orange_est          :real8;
  nom_red_est             :real8;
  nom_green_est           :real8;
  nom_orange_est          :real8;
  red_residual            :real8;
  green_residual           :real8;
  orange_residual          :real8;
  nom_mean_length,minimum :real8;
  wavelength              :real8;
  amb_length               :real8;
  amb_est                  :real8;
  residuals                :array[1..3,-20..20]of real8;
  key                     :byte;

begin

{resets length and wavelength variables}

precis_length:=0.0;
red_wave_corr:=9.99;
green_wave_corr:=9.99;
orange_wave_corr:=9.99;

{calculates wavelengths of operational lasers to ambient conditions}

if (red_used) then
  red_wave_corr:=wave_correction(red_wavelength,air_temp,air_pressure,air_humidity);
if (green_used) then
  green_wave_corr:=wave_correction(green_wavelength,air_temp,air_pressure,air_humidity);
if (orange_used) then
  orange_wave_corr:=wave_correction(orange_wavelength,air_temp,air_pressure,air_humidity);

{converts nominal length to ambient conditions}
{estimates nominal red interference order at ambient conditions}

amb_length:=nom_to_amb(nom_length,bar_temp);
amb_order_red:=round4(amb_length*2/red_wave_corr);

{resets fringe fraction totals for before averaging 9x9 points at image
centre}

red_fraction:=0;
green_fraction:=0;
orange_fraction:=0;
for x:=cent_x-4 to cent_x+4 do
  begin
    for y:=cent_y-4 to cent_y+4 do
      begin
        if (red_used) then red_fraction:=red_fraction+phi_red[x,y];
        if (green_used) then green_fraction:=green_fraction+phi_green[x,y];
        if (orange_used) then
          orange_fraction:=orange_fraction+phi_orange[x,y];
      end;
    end;
  red_fraction:=red_fraction/81;
  green_fraction:=green_fraction/81;
  orange_fraction:=orange_fraction/81;

{displays some headings and other information on the screen}
{chooses order scanning range: 20 for normal, 50 if debugging}

if (debugging) then rewrite(debug_file);
if (debugging)
  then range:=50
  else range:=20;

```

```

{scans over range of orders, around ambient estimated order}
{calculates solutions to interference equation at each order}
{and converts back to nominal conditions}

for i:=-range to range do
begin
amb_red_est:=(i+amb_order_red+red_fraction)*red_wave_corr/2;
nom_red_est:=amb_to_nom(amb_red_est,bar_temp);

{if green wavelength operating, works out corresponding green order at}
{ambient, and then solves interference equation for green data}
{then converts to nominal conditions}

if (green_used) then
begin
amb_order_green:=round4(2*amb_red_est/green_wave_corr-green_fraction);
amb_green_est:=(amb_order_green+green_fraction)*green_wave_corr/2;
nom_green_est:=amb_to_nom(amb_green_est,bar_temp);
end;

{repeats this process if orange wavelength operating}

if (orange_used) then
begin
amb_order_orange:=round4(2*amb_red_est/orange_wave_corr-
orange_fraction);

amb_orange_est:=(amb_order_orange+orange_fraction)*orange_wave_corr/2;
nom_orange_est:=amb_to_nom(amb_orange_est,bar_temp);
end;

{declares red answer correct, calculates orange and green residuals}
{which are departures from the red answer, in units of fringes}

nom_mean_length:=nom_red_est;
amb_length:=nom_to_amb(nom_mean_length,bar_temp);

if (green_used) then
begin
amb_est:=nom_to_amb(nom_green_est,bar_temp);
green_residual:=(amb_est-amb_length)/green_wave_corr;
end;

if (orange_used) then
begin
amb_est:=nom_to_amb(nom_orange_est,bar_temp);
orange_residual:=(amb_est-amb_length)/orange_wave_corr;
end;

{resets residuals to zero if wavelength not used}

if not(green_used) then green_residual:=0.0;
if not(orange_used) then orange_residual:=0.0;
red_residual:=0.0;

{writes debugging information to file, if requested}

if (debugging) then
begin
write(debug_file,i+amb_order_red,chr(9),green_residual:5:3,chr(9),
orange_residual:5:3,chr(9));
writeln(debug_file,abs(green_residual) + abs(orange_residual));
end;

{for central +- 20 orders, stores residuals in an array}

if (abs(i) <=20) then
begin
residuals[1,i]:=red_residual;
residuals[2,i]:=green_residual;
residuals[3,i]:=orange_residual;
end;
end;

{scans through array, picking lowest absolute residuals}
{marks this as best solution}

minimum:=99999.9;

for i:=-20 to 20 do
begin
if ((abs(residuals[1,i]) + abs(residuals[2,i]) + abs(residuals[3,i])) <
minimum)
then

```

```

begin
  minimum:=abs(residuals[1,i]) + abs(residuals[2,i]) +
  abs(residuals[3,i]);
  best_solution:=i;
end;
end;

{for 3 orders either side of best solution, displays results}

for i:=best_solution-3 to best_solution+3 do
begin
  amb_red_est:=(i+amb_order_red+red_fraction)*red_wave_corr/2;
  nom_red_est:=amb_to_nom(amb_red_est,bar_temp);

  if (green_used) then
    begin
      amb_order_green:=round4(2*amb_red_est/green_wave_corr-
green_fraction);
      amb_green_est:=(amb_order_green+green_fraction)*green_wave_corr/2;
      nom_green_est:=amb_to_nom(amb_green_est,bar_temp);
    end;

  if (orange_used) then
    begin
      amb_order_orange:=round4(2*amb_red_est/orange_wave_corr-
orange_fraction);

      amb_orange_est:=(amb_order_orange+orange_fraction)*orange_wave_corr/2;
      nom_orange_est:=amb_to_nom(amb_orange_est,bar_temp);
    end;

  nom_mean_length:=nom_red_est;
  amb_length:=nom_to_amb(nom_mean_length,bar_temp);

  if (green_used) then
    begin
      amb_est:=nom_to_amb(nom_green_est,bar_temp);
      green_residual:=(amb_est-amb_length)/green_wave_corr;
    end;

  if (orange_used) then
    begin
      amb_est:=nom_to_amb(nom_orange_est,bar_temp);
      orange_residual:=(amb_est-amb_length)/orange_wave_corr;
    end;

  if not(green_used) then green_residual:=0;
  if not(orange_used) then orange_residual:=0;
  red_residual:=0.0;

  residuals[1,i]:=red_residual;
  residuals[2,i]:=green_residual;
  residuals[3,i]:=orange_residual;

{prints information, and flags best solution}

write((i+amb_order_red):8,red_residual:10:3,green_residual:10:3,orange_resi
dual:10:3,
' ,amb_length*1000:11:6);
if (i = best_solution)
then
begin
  write('  <-- BEST SOLUTION');
  precis_length:=nom_mean_length;
  uncorrected_length:=nom_to_amb(precis_length,bar_temp);
  best_red_residual:=residuals[1,i];
  best_green_residual:=residuals[2,i];
  best_orange_residual:=residuals[3,i];
end;
writeln;
end;
end;

```

6.4.4 Order scanning

Although the expected repeat distance for the three-wavelength technique is 30 orders (± 15 either side of nominal), it is often possible to extend this range, with care. At the coincidence at $n_I = 30$, the difference between the red and green coincidences is -0.061 green fringes and between the red and orange coincidences is 0.031 orange fringes.

$$\begin{aligned} & 30.000 \times \lambda_I \\ &= 31.031 \times \lambda_2 \\ &= 34.939 \times \lambda_3 \end{aligned}$$

It is shown in chapter 7 and chapter 10 that the uncertainty of the refractive index correction is $\pm 2.5 \times 10^{-8}$ at a confidence level of 95%. Thus it is 95% certain that the errors in absolute fringe fraction measurements will be within $\pm 0.08 \lambda_I$, $\pm 0.09 \lambda_2$, $\pm 0.08 \lambda_3$ for bars up to 1 m in length. For shorter bars the errors will be expected to be correspondingly smaller and so it is possible to extend the order scanning. It has been found that ± 20 orders is a suitable scanning range though the software allows this to be extended to ± 50 orders for debugging purposes.

6.5 FLATNESS AND PARALLELISM MEASUREMENTS

As well as measuring the central length of the bar, the interferometer also measures the flatness of the exposed face of the bar and the parallelism of the faces, measured as the variation in length of the bar measured at different points on the exposed face. Because these are only relative measurements (from one point to another) and do not require absolute determination, the values of flatness and parallelism are calculated from only the red wavelength phase map.

6.5.1 Measurement of parallelism (variation)

The parallelism (strictly the deviation from parallelism) is defined as the difference between the maximum and minimum lengths, measured at any points on the measurement faces, measured perpendicular to the wrung face (see § 1.4.5).

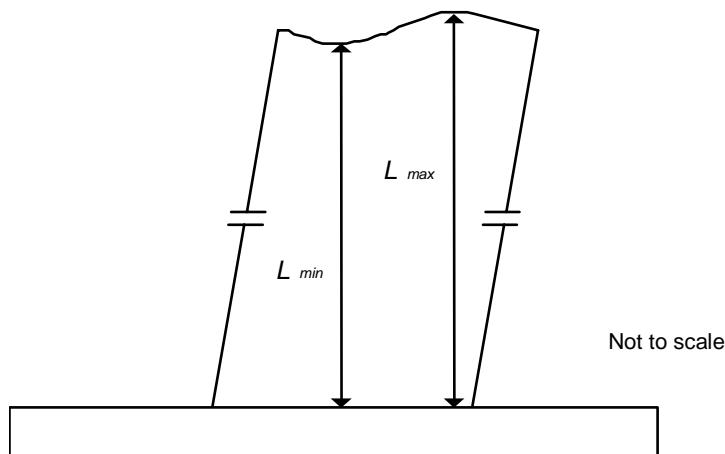


Figure 6.13 - Measurement of deviation from parallelism of a length bar's measurement faces

Note that the definitions of L_{max} and L_{min} are the same as for the central length, *i.e.* the distance, perpendicular to the wrung surface, to the measurement point. This is why the interferometer is aligned with the platen face perpendicular to the measurement beam, rather than the exposed face of the bar (see § 4.1.8). The phase measurements are also referred to the best fit surface through the platen data, which should correspond closely with the mean tangent to the platen surface at any point, if the platen surface is smoothly varying.

The parallelism is thus calculated as the difference between the maximum and minimum values in the phase data measured on the surface of the bar (the software mask is used to check that the data corresponds to the surface of the bar).

```
{having fitted least squares plane to data on the bar}
{now finds variation (parallelism)}

min:=9999.9;
max:=-9999.9;
for x:=0 to 126 do
begin
  for y:=0 to 126 do
    begin

{checks to see if data corresponds to bar surface}

      if (imagestore(11,x,y)=254) then
        begin

{data is on bar, takes max and min values}

        if (phi_red[x,y] < min) then min:=phi_red[x,y];
        if (phi_red[x,y] > max) then max:=phi_red[x,y];
      end;
    end;
end;

variation:=max-min;
```

When measuring gauge blocks, this result is termed the *variation* in length, since the parallelism is defined differently according to which standard is being used. Thus in the results, the parallelism is also referred to as variation.

6.5.2 Measurement of flatness

The (deviation from) flatness is defined as the minimum distance between two parallel planes which just envelop the measuring face.

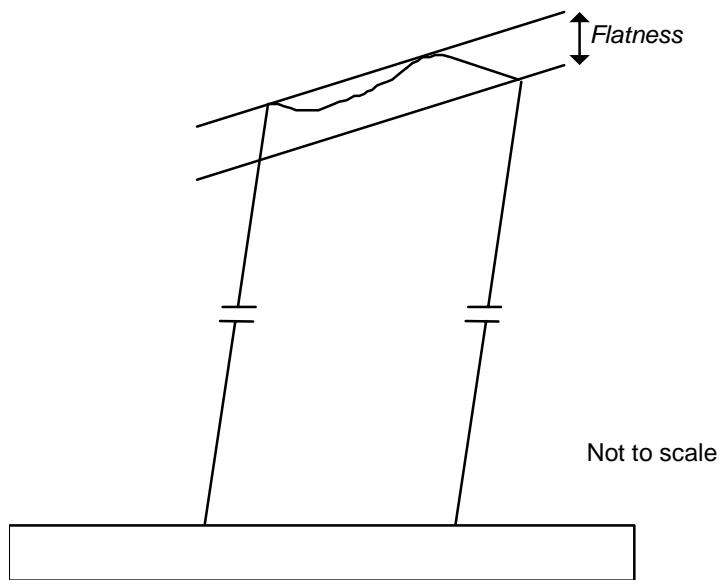


Figure 6.14 - Measurement of deviation from flatness of length bar's measurement face

This definition is difficult to realise in practice due to the difficulty of fitting two parallel planes to the data. Except for pathological cases, this definition of flatness is equivalent to that used in the interferometer, which is the difference between the maximum positive and negative departures from the best fit plane through the surface of the bar.

To measure the flatness, the program performs a least squares fit of a plane to the data on the bar. The difference between the maximum positive and negative departures from this plane is the flatness.

```

{fits least squares plane to data on bar}
{now finds flatness}

min:=9999.9;
max:=-9999.9;
for x:=0 to 126 do
begin
  for y:=0 to 126 do
    begin

{checks to see if data corresponds to bar surface}

      if (imagestore(11,x,y)=254) then
        begin

{data is on bar, now take difference between fitted and real data}
{max-min = flatness}

        if ((phi_red[x,y]-(b0+(b1*x)+(b2*y))) < min) then
min:=(phi_red[x,y]-(b0+(b1*x)+(b2*y)));
        if ((phi_red[x,y]-(b0+(b1*x)+(b2*y))) > max) then
max:=(phi_red[x,y]-(b0+(b1*x)+(b2*y)));
        end;
      end;
    end;
  flatness:=max-min;

```

6.5.3 Example measurements

The following are some example measurements of flatness and parallelism performed on length bars in the interferometer. For each bar, the red phase map is displayed, with numerical data for the flatness and parallelism (variation).

These measurements were performed with bars from the same set (Set 1455). In these printouts the end of the bar appears slightly oval, this is due to the camera and framestore pixels corresponding to rectangles in the image, rather than squares. A correction factor is used in the display software to try to correct the images.

The area around the bar which is masked and is set to zero phase can be seen in the phase maps as the flat area immediately around the edge of the bar. The smoothness of the platen surface can be judged from the phase data in the remainder of the image.

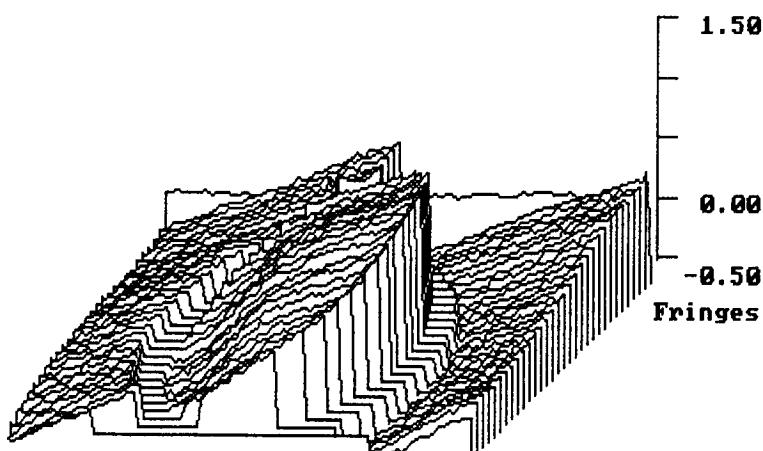
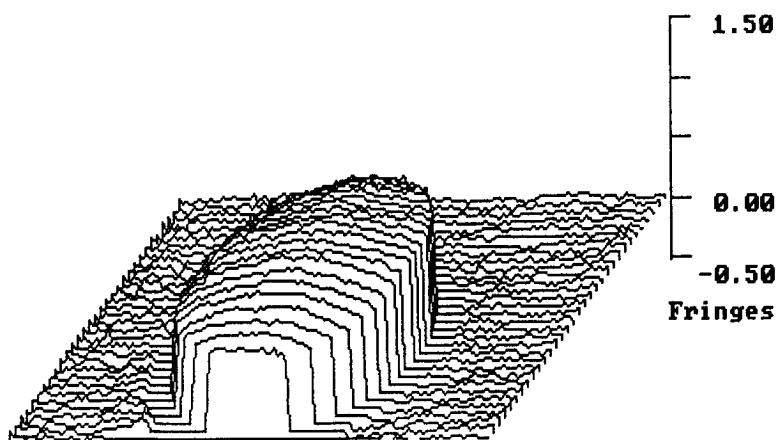
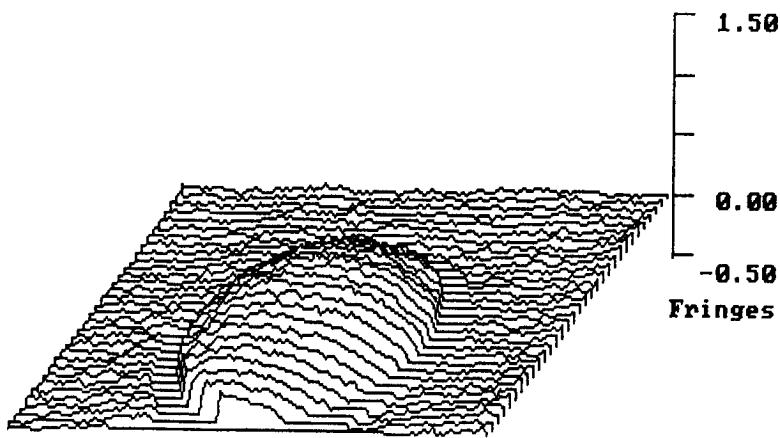


Figure 6.15 - Example measurement showing errors in the phase map due to improper discontinuity removal across top line, caused by incorrect positioning of platen in image

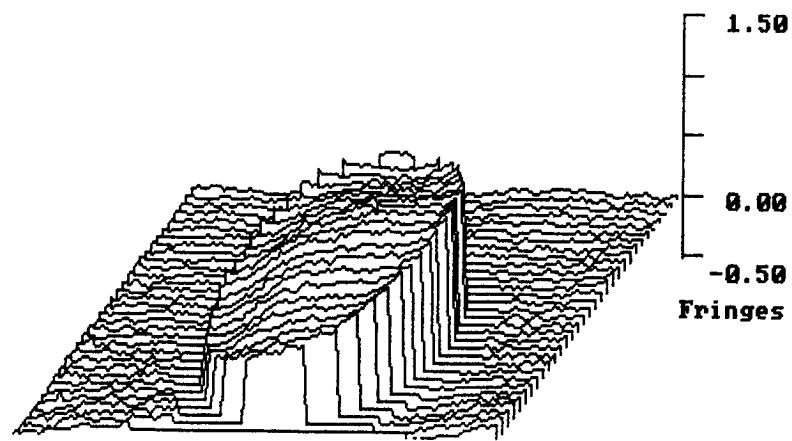
Flatness: 0.43 fringes
Variation: 0.48 fringes



Flatness: 0.55 fringes
Variation: 0.61 fringes



Flatness: 0.33 fringes
Variation: 0.74 fringes



Flatness: 0.42 fringes
Variation: 1.11 fringes

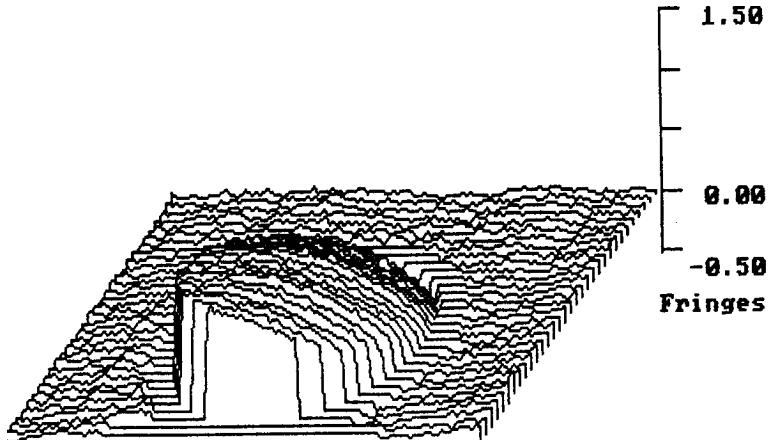


Figure 6.16 - Example measurements of flatness and parallelism (variation) of length bars

6.6 COMPUTER PROGRAM

The computer program is approximately 4500 lines long so only an overview of the routines is given here for reference. Figure 6.17 flowcharts the program which controls the interferometer.

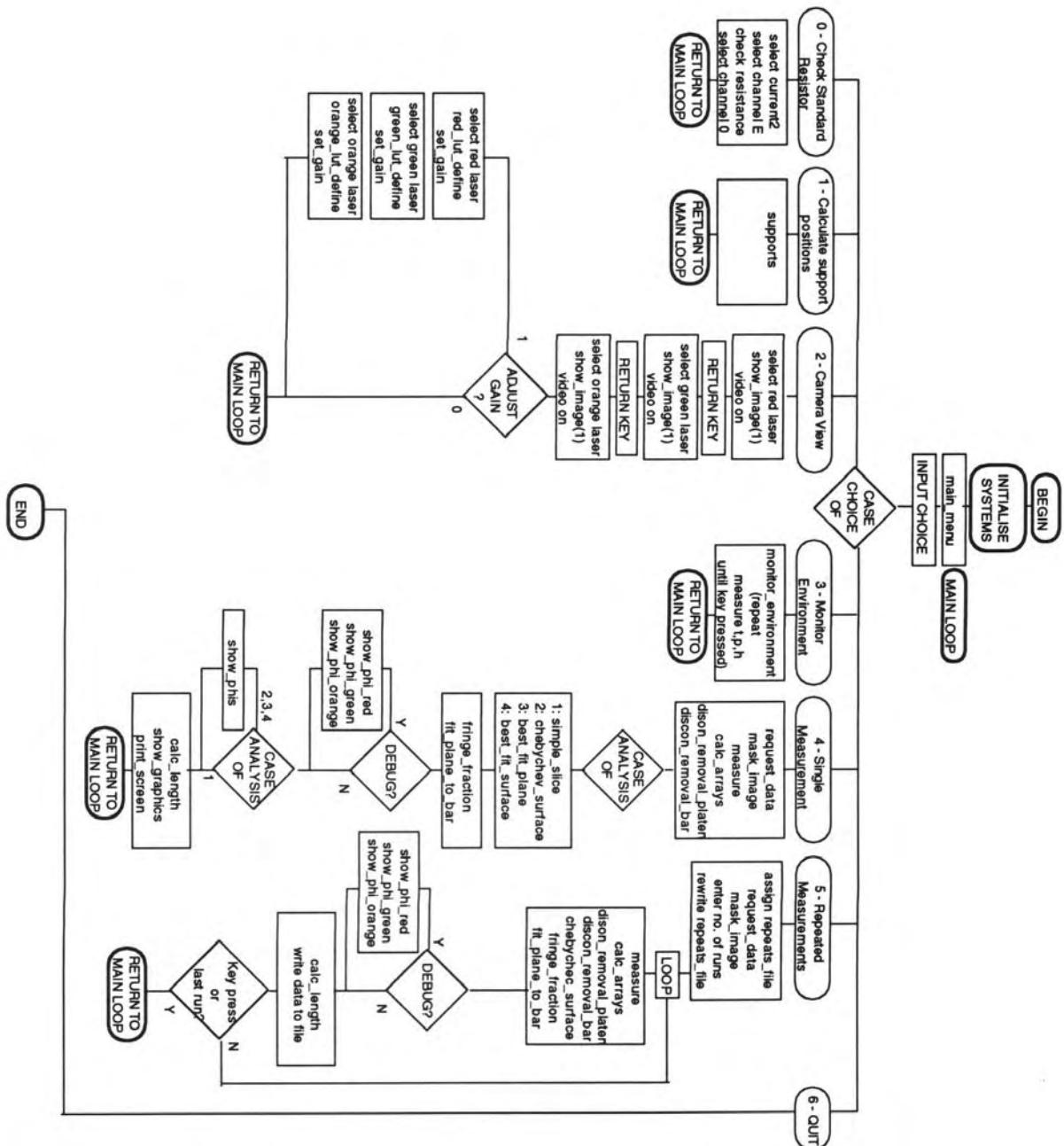


Figure 6.17 - Program flowchart (simplified)

After initialisation, the program is based around a main loop which allows the user six choices, selected from a menu. These choices are as follows.

0 - Check standard resistor

This option is used periodically to calibrate the Tinsley resistance bridge. It selects channel E, which has the $100\ \Omega$ resistor connected. The display of the bridge then shows the value of the resistor. The front panel potentiometer is then adjusted until the resistance reading is the same as the calibrated value for the resistor.

1 - Calculate support positions

This option is used to calculate the position of the supports for the length bars. An input of the bar length is required. The calculation assumes a platen thickness and diameter corresponding to the six specially manufactured interferometer platens. The calculation uses the analysis developed in appendix C. The results given are the positions of the supports from either end of the bar and the support separation.

2 - Camera view

This option is used either when aligning the interferometer (to switch on the laser beam) or when adjusting the tilt of the mirrors. The lasers are selected in sequence: red, green, orange. If necessary, there is a further option which allows the user to adjust the offset and gain of the digitiser to ensure that the full range of the A-to-D converter is used, but without saturation. A histogram is shown on screen similar to figures 6.4 and 6.5.

3 - Monitor environment

This option continuously monitors conditions inside the chamber in a loop, until the user selects halt. The temperatures are measured in sequence with the pressure, humidity and CO_2 concentration. The sample pump is halted before the pressure reading is taken. The results can be saved to a file on disc.

4 - Single measurement

This option is the one normally used for making measurements of length bars. Firstly it requests data from the user for: Reference, Nominal size of bar, nominal thermal expansion coefficient, which of the three positions on the carriage the bar is occupying

(to select the correct PRTs), and the type of analysis to use. The program then allows the user to adjust the mask using a video overlay system which shows the image of the bar on screen with superimposed cursor lines which the user adjusts to enclose the edge of the bar in the image. Next the full measurement procedure begins with selection of the relevant PRTs and measurement of the humidity and CO₂ content. The sample pump is then stopped and the PRT resistance measured. The air PRT is then read. The phase-stepping is then performed, including synchronisation to the modulation of the red laser and a suitable wait period between successive steps to allow the DPT to reach correct position. In the middle of the phase-stepping, the pressure is read. The second of the PRTs in contact with the bar is measured and the sample pump re-started.

Next the program calculates the `phi` and `alpha` arrays (phase and phase step size, respectively) for the three wavelengths. If the debugging mode is active, the three phase step maps are displayed on the framestore monitor as grey level maps with histograms showing the spread in phase step. The discontinuities are removed from the platen and bar surfaces and the appropriate analysis is used to fit to the data on the platen. The fringe fractions are then calculated as the difference between the measured data and the fitted data, divided by 360°. A least-squares plane is fitted to the phase data of the bar surface to allow calculation of the flatness and parallelism. If the debugging mode is active, the three phase maps are displayed on the framestore monitor as grey level maps. The three phase maps are displayed on the monitor in pseudo-three dimensional form (*e.g.* figure 6.16). Next the multiple-wavelength analysis is used to calculate the length of the bar. The results are displayed with an option for printout.

5 - Repeated measurements

This option performs the same as option 4, but with no graphical displays and with results saved to disc file. The user can select the number of measurement runs, which can be interrupted if necessary at the end of any measurement.

6 - Quit

This option simply terminates the main loop, closes files and quits the program.

Figure 6.18 shows an example printout from the interferometer for a 1000 mm length bar measured at approximately 20 °C. The bar temperature is 20.008 °C, the air temperature is 20.012 °C. The bar is flat to within 0.46 fringes (145 nm) and exhibits a

variation in length (parallelism error) of 1.16 fringes (367 nm). The measured length of the bar is 1000.003 601 mm, hence it shows a departure of 3601 nm from its nominal length. The results show a slight rounding error between the two stated values for departure of 1 nm.

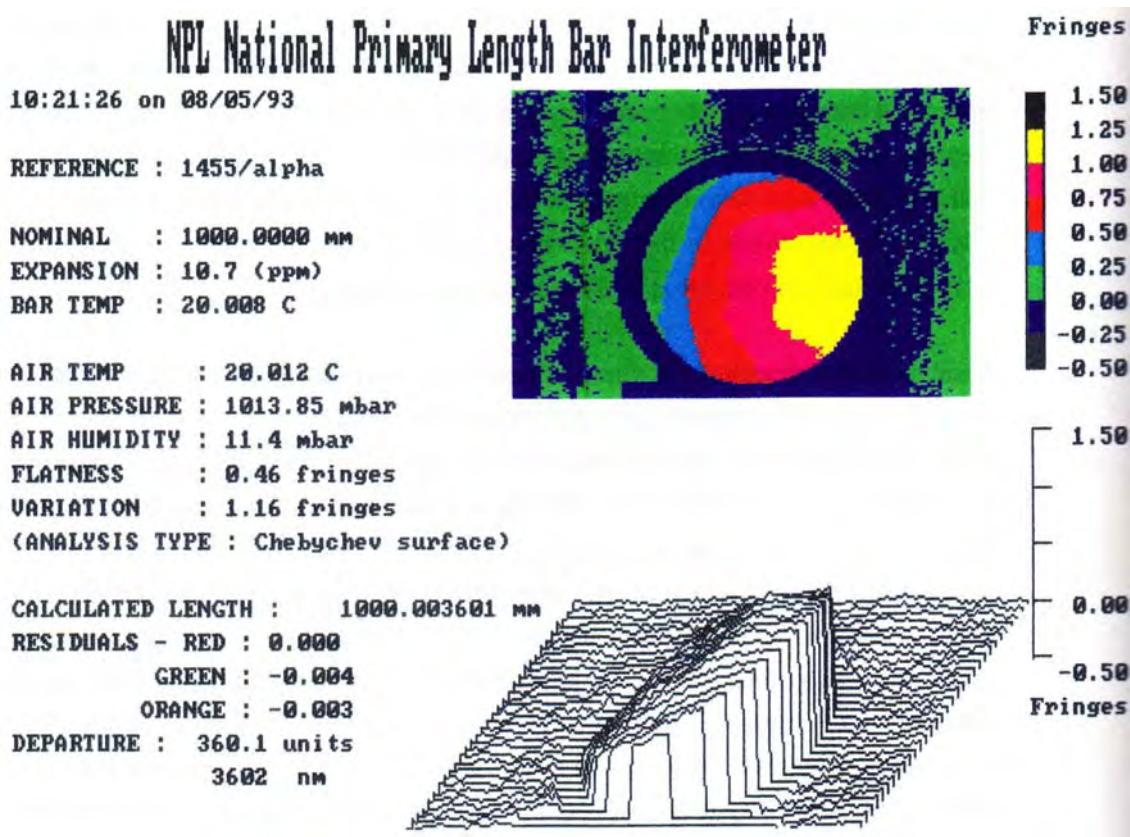


Figure 6.18 - Example printout from interferometer: results for a 1000 mm bar

6.7 DOUBLE-ENDED ANALYSIS

Initially the double-ended analysis was tried as a part of the normal interferometer program but this increased the code size beyond the limit imposed by the compiler - 64 K bytes of user generated code per compiled unit, module or program. Splitting the program into separate completed modules which could be combined at run time was tried, but cross-module global variable compatibility was difficult to achieve. The final solution was to produce a new program based on the original interferometer control program, but with non-essential procedures and functions removed to make space for the double-ended procedures. The new program has no options for environmental monitoring, disc file usage or single-ended measurements. It simply allows a camera view (for alignment purposes) and double-ended phase-stepping. New routines were

developed for: double-ended masking, double-ended discontinuity removal, double-ended data fitting and double-ended multiple-wavelength analysis. The main functionality of the program is the same as option 4 of the previous single-ended program, except for the following.

- 1 Masking now has to mask 2 bar images, a central discontinuity or join, and a larger area of masked-out data beneath the bar images.
- 2 When phase-stepping, it is necessary to re-align the interferometer for each wavelength because otherwise the extra tilt due to dispersion in the beamsplitter would defeat the discontinuity-removal algorithms (see figure 4.23).
- 3 The discontinuity removal for the bar now has to work on 2 separate bar faces.
- 4 The platen discontinuity removal has to 'bridge' the join in the centre of the image - it maps the pixels to the immediate left and right of the join to the same phase modulo 2π and then uses linear interpolation to fill the gap between them.
- 5 The 3-pass routine now needs a fourth pass in the lower-centre of the image to unwrap data in between the bars, near the supports.
- 6 Two fringe fractions are measured after using a least-squares-plane analysis - there is insufficient data to use multiple chebychev fits.
- 7 The multiple-wavelength calculation routine now uses double-ended fringe fraction results (see § 3.3.4).

The lower contrast of the fringes in the background of the image (previously the platen area) and the necessity to use best-fit-plane analysis result in larger phase-measurement errors in this region. This means that the fringe fraction measurements are not as accurate as the single-ended measurements and the reference surface (background) phase data is not flat. Measurements made using double-ended interferometry will not be as accurate as single-ended measurements and the scanning range of the multiple-wavelength analysis is restricted to prevent spurious results being selected. Results of some double-ended measurements are given in chapter 9.

The phase-stepping and refractive index and temperature measurements are performed as before. The spread in `alpha` (phase-step angle) appears larger than for single-ended measurements due to the low fringe contrast producing errors in the phase-step size calculation, but the mean step sizes for the three wavelengths are unchanged. This is

because it is the reference mirror which is stepped so the phase-stepping algorithm is the same as used for the single-ended measurements with the same step size. The fringe fractions are calculated as before, except that the signs are reversed: the fitted data of the background is subtracted from the measured phase maps then the phase values at the centres of the bar images are summed to give three complementary fringe fractions, *i.e.* $1-f_r$, $1-f_g$ and $1-f_o$. This is because equation (3.17) requires fringe fractions of the opposite sign to those measured in single-ended interferometry.

REFERENCES FOR CHAPTER 6

- [1] Ghiglia D C, Mastin G A & Romero L A Cellular automata method for phase-unwrapping *J. Opt. Soc. Am.* **4** (1987) 267-280
- [2] Stetson K A Phase-step interferometry of irregular shapes by using an edge-following algorithm *Appl. Opt.* **31** (1992) 5320-5325
- [3] Gierloff J J Phase unwrapping by regions *SPIE* **818** (1987) 2-9
- [4] Rolt F H *Gauges and Fine Measurement* (MacMillan & Co.: London) (1929) 49-50, 212-213
- [5] Creath K Step height measurement using two-wavelength phase-shifting interferometry *Appl. Opt.* **26** (1987) 2810-2816
- [6] Cheng YY & Wyant J C Two wavelength phase shifting interferometry *Appl. Opt.* **23** (1984) 4539-4543
- [7] Cheng YY & Wyant J C Multiple wavelength phase-shifting interferometry *Appl. Opt.* **24** (1985) 804-807
- [8] Pugh D J & Jackson K Automatic gauge block measurement using multiple wavelength interferometry *Proc. SPIE* **656** (1986) 244-249
- [9] Guzhov V I & Solodkin Y N Using the properties of integers to decode interferograms *Opt. Spectrosc. (USSR)* **65** (1988) 661-665
- [10] Walsh C J Limit to multiwavelength interferometry imposed by frequency instability of the source *Appl. Opt.* **26** (1987) 29-31

CHAPTER 7

REFRACTOMETRY

“Eigentlich weiss man nur wenn man wenig weiss, mit dem Wissen wächst der Zweifel.”

(“We know accurately only when we know little; with knowledge doubt increases.”)
Goethe

7.1 REFRACTIVITY OF AIR

When performing length interferometry in air, it is important to correct the laser wavelength for the refractivity of the air through which it passes. The correction factor, the refractive index, is applied to the vacuum wavelength of the light emitted by the laser

$$\lambda = \frac{\lambda_{vac}}{n} \quad (7.1)$$

where λ is the wavelength in air, λ_{vac} is the wavelength in vacuum, and n is the refractive index of air, for the ambient conditions.

The refractivity [1] of a gas, $(n-1)$, can be expressed as the product of 2 factors: the dispersion factor K_λ and the density factor D_{TP} . The dispersion factor depends only on the wavelength and the density factor is independent of the wavelength. Hence

$$(n - 1)_{TP} = K_\lambda D_{TP} \quad (7.2)$$

7.1.1 Dispersion factor

Svensson [2] demonstrated the invariance of the dispersion factor with temperature and Erickson [3] demonstrated its invariance with pressure. Due to this invariance, the dispersion factor can easily be determined to high accuracy by relative measurement, and in 1953 the dispersion formula was published by Edlén [4], based on the work of other authors' investigations.

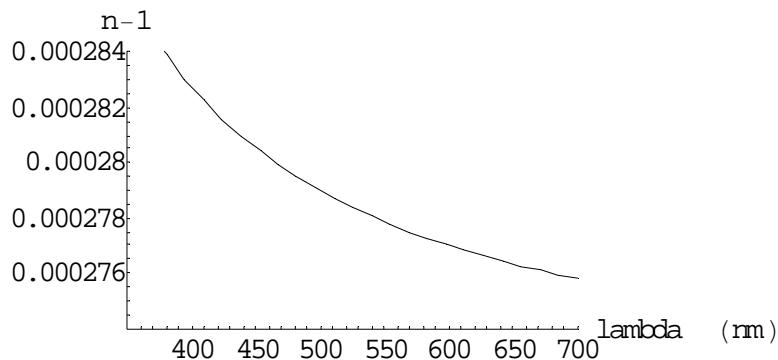


Figure 7.1 - Dispersion of standard air over the region 350 nm - 700 nm

This formula for the dispersion factor was expressed in the Sellmeier form

$$(n - 1) = \sum_i A_i \frac{1}{\sigma_i^2 - \sigma^2} \quad (7.3)$$

The σ_i represent wave-numbers of certain resonance frequencies, σ is the experimental wave-number. This is the basic form of the more recent dispersion formula (see below) that is now used. The Sellmeier formula is based on the idea that the refractivity of air is due to discrete absorption frequencies. This can be modelled by using Maxwell's equations in a dielectric medium in which an electric field interacts with an idealised harmonic oscillator, with various resonant frequencies. This is, however, not quite correct, as these absorption bands are not discrete frequencies, but continua. The term with the constant 38.9 in the formula below represents the effect of the absorption continuum of O₂, peaking at 145 nm, with a range of 40 nm. The constants 130 and 8342.13 represent the effect of the absorption continua of N₂ and O₂, starting at 100 nm, levelling off at 50 nm, before tailing off for lower wavelengths. Thus there is an empirical equation for K_λ , the dispersion factor, based on measurements of the dispersion of air.

7.1.2 Density factor

Edlén [5] gives a derivation for the density factor, D_{TP} based on the Lorentz-Lorentz equation, and the gaseous equation of state

$$\frac{PV}{RT} = 1 - \varepsilon_T P \quad (7.4)$$

and measured values of $\frac{PV}{RT}$ for dry air at different conditions. The form of the derived equation is

$$D_{TP} = \frac{P(1 + \beta_T P)}{(1 + \alpha T)} \quad (7.5)$$

This corresponds to equation (7.7) below relating the refractivity of air at non-standard pressures and temperatures, to that at standard conditions.

7.2 EDLÉN'S EQUATIONS FOR THE REFRACTIVITY OF AIR

In 1965 Edlén reviewed the most recent work [6,7], collated findings and issued new formulae for the dispersion of air [5]. The formulae derived in that paper have since been widely used to correct for the refractivity of air, with a minor correction to the humidity term suggested by Birch and Downs [8].

7.2.1 Refractivity of standard air

According to Edlén, standard air (dry air at 1 atmosphere, 15°C, containing 300 ppm by volume of CO₂) has a refractivity of

$$(n - 1)_s \times 10^8 = 8342.13 + \frac{2406030}{130 - \sigma^2} + \frac{15997}{38.9 - \sigma^2} \quad (7.6)$$

where n is the refractive index, σ is the vacuum wave-number in μm⁻¹. As an example, for air at 15 °C, $\lambda = 633$ nm, $(n-1)_s = 0.000\ 276\ 517$.

This is found to be accurate to about 1×10^{-8} over the wavelength range from 200 nm to the infra red, and represents the dispersion factor, described above.

7.2.2 Corrections for temperature, pressure, water vapour and CO₂

The effects of temperature and pressure are calculated from the density factor

$$(n - 1)_{TP} = (n - 1)_s \times \frac{P}{720.775} \left[\frac{1 + P(0.817 - 0.0133T) \times 10^{-6}}{1 + 0.0036610T} \right] \quad (7.7)$$

where P is the air pressure in torr, and T is the air temperature in °C. At 1 atmosphere (1013.25 mbar, 760 torr) and 20 °C, $(n-1)_{TP} = 0.000\,271\,786$.

A correction can also be applied for water vapour (humidity)

$$n_{TPf} = n_{TP} - f(5.722 - 0.0457\sigma^2) \times 10^{-8} \quad (7.8)$$

for air containing f torr of water vapour.

Birch and Downs [8] suggest an amendment to the last equation, whereby it becomes

$$n_{TPF} = n_{TP} - f(3.7209 - 0.0343\sigma^2) \times 10^{-8} \quad (7.9a)$$

but with f in mbar, not torr. When converted for measurements of f in torr, the wavelength-dependent term is unchanged from Edlén and the equation becomes

$$n_{TPF} = n_{TP} - f(4.9608 - 0.0457\sigma^2) \times 10^{-8} \quad (7.9b)$$

This is to correct an error in the water vapour term discovered using an absolute gas refractometer. Muijlwijk [9] suggested alterations to the standard formulae, on the grounds that normal laboratory conditions are now typically 20 °C, and 400 ppm by volume of CO₂. However, he acknowledged the fact that the equations that he presented agreed with Edlén's original equations to within 1 part in 10⁸, and thus for reasons of standardisation with other existing instruments, the refractivity equations used in the work of this thesis are those due to Edlén, 1966, with the Birch and Downs modification to the humidity term. The Birch and Downs revised humidity term has been verified by Beers and Doiron [10].

Birch and Downs expect the uncertainty of their modified Edlén equations to be $\pm 3 \times 10^{-8}$ unless more accurate measurements of the pressure, temperature and humidity are made. The amount of Carbon Dioxide must also not deviate from the standard value of 300 ppm by volume, unless its value is measured directly, as is the case with the NPLBI.

The effect of CO₂ concentration on the refractivity, according to Edlén, is given by

$$(n-1)_x = [1 + 0.540(x - 0.0003)](n-1)_s \quad (7.10)$$

where x is the CO₂ concentration, by volume. The effects of other gases commonly found in the atmosphere are negligible, unless specific contamination is introduced.

7.3 EFFECTS OF PRESSURE, TEMPERATURE, HUMIDITY AND CO₂ ON REFRACTIVITY

Table 7.1 summarises the effects of the air temperature, pressure, water vapour content (humidity) and CO₂ content on the refractive index.

Parameter	Typical value	Typical daily variation in calibration laboratory	Effect of variation on refractive index
temperature	20 ° C	± 0.2 °C	± 1.85 x 10 ⁻⁷
pressure	1013.25 mbar	± 20 mbar	± 5.36 x 10 ⁻⁶
humidity	13 mbar	± 2 mbar	± 5.45 x 10 ⁻⁸
CO ₂ content	300 ppm	± 100 ppm	± 1.47 x 10 ⁻⁸

Table 7.1 - Typical variations in air temperature, pressure, humidity and CO₂ content

Thus air pressure and temperature have the largest effect on the refractive index, hence accurate measurement of these parameters is essential if the refractive index is to be calculated accurately using Edlén's equations.

Parameter	Increment	Effect on refractive index
temperature	1 ° C	-9.26 x 10 ⁻⁷
pressure	1 mbar	2.68 x 10 ⁻⁷
humidity	1 mbar	-2.73 x 10 ⁻⁸
CO ₂ content	100 ppm	1.47 x 10 ⁻⁸

Table 7.2 - Effect of air temperature, pressure, humidity and CO₂ content on refractivity

7.3.1 Pressure measurement

The pressure transducer used in the PLBI is a Druck DPI 140. This instrument has a range of 35 - 1150 mbar (absolute), with a resolution of 0.01 mbar. The manufacturer's estimated accuracy of the instrument is ± 0.02% of reading ± 0.01% of full scale, or approximately ± 0.23 mbar, though with direct calibration against NPL primary standards, an estimated uncertainty of ± 0.08 mbar can be achieved.

The DPI 140 is situated in the equipment rack and is connected via a tube to the chamber. The opening of the tube is positioned on the carriage, close to the length bars,

ensuring that the pressure sense point is close to the measurement beam. To prevent the pump used to sample the chamber air from affecting the pressure, it is switched off before the pressure readings are made. The DPI reading is available over the IEEE bus, and is read by the computer, which applies a software correction to the reading according to the latest calibration.

The DPI works by measuring the resonant frequency of an oscillating cylinder, which is dependent on the density of the air surrounding the cylinder, which in turn is dependent on the air pressure. The measurement is thus a measurement of density and is affected, to second order, by humidity and other gaseous contaminants. The effect of these has been minimised by the use of the sealed chamber, situated in a humidity-controlled laboratory. From the manufacturer's handbook, it is seen that the errors given in table 7.3 arise in the pressure reading due to a change from 0% Relative Humidity (RH) to 70%, at the stated temperatures.

Temperature (°C)	Error (% reading)	Error at 1 atm (mbar)
10	0.004	0.04
25	0.010	0.10
30	0.013	0.13

Table 7.3 - Errors in pressure reading due to change from 0% to 70%RH

7.3.2 Temperature measurement

The temperature of the air inside the chamber is measured *in situ* by a platinum resistance thermometer. The PRT is enclosed in a small heatsink, placed on an insulating support, close to the length bars. The PRT thus measures the temperature of the air within a small region, which requires the temperature along the measurement beam to be uniform. An alternative would be to use many PRT sensors along the path, but the extra time taken to measure these sensors (approximately 20 seconds each) would allow drift of the readings of temperature and pressure and thus outweigh the advantage of having more measurement points. The temperature gradients along the beam paths are very small and do not affect the air temperature measurement: when the chamber is temperature-controlled at 20 °C, the temperature gradient along the measurement path has been measured as less than 0.001 °C per metre, and hence only one temperature measurement is required. The drift rate is less than 0.002 °C per hour.

The resistance of the PRT is measured using a Tinsley Senator 5840D precision AC resistance bridge which is based on an NPL design. The bridge uses a series of computer-switched precision transformer windings to balance the voltage in the two

arms of a Wheatstone bridge. The decade-windings are switched in or out according to the voltage difference between the two arms. The 1:1 windings are used first, followed by the 10:1, then the 100:1 and so on, gradually decreasing the out-of-balance signal. A digital display can be configured to show either the resistance reading, resistance ratio to an external standard resistor, or temperature calculated by an external computer, connected via the IEEE bus, along which the resistance reading or an out-of-balance indicator signal can be sent.

The PRT resistance is measured in a 4-terminal configuration to account for the resistance of the cables. The bridge contains a standard resistor, which is temperature controlled. The resolution of the bridge is 0.000 01 ohms, corresponding to a temperature resolution of 0.000 03 °C. When calibrated against an external 100 ohm standard resistor, the accuracy of the bridge is estimated to be $\pm 0.000\ 47\ ^\circ\text{C}$ (0.47 mK).

Each PRT used with the bridge is individually calibrated at the water triple point and the melting point of gallium, 0.01 °C and 29.7646 °C, respectively. The uncertainties in the realisations of these standard temperatures are $\pm 0.000\ 5\ ^\circ\text{C}$ and $\pm 0.000\ 5\ ^\circ\text{C}$, respectively. Calibration of the resistance of the PRTs at these temperatures allows calculation of corresponding resistance-temperature coefficients according to the International Temperature Scale of 1990 (ITS-90). Following a measurement of the resistance of a PRT, the ITS-90 coefficients allow the corresponding temperature to be calculated to within an accuracy of 0.000 13 °C (see § 8.3).

7.3.3 Humidity measurement

The water vapour content (humidity) of the air is measured by a Michell S3000 dewpoint hygrometer. This device measures the dewpoint of air passing over a temperature-controlled mirror by controlling the temperature of the mirror to the point where dew forms on its surface. A photocell senses the intensity of light reflected from the mirror surface - this signal decreases when dew forms on the mirror. The temperature of the mirror is monitored by a PRT. A simple formula [11] is used to convert the dewpoint temperature, T_{dp} , into a partial pressure of water vapour, f , in mbar.

$$f = \log^{-1} \left(\frac{7.5 \times T_{dp}}{237.3 + T_{dp}} + 0.78571 \right) \quad (7.10)$$

The resolution of the S3000 is 0.1 °C dewpoint and is calibrated against traceable humidity standards to provide an accuracy of ± 0.25 °C. The dewpoint reading is converted into a voltage which is read by a DVM with an IEEE interface. Overall accuracy of humidity measurement is estimated to be ± 0.3 °C (± 0.24 mbar vapour pressure).

Due to the heat output of the peltier cooler attached to the dewpoint mirror, the sensor is mounted outside the chamber. Air is sampled from the chamber from a point near the length bars by a PTFE-coated re-circulating pump. The air flows along PTFE/stainless steel tubes, through the S3000 sensor, through the CO₂ sensor, and is returned to the chamber. The pump surfaces are PTFE and are sufficiently clean for analytical gas sampling operations and hence do not contaminate the returned air. The flow rate of the sampled air is approximately 0.5 litres min⁻¹.

7.3.4 Carbon dioxide measurement

The CO₂ content of the sampled air is measured using an Edinburgh Instruments GASCARD CO₂ sensor. This measures the relative absorption of light passing through a sample tube of gas. The sampled gas flows through a single chamber, approximately 10 cm long, though which a wide-band light source shines. The intensity detected at certain wavelengths is used to provide an output signal (monitored by an IEEE DVM) which corresponds to the CO₂ molar density. The instrument has a resolution of ± 18 ppm CO₂ and an uncalibrated accuracy of ± 60 ppm. The device is calibrated using standard gases at 0 ppm and 370 ppm (by volume) to achieve an estimated accuracy of ± 30 ppm.

7.3.5 Air parameter measurement order

The measurements of air temperature, pressure, humidity and CO₂ concentration are performed at the same time as the phase-stepping measurements. The order of the individual measurements is given below. The order of measurement of devices is designed so that parameters which are slowly varying (humidity, CO₂, bar temperature) are read at the start and end of the measurement, whereas more quickly varying parameters (air pressure and temperature) are read at the same time as the phase-stepping is performed. The CO₂ and humidity readings are taken before the sample pump is temporarily halted.

- (1) The air temperature PRT is selected.
- (2) The humidity sensor is read by the computer.
- (3) The CO₂ sensor is read by the computer.
- (4) The sample pump is stopped.
- (5) The air temperature PRT is read.
- (6) One of the two PRTs for measuring the temperature of the length bar is read.
- (7) The phase-stepping is performed at three wavelengths.
- (8) The air pressure is read by the computer.
- (9) The second bar temperature PRT is read.
- (10) The sample pump is re-started.

From the measurements of pressure, temperature, humidity and CO₂ content, the refractivities at the three measurement wavelengths are calculated using the modified Edlén equations. The wavelengths corrected for the effects of refractive index are used in the length calculation routine.

7.4 OTHER WAVELENGTH COMPENSATION TECHNIQUES

7.4.1 Two-wavelength compensation

Matsumoto and Honda [12] used a two-wavelength interferometer based around a YAG laser with a frequency-doubling crystal (KTiOPO₄) producing both 1.06 µm and 532 nm wavelengths. The interferometer averaged fringe counts at the two wavelengths to derive a wavelength-corrected fringe count which corresponded to the length being measured by the moving part of the interferometer. An accuracy of length measurement of 0.2 µm over 1 m of travel was obtained in air.

7.4.2 Gas refractometry

The refractive index of air can be measured directly in a gas refractometer. The refractive index, or the refractivity (the refractive index -1) of a sample of gas is measured relative to that of a known standard, often vacuum.

In order to check the accuracy of the results of the modified Edlén equation and as a possible second technique for refractive index determination, a gas refractometer was

constructed at NPL by Mr P D West, based on the design of the NPL Gas Refractometer of Birch and Downs [13,8].

7.4.2.1 Gas refractometer design

The design of the refractometer is shown in figure 7.2.

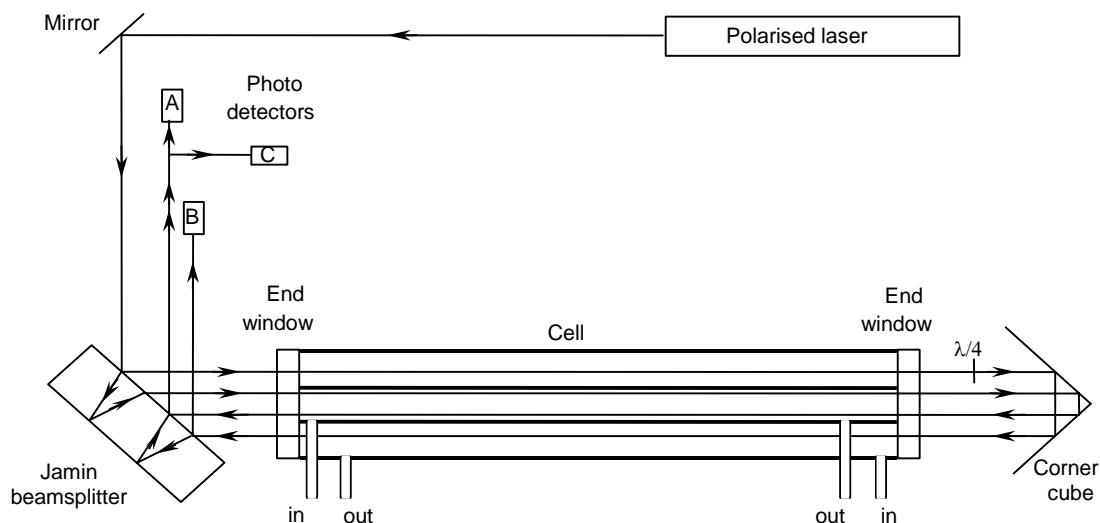


Figure 7.2 - Gas refractometer schematic

This instrument is based around a polarised interferometer using a Jamin beamsplitter and a linearly polarised laser. A concentric two-chamber stainless steel gas cell with quartz end plates is used to contain the sample and reference gases. The inner cell is used for the reference, which is usually vacuum and the outer cell for the sample gas. The external solenoid valves and pipework allow interconnection of the two chambers allowing serial flow of the sample gas though both chambers, gas-outer & vacuum-inner, or both vacuum.

Vacuum is provided by a simple rotary pump (Edwards Model 1.5), which achieves a vacuum of 3×10^{-2} mbar. The three photo detectors are gain matched: two provide the orthogonal output signals from the interferometer, the third is used to adjust the gain of the other two detection channels, based on the intensity of the laser beam.

The internal length of the cell is approximately 316.6 mm, corresponding to 10^6 interference fringes, at the laser wavelength, thus when comparing the optical path lengths of the two cell chambers, a difference in refractive index of 1×10^{-6} will be equivalent to 1 interference fringe. This is the basis of operation of the interferometer.

7.4.2.2 Refractometer operation

Sample gas is flowed through both chambers using the sampling pump for a period of 10 minutes. The pump is then switched off to allow the gas to stabilise at ambient pressure. A fringe counter, connected to two photocells which monitor the two orthogonal outputs of the interferometer, is zeroed. The sample pump is re-started and the inner chamber evacuated. As the air is removed, the change in refractive index causes a change in optical path length of the reference cell. The signals from two orthogonal outputs then exhibit a sinusoidal variation corresponding to the passage of interference fringes across the two detectors. The fringe counter tracks these fringes by tracking the two outputs.

When a sufficient vacuum has been obtained in the inner chamber, the display of the fringe counter will be the refractivity of the sample gas, with a factor of 10^{-6} . When a reading is required, the sample pump is stopped for 60 secs before the reading is taken. As an example, a fringe count of 267.123 corresponds to a refractive index of 1.000 267 123.

7.4.2.3 Corrections and errors

Due to constructional and operational constraints, a number of corrections have to be applied to the reading of the fringe counter to obtain the correct refractive index of the sample gas.

(1) Cell length

Due to engineering tolerances, the cell length is not exactly 1 million half-wavelengths long, thus the fringe count is first multiplied by a correction factor of 0.999716 (determined from accurate measurements of the cell length using a Co-ordinate Measuring Machine and a Gauge Block Comparator).

(2) Imperfect vacuum

The refractive index of absolute vacuum is 1 by definition. The refractive index increases by 2.7×10^{-9} per 10^{-2} mbar, thus due to the imperfect vacuum achieved by the rotary pump, a correction of 8.1×10^{-9} is added to the fringe count.

(3) Cell expansion

The cell is made of stainless steel which has a thermal expansion coefficient of $14.7 \times 10^{-6} \text{ K}^{-1}$. Thus the cell length is dependent on temperature. A correction factor of $5 \times 10^{-9} \text{ K}^{-1}$ departure from 20°C is applied to the fringe count.

(4) Gas adsorption

It has been shown [14] that when moist air is admitted to a previously evacuated cell in order to measure its refractivity, some water molecules are absorbed by the metallic walls of the cell, leading to a decrease in the humidity of the air with a corresponding increase in the refractivity of up to 1×10^{-7} . This effect can be removed by continuously flushing with ten cell volumes of gas, prior to making a measurement. Alternatively, the cell can be used in the reverse configuration with the two chambers of the cell initially flushed with air before one is evacuated (as used in the refractometer used with the NPLBI).

However, the process of physisorption [15] can still lead to an error. Here, water molecules in the stationary air are adsorbed onto the stainless steel wall, leading to an increase in the refractivity of up to 1.5×10^{-7} . The effect can be minimised by increasing the flow rate of the gas prior to stopping the flow for measurement and by reducing the overall size of the refractometer cell tubes.

(5) Pressure effects

Unless the pressures at the sample point and within the cell are equal, there will be a pressure-dependent refractivity gradient between these two points, and the refractivity measured by the refractometer will not be the same as that at the sample point. Thus, the effect of gravity on the air pressure must be taken into account, and the cell placed at the same height (above sea level) as the sample point. The earth's pressure gradient near the surface causes a change in refractive index of $-3 \times 10^{-8} \text{ m}^{-1}$ ascent.

A second order pressure effect is due to the optical path length changes induced in the cell windows by the pressure change as air is evacuated from one of the cell chambers. For a 320 mm long cell, this has been shown [16] to cause a change in the measured refractivity of up to 1×10^{-8} .

(6) Temperature effects

The refractometer measures the refractive index of the gas under the conditions present inside the cell chamber. If there is a temperature gradient between the sample point and the cell, then the refractive index will be different at these two points and a correction factor, based on Edlén's equation will be required. When the length bar interferometer is being used to perform thermal expansion measurements, the temperature of the air inside the sealed chamber can be as high as 30 °C. The temperature of the air surrounding the refractometer is approximately 20 °C. The specific heat capacity of air is so low, that the air sampled from the interferometer chamber cannot supply enough energy to heat the refractometer cell by more than approximately 0.5 °C, and so a large temperature gradient exists between the chamber and the refractometer cell. This will cause a change in refractive index due to not only the temperature change (which can be measured) but also a possible change due to possible outgassing of the connection pipes or water condensation inside the cell, which cannot be measured.

(7) Fringe interpolation

As mentioned in section 2.3.1, errors arise in the fringe interpolation of the fringe counter due to effects such as gain mis-match between the two signals, offsets and phase errors. Some degree of computer correction is possible [17] by fitting ellipses to the data, although this takes time and requires low-noise signals. In the present refractometer, drift of the electronics or the alignment of the interferometer causes the offset and gain for the two channels to vary, leading to fringe interpolation errors.

7.4.2.4 Comparison between Edlén and refractometer

To compare the performances of the modified Edlén equations and the refractometer, they were both used to measure the refractive index of air sampled from the interferometer chamber.

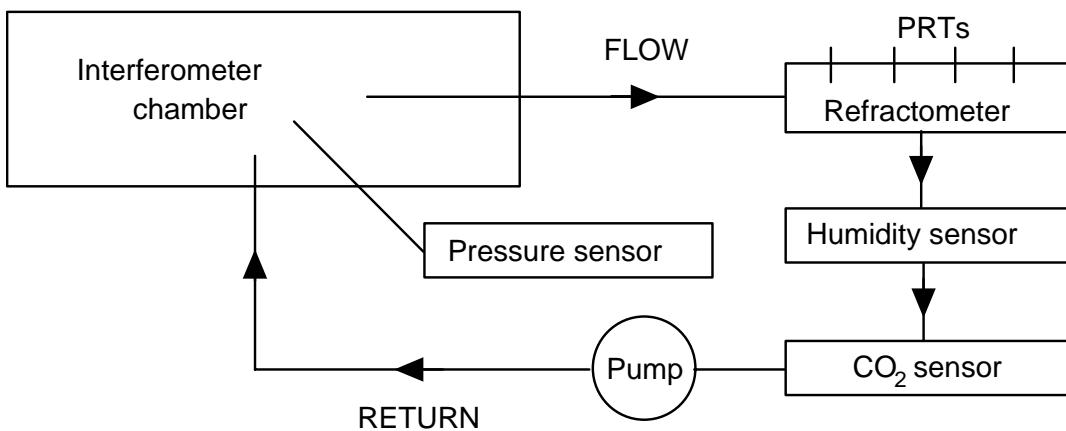


Figure 7.3 - Comparison of Edlén and refractometer

The value of the refractive index measured by the refractometer was compared against that calculated using the modified Edlén equation based on measurements of the air parameters. The air temperature used for the calculations was that inside the refractometer cell and the pressure was measured inside the interferometer chamber.

General agreement between the two methods was $\pm 3 \times 10^{-8}$, as can be seen in figure 7.4. On analysis of the data, it can be seen that the difference between the two methods is periodic with a period equal to one refractometer fringe, as shown in figure 7.5. This indicates that the two channels do not remain in perfect phase quadrature with equal offsets and gains. This was also detected by viewing the outputs of the two channels as a lissajous figure on an oscilloscope.

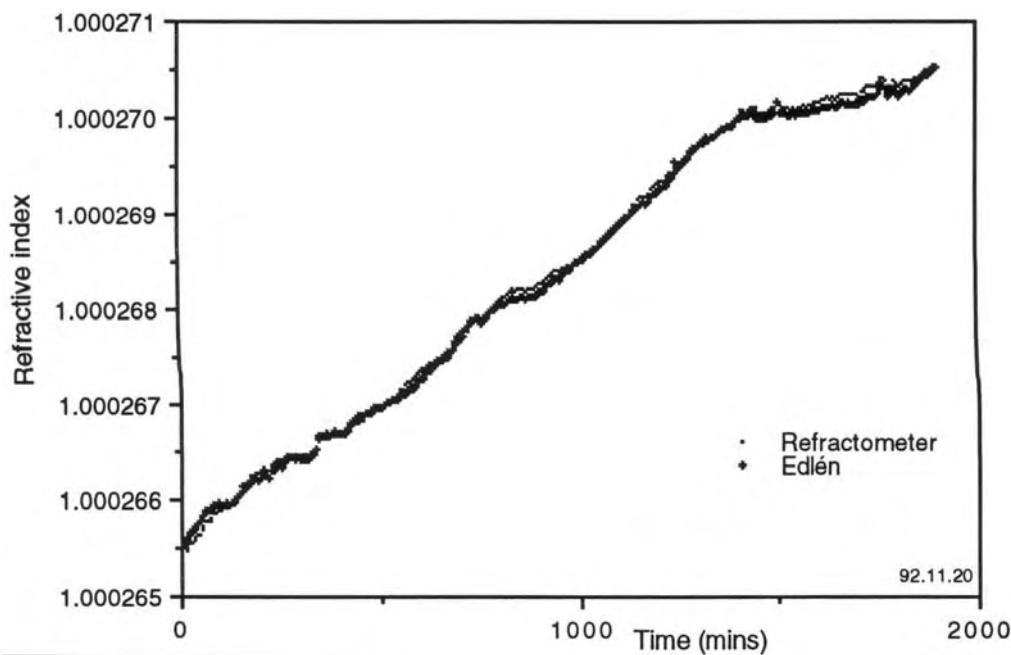


Figure 7.4 - Comparison of Edlén and refractometer

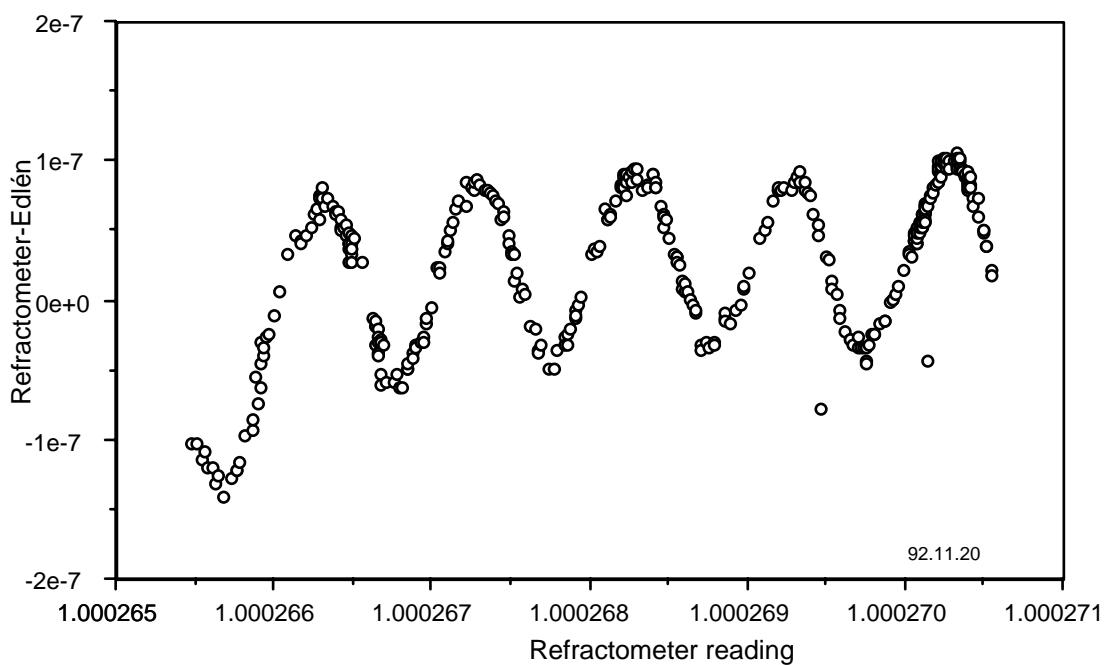


Figure 7.5 - Periodic error due to drift

Thus although an absolute gas refractometer can be used to measure the refractive index of the air, the accuracy of measurement is generally no better than that achievable by the use of Edlén's equations coupled with sensitive, accurate transducers. Because of the fact that the composition of the air inside the chamber does not change rapidly, if at all, the calculation of the refractive index should be at least as accurate as the determination using the refractometer. At elevated temperatures, unless the system is re-designed with an internal refractometer, of a more stable design, Edlén is the only accurate solution to measure the refractivity of the air inside the interferometer chamber.

Birch *et al* [18] compared the results of air refractivity measurements made using an automatic refractometer with those obtained using commercial sensors and the Edlén equations. Their conclusion was that the modified Edlén equation is the most accurate form to use for calculating air refractivity. An uncertainty of $\pm 3 \times 10^{-8}$ is achievable. When account is taken of excess CO₂ levels (above the nominal 300 ppm assumed by Edlén), the modified Edlén equation is accurate to about $\pm 1 \times 10^{-8}$.

7.5 ADDENDUM

Birch & Downs [19] have recently published updated Edlén equations for the refractive index of air. These new formulae take into account: the adoption of SI units (the pascal is now the recommended unit of pressure, not the torr); the establishment of the International Temperature Scale of 1990, superseding IPTS-68; improvements in the equation for the density of moist air [20]; increased CO₂ levels; and revision of the water vapour term. The new standard equations are as follows.

$$(n - 1)_{TP} = \frac{P(n - 1)_s}{96095.43} \times \frac{1 + 10^{-8}(0.601 - 0.00972T)P}{1 + 0.0036610T} \quad (7.11)$$

$$(n - 1)_s \times 10^{-8} = 8343.05 + \frac{2406294}{130 - \sigma^{-2}} + \frac{15999}{38.9 - \sigma^{-2}} \quad (7.12)$$

$$n_{TPf} = n_{TP} - f(3.7345 - 0.0401\sigma^{-2}) \times 10^{-10} \quad (7.13)$$

where T is the temperature in °C, f is the water vapour pressure in Pa, σ is the wavelength in μm and P is the air pressure in Pa. These new formulae are expected to have an uncertainty, over the wavelength range 350 - 650 nm, of ± 1 × 10⁻⁸. These equations will be incorporated into the software which calculates refractive index inside the interferometer.

[2002 re-publication note

Equation 7.12 has since been corrected by Birch & Downs to become:

$$(n - 1)_s \times 10^{-8} = 8342.54 + \frac{2406147}{130 - \sigma^{-2}} + \frac{15998}{38.9 - \sigma^{-2}}$$

There have been several later revisions of these equations by other authors]

REFERENCES FOR CHAPTER 7

- [1] Barrel H & Sears J E The refraction and dispersion of air for the visible spectrum *Phil. Trans. Roy. Soc. London A* **238** (1939) 1-64
 - [2] Svensson K *Arkiv Fysik* **16** (1960) 361
 - [3] Erickson K E Investigation of the invariance of atmospheric dispersion with a long-path refractometer *J. Opt. Soc. Am.* **52** (1962) 777-780
 - [4] Edlén B The dispersion of standard air *J. Opt. Soc. Am.* **43** (1953) 339-344
 - [5] Edlén B The refractivity of air *Metrologia* **2** (1966) 71-80
 - [6] Rank D H, Saksena G D & McCubbin T K Measurements of the dispersion of air from 3651 to 15300 Angstroms *J. Opt. Soc. Am.* **48** (1958) 455-458
 - [7] Peck E R & Khanna B N Dispersion of air in the near-infrared *J. Opt. Soc. Am.* **52** (1962) 416-419
 - [8] Birch K P & Downs M J The results of a comparison between calculated and measured values of the refractive index of air *J. Phys. E: Sci. Instrum.* **21** (1988) 694-695
 - [9] Muijlwijk R Update of the Edlén formulae for the refractive index of air *Metrologia* **25** (1988) 189
 - [10] Beers J & Doiron T Verification of revised water vapour correction to the refraction of air *Metrologia* **29** (1992) 315-316
 - [11] British Standard BS 1339 (1965) (London: British Standards Institution)
 - [12] Matsumoto H & Honda T High accuracy length-measuring interferometer using the two-colour method of compensating for the refractive index of air *Meas. Sci. Technol.* **3** (1992) 1084-1086
 - [13] Downs M J & Birch K P Bi-directional fringe counting interference refractometer *Precis. Eng.* **5** (1983) 105-110
 - [14] Birch K P & Downs M J Errors in the determination of the refractive index of air *Appl. Opt.* **28** (1989) 825-826
 - [15] Birch K P, Downs M J & Ward R E The measurement of humidity variations in gases resulting from the dispersion of water onto surfaces *J. Phys. E: Sci. Instrum.* **21** (1988) 692-694
 - [16] Birch K P, Downs M J & Ferris D H Optical path length changes induced in cell windows and solid etalons by evacuation *J. Phys. E: Sci. Instrum.* **21** (1988) 690-692
-

- [17] Birch K P Optical fringe subdivision with nanometric accuracy *Precis. Eng.* **12** (1990) 195-199
- [18] Birch K P, Reinboth F, Ward R E & Wilkening G The effect of variations in the refractive index of industrial air upon the uncertainty of precision length measurement *Metrologia* **30** (1993) 7-14
- [19] Birch K P & Downs M J An updated Edlen equation for the refractive index of air *Metrologia* **30** (1993) 155-162
- [20] Giacomo P Equation for the determination of the density of moist air (1981) *Metrologia* **18** (1982) 33-40

CHAPTER 8

Thermal Expansion

“Absence of evidence is not evidence of absence.”
M Rees

8.1 THERMAL REQUIREMENTS

When measuring the length of a length bar in the interferometer there is a potentially large source of error due to the fact that the bars are made of steel, which has a linear coefficient of thermal expansion of approximately $10.7 \times 10^{-6} \text{ K}^{-1}$. This means that the length of the bar will vary with temperature and hence must be referred to a standard temperature. Currently the standard reference temperature for metrological laboratories [1] is 20 °C. Thus there two options for the measurement of length bars:

- (1) measure the length of the bar at exactly 20 °C
- (2) measure the length of the bar at some other temperature and correct the measured length to 20 °C by using a value of the linear thermal expansion coefficient, α .

Both of these options require accurate measurement of temperature, in (1) to be sure that the bar is at exactly 20 °C, and in (2) to apply a correction for the departure of temperature from 20 °C.

The problem with option (1) is that it is difficult to stabilise the temperature of all the bars in the interferometer at exactly 20 °C. Stable temperature conditions usually require good thermal conductivity (in this case of the air in the chamber) and a suitable reference temperature standard such as a melting point or a triple point. In the International Temperature Scale of 1990 (ITS90) [2] the nearest reference temperatures are at the triple point of water ($0.01 \text{ }^{\circ}\text{C} \pm 0.0005 \text{ }^{\circ}\text{C}$) and the melting point of gallium ($29.7646 \text{ }^{\circ}\text{C} \pm 0.005 \text{ }^{\circ}\text{C}$). These are not sufficiently close to 20 °C to allow accurate temperature stability at 20 °C.

The problem with option (2) is that, according to the standards for gauge blocks and length bars, the coefficient of expansion for steel gauge blocks and length bars can vary, or is not defined:

"the generally accepted value ... for steel is 11.5 parts in a million per degree Celsius" (BS 4311 -metric gauge blocks)

[no mention] (BS 888 - imperial gauge blocks)

"in the temperature range 10 °C to 30 °C shall be $(11.5 \pm 1.0) \times 10^{-6} \text{ K}^{-1}$ " (DIN 861 - metric gauge blocks)

"shall be within the limits $(11.5 \pm 1.0) \times 10^{-6}$ per degree Celsius" (ISO 3650 - metric gauge blocks)

[no mention] (BS 5317 - metric length bars)

[no mention] (BS 1790 - metric & imperial length bars)

There is increased awareness of the importance of the coefficient of thermal expansivity, as reflected in the wording of the latest British Standard [3] for gauge blocks:

"It is essential for gauge block manufacturers to use a grade and quality of material which is consistent and to control the processes of manufacture to enable the coefficient of expansion, within the temperature range 10 °C to 30 °C, to be within a tolerance of $\pm 0.5 \times 10^{-6}$ per °C of its stated value."

Hence α may vary between 10.5 to $12.5 \times 10^{-6} \text{ K}^{-1}$ between length bars in the same set (since they are often manufactured at different times from different batches of material) or by $\pm 0.5 \times 10^{-6} \text{ K}^{-1}$ for gauge blocks manufactured to the latest version of BS 4311. The different depth of hardening of bars may also lead to a variation, since short bars are hardened throughout their length whereas longer bars are hardened only partially. According to BS 5317:

"25 mm bars shall be hardened throughout their length. Bars over 25 mm up to and including 125 mm shall be hardened either throughout their length or at the ends only for a distance of not less than 4 mm. Longer bars shall be hardened at the ends only for a distance of about 6 mm and not less than 4 mm from each end."

The hardened and un-hardened materials have different thermal expansion coefficients and hence the bulk average coefficient will depend on the length of the hardened zone, all other factors being equal.

Because of these variations and the emerging requirements from customers for higher accuracy length bar calibrations for length bars which may be used at temperatures other than 20 °C, it was decided that the interferometer would measure the length of bars at a temperature close to 20 °C (the final figure achieved is 20 °C ± 0.03 °C) and correct the length to 20 °C using a nominal value of α . For the highest accuracy measurements, the interferometer would also operate as a dilatometer, i.e. it would

measure the lengths of bars at different temperatures over the range from 20 °C to 30 °C and thus derive an accurate value of α which could be used to accurately correct measured lengths to 20 °C or other temperatures.

8.2 TEMPERATURE CONTROL SYSTEM

Options considered for the temperature control of the interferometer included resistive heating wires, Thermofoil [4] heaters, Peltier effect devices and temperature-controlled flowing fluids. On grounds of cost, ease of use and ability to cool as well as heat, the design chosen was that of temperature-controlled water flowing in pipes inside the interferometer. A commercial temperature-controlled water bath and circulator was chosen to control the temperature of the water and to pump it around the pipes.

The baseplate of the interferometer is mounted on insulating nylon supports spaced at every 100 x 100 mm square. A copper pipe (8 mm diameter) is held against the bottom of the baseplate using steel clips. The pipe is wound into a spiral, shown in figure 8.1, with the pipe doubled-back against itself. The reason for this spiral is that when heat is being supplied to the chamber, the cooler return water runs alongside the hotter inflowing water. The coolest water outflow is next to the hottest inflow, thus the net temperature of flowing water at any point along the piping is approximately constant and equal to the mean of the inflow and outflow temperatures. There is a similar spiral of copper piping in the lid, which is held against the aluminium surface of the inside of the lid. Insulation material is used in the lid and against the side walls of the chamber and the interferometer is operated inside a temperature controlled laboratory ($20\text{ }^{\circ}\text{C} \pm 0.2\text{ }^{\circ}\text{C}$).



Figure 8.1 - Spiral of pipework on lid and baseplate

The water circulator is a Haake F3-CH unit which uses proportional-integral-derivative (PID) control to control a heater and refrigeration unit. The input signal for the controller is the temperature of a PRT placed in thermal contact with the baseplate, near the corner where the pipes are connected. The water flow from the Haake is split into two, one of which flows into the pipe in the lid, the other flows (via a valve) into the pipe below the baseplate. The accuracy of the Haake temperature control circuit is ± 0.02 °C, though any small temperature fluctuations will be integrated out by the thermal mass of the water and the chamber. The water flow rate is approximately 15 litres min⁻¹. The Haake uses both a 1 kW heater and a 0.4 kW refrigerator, operating in push-pull mode with the heater under PID control. A front panel control allows selection of the set-point temperature in 0.1 °C steps. The range of the temperature controller is dependent on the heat exchange liquid used. For water, the temperature is limited to the range 0 °C to 60 °C for safety requirements.

Initial experiments showed that the temperature inside the chamber was not uniform (at temperatures away from 20 °C). It was discovered that the temperature of the (unheated) side walls of the chamber were a few degrees cooler than the baseplate which was hotter than the lid. To solve this, three modifications were made. Firstly, the level of insulation was increased. A 50 mm thick box of Celotex™ Thermal Sheathing [5] was built around the interferometer, sitting on the edge of the optical table. Secondly, the water flow in the pipe below the baseplate was reduced, until the temperatures of the baseplate and lid were within 0.1 °C of each other. Thirdly, heat shields were mounted on the edges of the baseplate - these are thin sheets of sand-blasted aluminium which are heated by conduction from the baseplate. These ‘shield’ the inside of the interferometer from the cooler side walls. (These can be seen in figure 3.25). Figure 8.2 shows the heating/cooling/insulation of the interferometer.

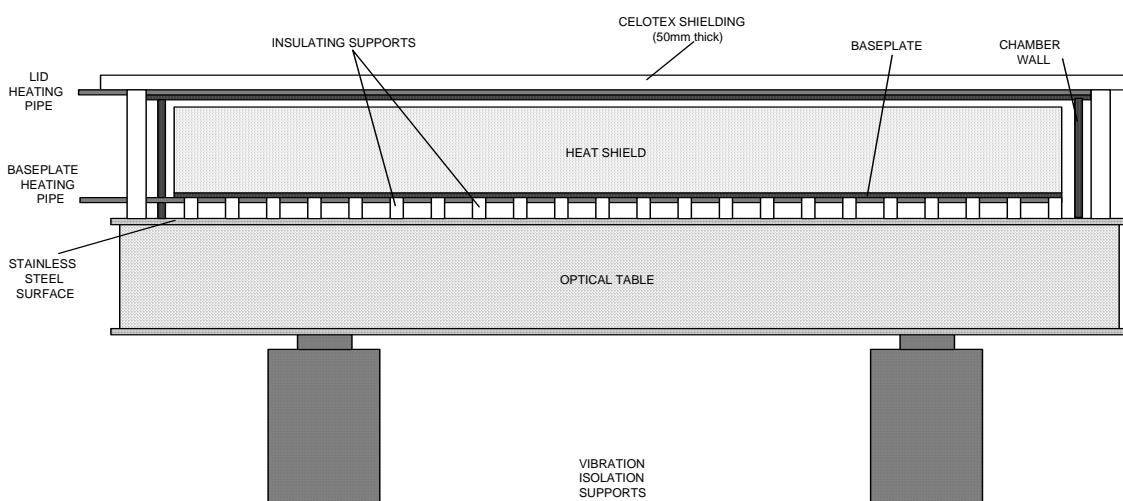


Figure 8.2 - Heating, cooling and insulation of interferometer

The non-uniform temperature caused turbulence and convection of the air inside the chamber leading to refractive index variations which distorted the fringes, making length measurement difficult and inaccurate. After balancing the temperatures of the top and bottom panels, an acceptable level of temperature homogeneity was achieved in both the lengths bars and the air inside the chamber. Details of the verification of this temperature homogeneity and an assessment of residual inhomogeneity are given below. There is a period of convection during the heating/cooling phase after a new set-point temperature is selected, but as the air reaches this temperature and the temperatures stabilise, no turbulence is visible. The video images show straight, stable fringes and the measurements of flatness and parallelism are similar to those made at 20 °C, showing that there is no distortion of the faces of the bar. If the chamber is opened whilst at a raised temperature, the fringes become distorted due to turbulence and thermal distortion of the platen due to the thermal shock

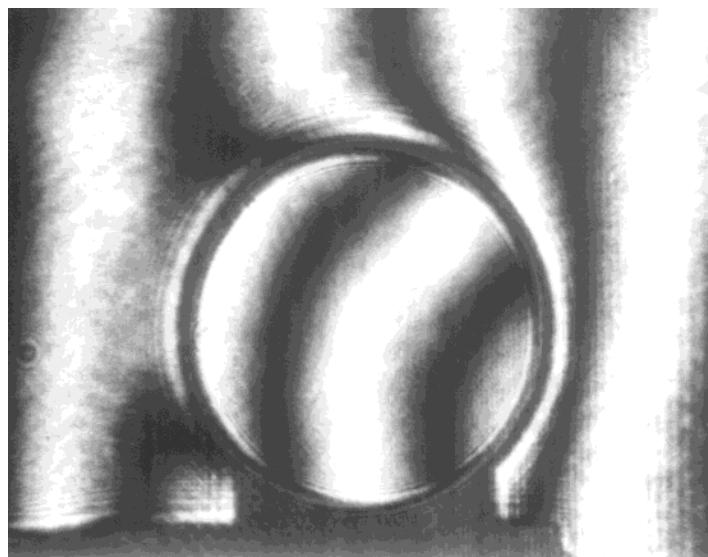


Figure 8.3 - Fringe distortion due to opening of chamber at raised temperature

8.3 TEMPERATURE MEASUREMENT SYSTEM

8.3.1 PRTs and resistance bridge details

All the temperatures inside the chamber are measured using 4-wire miniature platinum resistance thermometers (PRTs) conforming to standard DIN 43760 (1980) [6] and having resistance values within $\pm 0.01\%$ of those specified in that standard at 0 °C (referred to as “1/10 DIN” tolerance). These devices consist of a small coil of pure platinum wire

which has been manufactured by drawing platinum through a small sapphire or diamond die. The coil is suspended in a ceramic housing and has four pure platinum electrodes connected, two to either side of the coil, forming a standard four-terminal resistor. The resistance of the PRT is nominally $100\ \Omega$ at $0\text{ }^\circ\text{C}$, rising to approximately $107.8\ \Omega$ at $20\text{ }^\circ\text{C}$. The ceramic is bonded into a stainless steel sheath, 3 mm diameter and 25 mm long. Four thin wires, individually insulated and 2 m long are connected to the platinum leads. These wires are fed into 2 m long silicon tubing, terminated in a four-terminal LEMO plug. The plugs are inserted into feedthrough connectors in the chamber wall, which in turn are connected to the resistance bridge using individual BNC cables, two per PRT. These cables are designated the Potential (P) and Current (C) connections for the PRT. One cable is used to measure the voltage drop across the PRT when the other cable is supplying a current of 0.2 mA. The self-heating effect of the current in the PRT is approximately 1 - 2 mK, but is taken into account during the calibrations of the PRTs (see § 8.3.2).

The resistance bridge determines the ratio of the resistance of the PRT to the resistance of an internal temperature-controlled standard resistor. The bridge is calibrated using an external $100\ \Omega$ standard resistor, which is calibrated against standards traceable to the NPL realisation of the ohm, using the quantum Hall resistance (see § 7.3.2, § 10.4.2 and § 10.4.7 for further details of the resistance bridge). The PRTs are connected to the bridge via a 15-way selector switch, which is controlled by commands sent to the bridge's IEEE interface from the control computer.

8.3.2 Calibration of PRTs

The PRTs are calibrated by Temperature Section, NPL. To reproduce the conditions in which they are used, the calibration is performed with them connected to a similar resistance bridge, using the same connectors as used in the interferometer. The PRTs are calibrated by measuring their resistances at the triple point of water ($0.01\text{ }^\circ\text{C}$, $\pm 0.000\ 5\text{ }^\circ\text{C}$) and the melting point of gallium ($29.7646\text{ }^\circ\text{C} \pm 0.000\ 5\text{ }^\circ\text{C}$). These are two of the recommended fixed points of the International Temperature Scale (1990). The calibrations are carried out using the same current as in the interferometer. Thus the effects of self-heating are negated. From the values of the resistances measured at these two points, the corresponding ITS-90 coefficients can be calculated.

Periodically, the PRTs are checked by measuring their resistances at the triple point of water using a triple point cell.

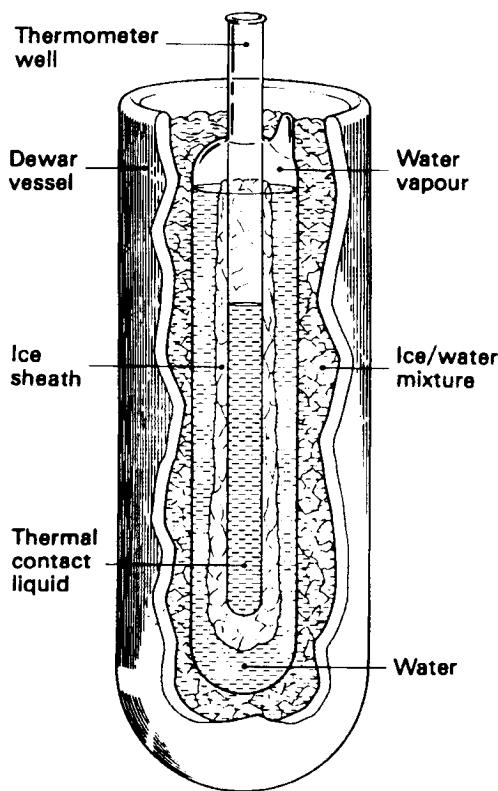


Figure 8.4 - Triple point of water cell used for temperature calibrations

8.3.3 Temperature measurements using ITS-90

From the triple point of equilibrium hydrogen at 13.8033 K to the freezing point of silver at 961.78 °C, the ITS-90 is realised using PRTs. The measurements of temperature are based on reference functions describing the behaviour of the resistance of standard PRTs, and deviation equations describing the departure of a PRT from this reference, measured during calibration. The functions are written in terms of the resistance ratios of the measured resistance at a particular temperature to the resistance at the triple point of water.

$$W = \frac{R(T)}{R(0.01^\circ\text{C})} \quad (8.1)$$

For the range of temperatures above 0 °C, the deviation equation is

$$W - W_{ref} = a(W - 1) + b(W - 1)^2 + c(W - 1)^3 + d \{W - W(660.323^\circ\text{C})\}^2 \quad (8.2)$$

where W and W_{ref} are the thermometer and reference resistance ratios, respectively. The determination of the coefficients a , b , c and d is made from measurements at various freezing points. The ITS-90 guidelines also permit single-point determinations with $b = c = d = 0$, and the measurement performed at the melting point of gallium (or the freezing

point of indium). This still requires determination of $R(0.01\text{ }^{\circ}\text{C})$ and hence is really a two-point calibration which results in a measured value of $R(0.01\text{ }^{\circ}\text{C})$ and a determined value of a .

To calculate a temperature, based on a measurement of resistance of a PRT requires the following procedure. W_{ref} is calculated from the deviation equation, using values of W and a . The appropriate reference function is then used to calculate the temperature. For the range $0\text{ }^{\circ}\text{C}$ to $961.78\text{ }^{\circ}\text{C}$, the reference function is

$$t_{90}/^{\circ}\text{C} = D_0 + \sum_{i=1}^9 D_i \left\{ \left[W_{ref} - 2.64 \right] / 1.64 \right\} \quad (8.3)$$

The constants, D_i , are given in table 8.1.

i	D_i
0	439.932 854
1	472.418 020
2	37.684 494
3	7.472 018
4	2.920 828
5	0.005 184
6	-0.963 864
7	-0.188 732
8	0.191 203
9	0.049 025

Table 8.1 - Constants used in ITS-90 reference equation

The reference equation is accurate to $\pm 0.000\text{ }13\text{ }^{\circ}\text{C}$. The calibration data for the PRTs used in the interferometer are given in table 8.2.

Channel	PRT	$R(0.01\text{ }^{\circ}\text{C}) / \Omega$	a
1	AJL1	100.006 69	-0.018 983 35
2	AJL2	100.001 51	-0.018 848 02
3	AJL3	100.001 86	-0.018 906 89
4	SP1	100.000 76	-0.018 765 43
5	AJL5	100.003 21	-0.018 926 23
6	SP2	100.009 14	-0.019 009 72
7	AJL7	100.007 51	-0.019 267 32
8	AJL8	99.993 67	-0.019 223 35

Table 8.2 - Calibration data for interferometer PRTs

The calibration data is stored in the computer program and is automatically used to calculate values of temperature, based on values of resistance measured by the resistance bridge. When measuring the temperature of a PRT, the program selects the PRT before waiting for the bridge to balance over the next 20 seconds or so. The computer program waits for the ‘balancing’ signal to be cleared, then waits until the temperature readings of the PRT are stable to within 1 mK over a few seconds. Thus no temperature measurements can be made if the temperature is changing rapidly.

8.4 STABILITY OF TEMPERATURES INSIDE CHAMBER

8.4.1 Measurements at 20 °C

Measurements of the temperatures inside the chamber show that the temperature control circuit works well, and controls the temperatures of the bars inside the chamber to $20\text{ }^{\circ}\text{C} \pm 0.03\text{ }^{\circ}\text{C}$, with resetability in this range after heating to $30\text{ }^{\circ}\text{C}$. The temperature control and stability are better than the $\pm 0.2\text{ }^{\circ}\text{C}$ air temperature control of the room, as shown in figure 8.5. A typical drift rate for the air and bar temperatures is less than 2 mK per hour.

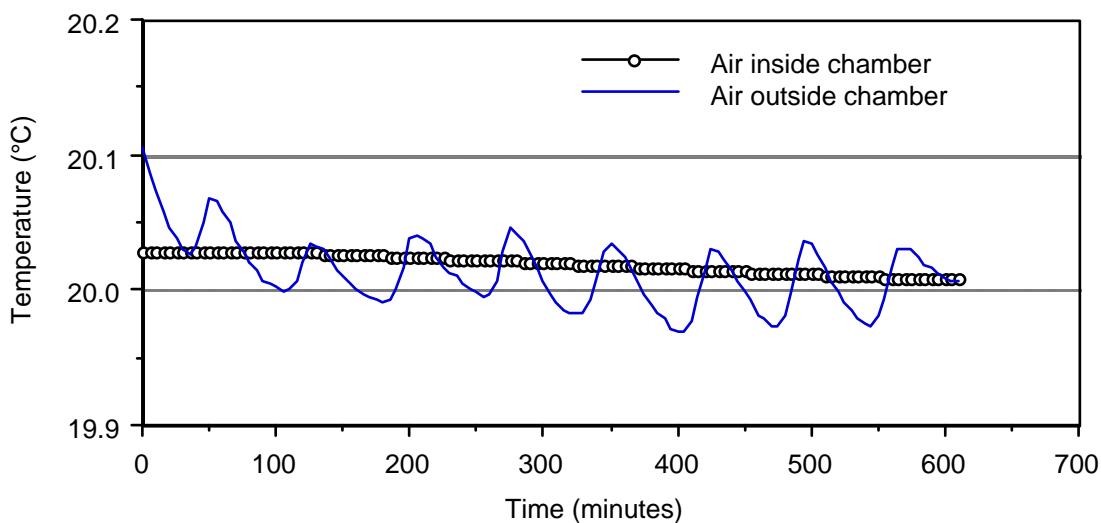


Figure 8.5 - Stability of air temperature inside chamber at 20 °C

Temperature gradients inside the chamber are also very small at 20 °C. The readings shown in figure 8.6 were taken with a 150 mm bar resting on PRTs 1 and 4, a 36 inch bar on PRTs 2 and 5, and a 400 mm bar on PRTs 3 and 6.

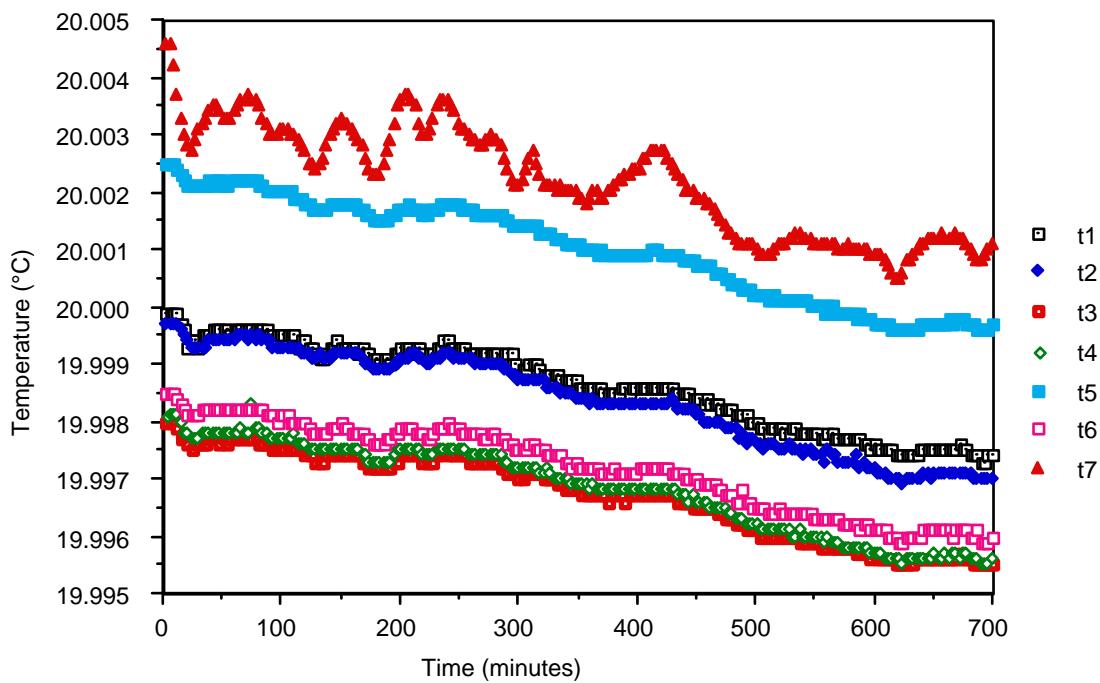


Figure 8.6 - Stability of bar and air temperatures at 20 °C

8.4.2 Heating from 20 °C to 30 °C

Normally, the temperature of the chamber is stepped over the range 20 °C to 30 °C in 2 °C increments when performing thermal expansion measurements. Each temperature step requires approximately 16 hours for the temperatures of the bars and the air to stabilise before measurements are made. This is usually performed overnight, allowing measurement the next day. Measurements are not performed until the temperatures are stable to within ± 1 mK over the time taken for measurement (approximately 2 minutes).

Normally, the PRT that controls the water circulator is placed in thermal contact with the baseplate of the interferometer. Tests have also been performed with the PRT placed inside the water bath of the circulator. As expected, the time required to heat the chamber was increased because the water temperature was stabilised at the set point temperature, rather than being raised higher, to provide faster heating. This is shown in figure 8.7.

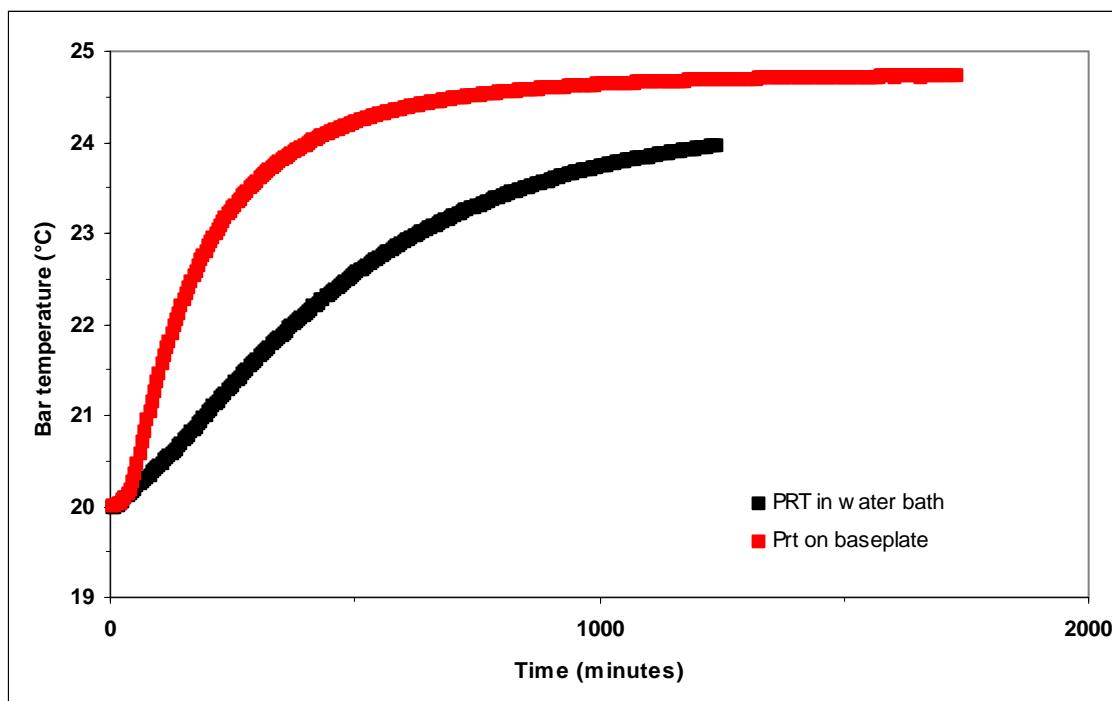


Figure 8.7 - Comparison of heating rates using PRT on baseplate or in water bath

As the temperature is raised, temperature gradients appear inside the chamber due to un-even heating. This leads to a variation in the temperature of the air along the path surrounding the length bar. This causes a temperature gradient in the length bar, with the hottest end being the one to which the platen is wrung.

The temperature gradient has been measured by attaching all PRTs inside the chamber to a 1000 mm length bar, except for the two PRTs which remained in the supports underneath the bar.

The temperature gradients of the air in the measurement path have been measured by placing 5 of the PRTs inside small heat sinks, and placing these in the air alongside a 36 inch bar, supported on the usual 2 PRT supports. The results are shown in figure 8.8, for a temperature of 25.76 °C.

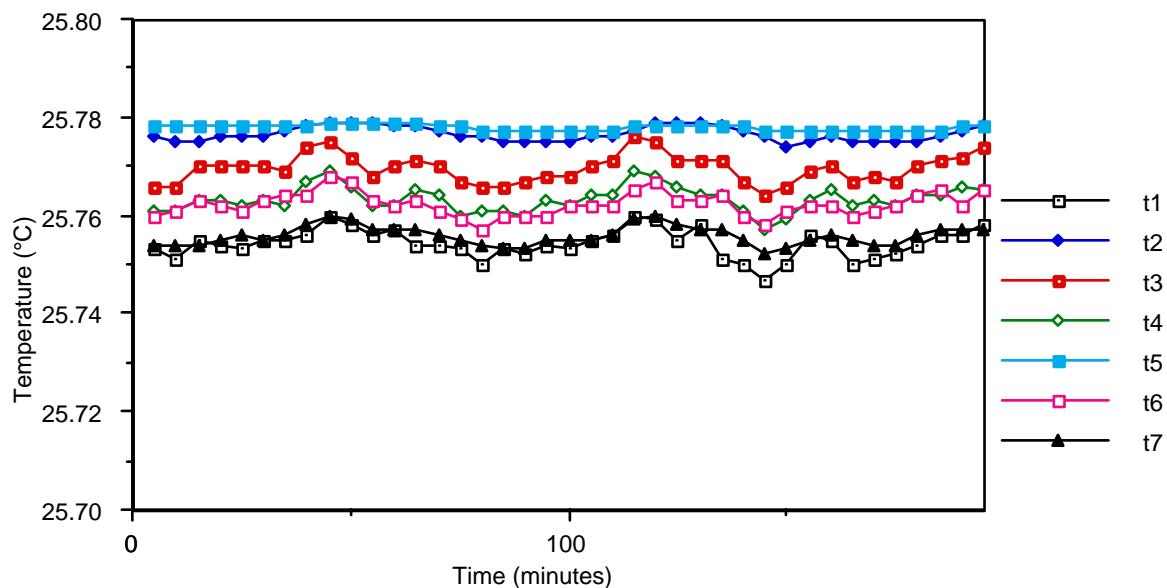


Figure 8.8 - Temperature measurements of air temperature gradients: t_2 & t_5 are the support temperatures, t_1 , t_3 , t_7 , t_4 & t_6 are air temperatures, in order from the unwrung end to the wrung end

8.5 CALCULATION OF THERMAL EXPANSION COEFFICIENTS

The linear coefficient of thermal expansion, α , is defined by equation (8.1).

$$\alpha = \frac{L_2 - L_1}{L_1(T_2 - T_1)} \quad (8.4)$$

where L_1 and L_2 are the lengths of the bar measured at temperatures T_1 and T_2 , respectively. Generally $T_2 > T_1$ and α is positive for steel, hence $L_2 > L_1$. Thus α may be measured by measuring the length of a bar at two known temperatures and then using equation (8.4). A more accurate value can be obtained by measuring at many temperatures and obtaining a set of temperature-length data pairs. This data is analysed as follows. Firstly, the temperatures are all referenced to 20 °C, i.e. 20 °C is subtracted from each temperature reading. A least-squares fit of a quadratic function is then performed. The fitted function is

$$L(T) = L_{20} + \alpha' T + \beta' T^2 \quad (8.5)$$

where T = Temperature - 20 °C, L_{20} = length of bar at 20 °C, α' = linear expansivity, β' = 2nd order non-linear expansivity, $L(T)$ = length of bar at T degrees above 20 °C.

The second order coefficient β' is included to take account of any non-linearity of the expansion. Generally β' is of the order of $\alpha'/1000$ in magnitude for length bar steel.

The required coefficients of thermal expansion α and β are obtained from the expansivities α' and β' by dividing by L_{20} . This gives

$$L(T) = L_{20}(1 + \alpha T + \beta T^2) \quad (8.6)$$

According to standard texts [7], α can be represented by $\alpha = a + bt + ct^2$ where a , b and c are constants, and t is the temperature. In this case, the value of α in (8.6) corresponds to a and β corresponds to b . The coefficient c is very small and over the temperature range encountered in normal laboratory conditions is completely negligible. For this thesis, equation (8.4) will be used to define α , and β will be considered as the departure from linear expansion, *i.e.* the variation of α with temperature.

8.6 ERRORS IN α AND β

A full error analysis of the calculated α and β values is possible by examining the errors of a least squares fit to data pairs, with errors in both variables, using Monte-Carlo techniques. However an order of magnitude estimate can be obtained from the usual theory of error propagation. This will be used to calculate the error in α . Second order β effects can be assumed to be negligible: it will be shown that these are in the ‘noise’ of the measurements. An order of magnitude analysis also gives more insight into what the main sources of error are.

8.6.1 Error propagation method - calculation of error in α and β

If α is calculated from (8.4), then the error $\Delta\alpha$ in the calculation of α is given by

$$(\Delta\alpha)^2 = \left(\frac{\partial\alpha}{\partial L_2} \Delta L_2 \right)^2 + \left(\frac{\partial\alpha}{\partial L_1} \Delta L_1 \right)^2 + \left(\frac{\partial\alpha}{\partial T_2} \Delta T_2 \right)^2 + \left(\frac{\partial\alpha}{\partial T_1} \Delta T_1 \right)^2 \quad (8.7)$$

where

ΔL_1 = error in measurement of L_1

ΔL_2 = error in measurement of L_2

ΔT_1 = error in measurement of T_1

ΔT_2 = error in measurement of T_2

Now

$$\frac{\partial \alpha}{\partial L_2} = \frac{1}{L_1(T_2 - T_1)} \quad , \quad \frac{\partial \alpha}{\partial L_1} = \frac{-L_2}{L_1^2(T_2 - T_1)} \quad (8.8)$$

$$\frac{\partial \alpha}{\partial T_1} = -\frac{L_2 - L_1}{L_1(T_2 - T_1)^2} \quad , \quad \frac{\partial \alpha}{\partial T_2} = \frac{L_2 - L_1}{L_1(T_2 - T_1)^2} \quad (8.9)$$

$$(\Delta \alpha)^2 = \left[\frac{\Delta L_2}{L_1(T_2 - T_1)} \right]^2 + \left[\frac{L_2 \Delta L_1}{L_1^2(T_2 - T_1)} \right]^2 + \left[\frac{\Delta T_2(L_2 - L_1)}{L_1(T_2 - T_1)^2} \right]^2 + \left[\frac{\Delta T_1(L_2 - L_1)}{L_1(T_2 - T_1)^2} \right]^2 \quad (8.10)$$

So far, the analysis is exact. Now use the substitution

$$\frac{L_2 - L_1}{L_1(T_2 - T_1)} = \alpha \quad (8.11)$$

The rest of the results will be exact, if α is correct.

$$(\Delta \alpha)^2 = \left[\frac{\Delta L_2}{L_1(T_2 - T_1)} \right]^2 + \left[\frac{L_2 \Delta L_1}{L_1^2(T_2 - T_1)} \right]^2 + \left[\frac{\Delta T_2 \alpha}{(T_2 - T_1)} \right]^2 + \left[\frac{\Delta T_1 \alpha}{(T_2 - T_1)} \right]^2 \quad (8.12)$$

To remove terms in L_2 , approximate $L_2 \sim L_1$. This is acceptable since there are no terms which involve $L_2 - L_1$.

$$(\Delta \alpha)^2 = \frac{1}{(T_2 - T_1)^2} \left[\frac{\Delta L_2^2 + \Delta L_1^2}{L_1^2} + \alpha^2 (\Delta T_2^2 + \Delta T_1^2) \right] \quad (8.13)$$

Thus $\Delta \alpha$ depends on the errors in the measurements of the temperatures and the lengths, and on the size of α and the temperature step between readings. This is an approximate result, which is exact if the value of α is known, and is not too large, *i.e.* for $L_2 \sim L_1$ to be valid. There is an important distinction to make when selecting contributions to ΔL_1 , ΔL_2 , ΔT_1 and ΔT_2 . Any terms which can be attributed directly to temperature error must not be included in the value for the length errors, even if they contribute a length uncertainty at either temperature, since this would include them twice in the error budget.

The value to use in this analysis for ΔL_1 is the uncertainty in length measurement at 20 °C, which is shown in chapter 10 to be approximately $(\pm 30 \pm 62 L_1)$ nm. The value

of α is taken as $10.7 \times 10^{-6} \text{ K}^{-1}$. The uncertainty in T_1 is shown in chapter 10 to be $\pm 1 \text{ mK}$. The errors in T_2 and L_2 require further consideration.

The error in T_2 has two sources: the non-linearity of the horizontal temperature gradient in the length bar and the vertical temperature gradient across the length bar. A maximum value of the horizontal gradient has been obtained from measurements of the 1000 mm length bar. At 30 °C, the departure from a linear-temperature gradient contributed a temperature measurement error (of the bulk mean temperature of the bar) of 5 mK. At the same temperature, there was a 1 mK temperature gradient across the vertical diameter of the bar, between the PRTs in the supports and those attached to the top of the bar. The magnitudes of these temperature gradients are temperature-dependent.

The sum of these contributions gives

$$\Delta T_2 = \pm 0.0006 (T_2 - T_1) \text{ K.}$$

The sources of error in the length measurement are as follows.

- (1) Inaccuracy of the Edlén equations due to horizontal air temperature gradients between the sensor and the mean position along the length of the bar. This was measured as $25 \text{ mK m}^{-1} \text{ K}^{-1}$. This is length-dependent because the longer the bar, the further away the sensor from the ideal measurement position at the centre of the bar.
- (2) Drift of the alignment of the interferometer. The alignment of the interferometer drifts as the optics are heated due to a differential expansion of the mirror mounts. This causes the reference and measurement beams to become slightly mis-aligned. Before any measurement is performed at a raised temperature, the beams are re-aligned. Further drift is expected to contribute an error of less than 1×10^{-9} per degree temperature excursion.
- (3) Raised temperature inaccuracies of the Edlén equations. The use of the Edlén equations at raised temperatures is expected to result in a length measurement error of $1 \times 10^{-9} \text{ K}^{-1}$.
- (4) The measurement will be subject to the other errors, encountered for measurements at 20 °C. The value give above is $(\pm 30 \pm 62 L_1) \text{ nm}$.

The sum of these contributions gives

$$\Delta L_2 = 2(T_2 - T_1) \text{ nm} + 30 \text{ nm} + 62 L_1 \text{ nm} + 2.3 L_1(T_2 - T_1) \text{ nm.}$$

Substituting these into equation (8.13) gives the following errors for α .

Temperature Step $T_2 - T_1$ (°C)	1	2	3	4	5	6	7	8	9	10
Bar length L_1 (mm)										
100	0.430	0.220	0.150	0.110	0.089	0.075	0.065	0.058	0.053	0.048
200	0.230	0.120	0.078	0.059	0.048	0.040	0.035	0.031	0.028	0.026
300	0.170	0.084	0.056	0.043	0.035	0.029	0.026	0.023	0.021	0.019
400	0.140	0.070	0.047	0.036	0.029	0.025	0.021	0.019	0.018	0.016
500	0.120	0.062	0.042	0.032	0.026	0.022	0.019	0.017	0.016	0.015
600	0.110	0.057	0.038	0.029	0.024	0.020	0.018	0.016	0.015	0.014
700	0.110	0.054	0.036	0.028	0.023	0.019	0.017	0.015	0.014	0.013
800	0.100	0.052	0.035	0.027	0.022	0.019	0.016	0.015	0.014	0.013
900	0.100	0.051	0.034	0.026	0.021	0.018	0.016	0.014	0.013	0.012
1000	0.098	0.049	0.033	0.025	0.021	0.018	0.016	0.014	0.013	0.012
1100	0.097	0.049	0.033	0.025	0.021	0.018	0.015	0.014	0.013	0.012
1200	0.095	0.048	0.032	0.025	0.020	0.017	0.015	0.014	0.013	0.012
1300	0.094	0.048	0.032	0.025	0.020	0.017	0.015	0.014	0.013	0.012
1400	0.094	0.047	0.032	0.024	0.020	0.017	0.015	0.014	0.012	0.012
1500	0.093	0.047	0.032	0.024	0.020	0.017	0.015	0.014	0.012	0.012

Table 8.3 - Error in measured value of α (10^{-6} K^{-1})

8.6.2 Least-squares fit to data with errors in both coordinates

To obtain robust estimates of the errors in the α and β values obtained by least-squares fitting of a quadratic to the length - temperature data requires the adoption of an error analysis such as that proposed by Cecchi [8] (and subsequently corrections by Moreno & Bruzzone [9]). This technique uses the error propagation law and the canonical least-squares equations to estimate the variances in the calculated linear and quadratic terms. A curve-fitting algorithm based on this analysis has been developed by Ben Hughes at NPL for performing exactly the same analysis as required here, for the NPL Gauge Block Dilatometer.

The data given in § 9.6 for the thermal expansivity of a 900 mm bar is reproduced here. Six pairs of data corresponding to the measured length and temperature of the length bar were used as the data for a least-squares fit of a quadratic, using *Mathematica*. The results for the α and β coefficients and the length at 20 °C were:

$$L_{20} = 900.000\ 570 \text{ mm}, \alpha = 10.633 \times 10^{-6} \text{ K}^{-1}, \beta = 8.6 \times 10^{-9} \text{ K}^{-2}$$

From the simple error analysis of the preceding section, the error in α was estimated to be $\delta\alpha = \pm 0.051 \times 10^{-6} \text{ K}^{-1}$, with $L_1 = 900 \text{ mm}$, $T_2-T_1 = 2 \text{ }^\circ\text{C}$ (the temperature step between readings).

The results of the curve fitting algorithm using the same data, but weighted according to the estimates of the errors in the temperature and length measurements, were:

$$L_{20} = 900.000\ 572 \text{ mm}, \alpha = 10.631 \times 10^{-6} \text{ K}^{-1}, \beta = 8.8 \times 10^{-9} \text{ K}^{-2}$$

with estimated variances of $\delta\alpha = \pm 0.04 \times 10^{-6} \text{ K}^{-1}$, $\delta\beta = \pm 4 \times 10^{-9} \text{ K}^{-2}$. Thus the calculated values for the α coefficient differ by only $0.002 \times 10^{-6} \text{ K}^{-1}$, and the β coefficients by $0.2 \times 10^{-9} \text{ K}^{-2}$. The estimates of the length of the bar at 20 °C differ by only 2 nm. The error in the value of α was overestimated by the simple error analysis by $0.011 \times 10^{-6} \text{ K}^{-1}$ (22%). The close agreement is because the errors in the length and temperature measurements are small compared to the values of the measurements. Thus it appears ‘safe’ to use the simple analysis for estimating the error in the measured expansion coefficient. For measurements of expansion coefficient, it is sufficient to use a simple least-squares analysis to calculate the length of the bar and its expansion coefficient, given the size of the other uncertainties (see chapter 10).

8.7 EXAMPLE OF THERMAL EXPANSION MEASUREMENT

As an example of a thermal expansion measurement, a 1000 mm length bar was measured over the temperature range 20 °C to 30 °C. The data were analysed in *Mathematica* and plotted in figure 8.10.

Bar temperature (°C)	Bar Length (mm)
20.009	1000.003 679
21.896	1000.023 850
24.719	1000.054 058
27.727	1000.086 357
30.034	1000.111 151

Table 8.4 - Measured thermal expansion data for a 1000 mm length bar

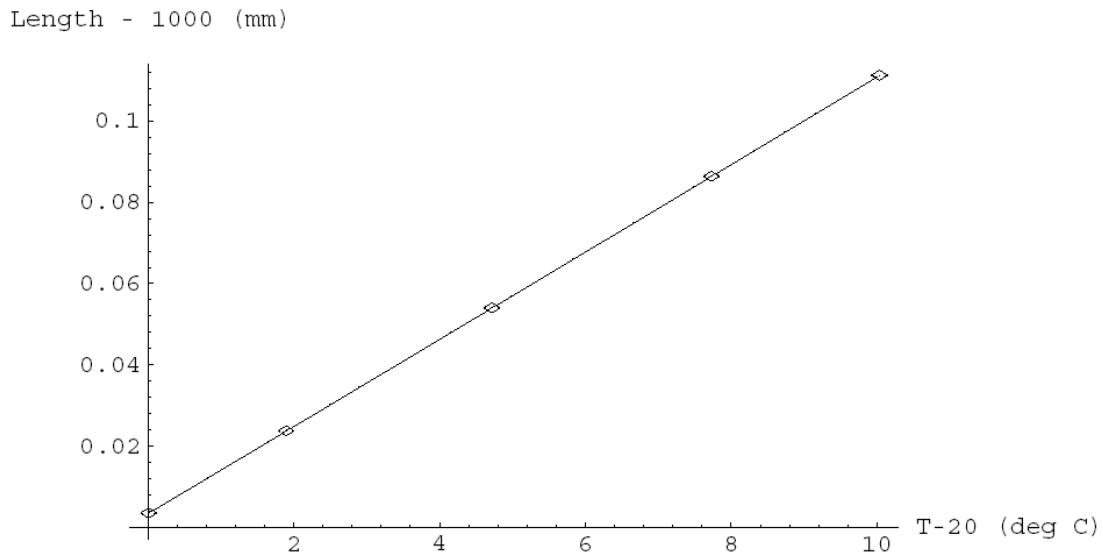


Figure 8.9 - Least squares quadratic fit to thermal expansion data for 1000 mm bar

After least squares fitting, the following coefficients were found:

$$L_{20} = 1000.003\ 580 \text{ mm}, \alpha = 10.678 \times 10^{-6} \text{ K}^{-1} \text{ and } \beta = 4.2 \times 10^{-9} \text{ K}^{-2}$$

Using these coefficients, the agreement with the actual measured lengths at different temperatures ranges from 2.7 nm at 20.009 °C to 4.7 nm at 30.034 °C.

Further examples of thermal expansivity measurements can be found in chapter 9.

REFERENCES FOR CHAPTER 8

- [1] International Standard ISO 1 (1931) (Geneva: International Organisation for Standardisation)
- [2] Rusby R L, Hudson R P, Durieux M, Schooley J F, Steur P P M & Swenson C A Thermodynamic basis of the ITS-90 *Metrologia* **28** (1991) 9-18
- [3] British Standard BS 4311(1993) (London: British Standards Institution)
- [4] Thermofoil™ heaters, Minco Products Incorporated, Minnesota USA
- [5] Celotex™ Thermal sheathing: glass-fibre reinforced rigid polyisocyanurate foam faced on both sides with aluminium foils, Celotex Ltd., Ealing, London
- [6] German Standard DIN 43760 (1980) (Berlin: Deutsches Institut für Normung e.V.)
- [7] Kaye G W C & Laby T H *Tables of Physical and Chemical Constants* 15 edn (1989) (Harlow, Essex: Longmans)
- [8] Cecchi G C Error analysis of the parameters of a least-squares determined curve when both variables have uncertainties *Meas. Sci. Technol.* **2** (1991) 1127-1128
- [9] Moreno C & Bruzzone H Parameters' variances of a least-squares determined straight line with errors in both coordinates *Meas. Sci. Technol.* **4** (1993) 625-636

CHAPTER 9

PERFORMANCE OF THE INTERFEROMETER

*“Deus ex machina.”
 (“A god from the machine.”)
 Menander*

9.1 ASSESSMENT OF THE INTERFEROMETER

One major problem in assessing the performance of the interferometer is that there is no other instrument or set of calibrated length bars at NPL with sufficient accuracy against which comparisons can be made over the full range of the new instrument. Other NPL instruments which can be used for length bar measurements are: the NPL Length Bar Machine (bars 100 mm - 1200 mm, uncertainty $\pm 68 \pm 350L$ nm), the National Standard Multi-axis Co-ordinate Measuring Machine (uncertainty approximately ± 100 - 300 nm), and the NPL Gauge Block Interferometer (length bars 25 - 300 mm, uncertainty $\pm 24 \pm 480L$ nm, L in metres). The only instrument capable of similar accuracy is the Kösters-Zeiss interferometer operated by PTB (Physikalisch Technische Bundesanstalt) - the German national standards laboratory. For brevity, the following acronyms will be used to describe the NPL instruments:

- GBI Gauge Block Interferometer
- LBM Length Bar Machine
- LBI Length Bar Interferometer (the subject of this thesis)

An intercomparison of long series gauge blocks of lengths 600 mm and 1000 mm is planned to take place in 1994, under the auspices of EUROMET. The new Length Bar Interferometer will take part in this intercomparison, as well as the Length Bar Machine. Until this date, NPL has only one length bar that has been measured elsewhere. This is a 36 inch master standard, used as a checking standard in the Length Bar Machine. This bar was measured at PTB in 1988, with results for central length and thermal expansion coefficient. Unfortunately, the faces of the bar have become scratched and pitted, necessitating them being re-lapped, thus shortening the bar. The thermal expansivity is however unaffected and so this bar is useful as a thermal expansion standard.

Chapter 10 contains a full uncertainty analysis for length measurements made by the interferometer - this is the theoretical uncertainty which can be achieved by the interferometer. Unless there is evidence to the contrary, this will be the reported performance of the instrument.

To see whether this level of performance is achieved in practice, a set of length bars (NPL set 1455) has been measured on all available equipment. The set has been measured twice in the LBM and GBI (100 - 1200 mm and 25 - 300 mm respectively) and twice in the LBI, with a gap of approximately one year between the repeated measurements. Measurements of thermal expansion have been made for the 36 inch standard and a group of bars from set 1455. Measurements have also been made with “zero-length” objects, *i.e.* the platen with no bar wrung to it, to assess the accuracy of the optics and the fringe fraction measurement. Thin film fringe fraction samples have also been measured - these are optical glass flats, coated with chromium, with a step height at the centre, in the shape of a gauge block, with a height of less than one fringe. These have been manufactured using thin film deposition techniques and etching.

9.2 FRINGE FRACTION MEASUREMENTS

Two different results have been obtained for the measurements of the thin film fringe fraction samples using the two available data-fitting procedures: Chebychev surface (CS) and best fit plane (BFP). The BFP algorithm is a simple least-squares fit of a plane ($z = ax + by + c$) to the phase data of the platen (for each wavelength). The CS analysis simply fits Chebychev polynomials on a line-by-line basis to 127 horizontal lines through the data. The individual lines are not connected with each other. A full least-squares surface fit, of the form $z = ax + by + cx^2 + dy^2 + e$ will be programmed, when time permits. The two procedures produce slightly different results, depending on the flatness of the platen surface: both will fit well to flat platens, whereas the BFP algorithm will depart from the CS for curved platen surfaces. The answer to the question of which is correct depends on the exact definition of the length of the bar. British Standard BS 5317 simply states that the length is the distance between the centre of the face and a flat surface in wringing contact with the other end. The degree of flatness of the platen is not specified, nor the way in which the corresponding platen surface is determined from the measured data. Due to size/weight considerations, the platens have to be relatively thin, and they are supported by wringing, rather than using proper kinematic supports, and so good flatness is difficult to achieve. Typical flatness of the wrung platens is $\lambda/20$. It is up to the operator to select the desired analysis and to interpret the results. It should be noted however, that the errors due to the data fitting are scaled according to the wavelength - they have the same effect on all three phase

maps, and hence do not influence the choice of nominal order in the multiple-wavelength analysis.

The fringe fraction samples were measured in the Gauge Block Interferometer and in the new interferometer. The GBI uses a best fit least-squares surface, the new interferometer used both CS and BFP analyses.

Fringe fraction measurement ($\lambda = 633$ nm)		
GBI	LBI (BFP)	LBI (CS)
0.313	0.310	0.277
0.647	0.663	0.651

Table 9.1 - Fringe fraction sample results (mean of 50 measurements) BFP - Best Fit Plane, CS - Chebychev Surface

The results using the BFP analysis are in better agreement with the GBI than the Chebychev Surface results. This is because the GBI uses a similar analysis technique. Due to the optical adjustments required to image these samples properly in the new interferometer, the fringe fractions are not measured at the same position as in the GBI. Both samples are out of flat by at least 0.1 fringe (32 nm) thus some differences between fringe fractions measured in the GBI and the LBI should be expected. Considering only the BFP measurements, the GBI and LBI are in agreement to within 0.016 fringe (5 nm).

9.3 ZERO-LENGTH MEASUREMENTS

The results of the zero-length measurements (measurements of platen surfaces with nothing wrung on) are similarly affected by the choice of analysis. Platen number 1 was selected as being the flattest and least scratched of the six interferometer platens. Fifty measurements were made of the platen using both analysis types. The mean results, spread in the results, and the standard deviation are given in table 9.2. These were measured using the mask set up for a length bar, *i.e.* the central region of the phase data is taken to correspond to the surface of the bar, with the surrounding region corresponding to the platen. For a totally flat platen, the measured length should be zero.

Analysis	Mean (nm)	Spread (nm)	Std Devn (nm)
BFP	+ 24.1	3	0.44
CS	+ 6.36	4	0.56

Table 9.2 - Results of zero-length measurements

The two techniques agree on the flatness of the platen as 53 nm (± 2 nm) over the area where the bar would be wrung. The BFP analysis is not suitable for a measurement such as this where the platen is not flat. The CS analysis consistently resulted in a measured length of 6.36 nm, with a spread of 4 nm. This corresponds to a fringe fraction of 0.020 for the red wavelength.

9.4 CENTRAL LENGTH MEASUREMENTS

The measurements of bars from set 1455 were all performed using chebychev surface analysis because this provides more accurate results than the BFP analysis for non-flat platens. The bars were measured both ways round, *i.e.* with the platen wrung to each face in turn. The quoted result is the difference between the mean of these two measurements and the nominal size of the bar, *i.e.* it is the departure from nominal length. The results for the GBI and LBM are also means of two orientations. The difference between the 1st and 2nd wringing results ($|FW - SW|$) for the LBI is given.

The results of the GBI have been corrected to the horizontal position, allowing for prismatic compression of the bar under its own weight (see Appendix D).

Nominal length (mm)	Deviation from nominal (nm)	Difference in LBI results $ FW - SW $ (nm)	LBI -LBM (nm)	LBI -GBI (nm)
100	1666	2	76	59
125	-313	14	-5	-47
150	-571	11	-2	21
175	457	3	27	8
200	125	*	33	1
225	613	3	25	53
250	-926	27	34	-6
275	-449	2	23	70
300	343	10	159	50
400	-1125	64	70	N/A
500	1170	42	69	N/A
600	-1454	31	37	N/A
700	2172	19	161	N/A
800	-776	12	24	N/A
900	452	12	11	N/A
1000	3555	22	-77	N/A
1200	5362	8	108	N/A

* The 200 mm bar could only be wrung one way round due to a burr

Table 9.3 - Comparison results of length bars measured in three instruments

The differences between the results are graphed in figure 9.1.

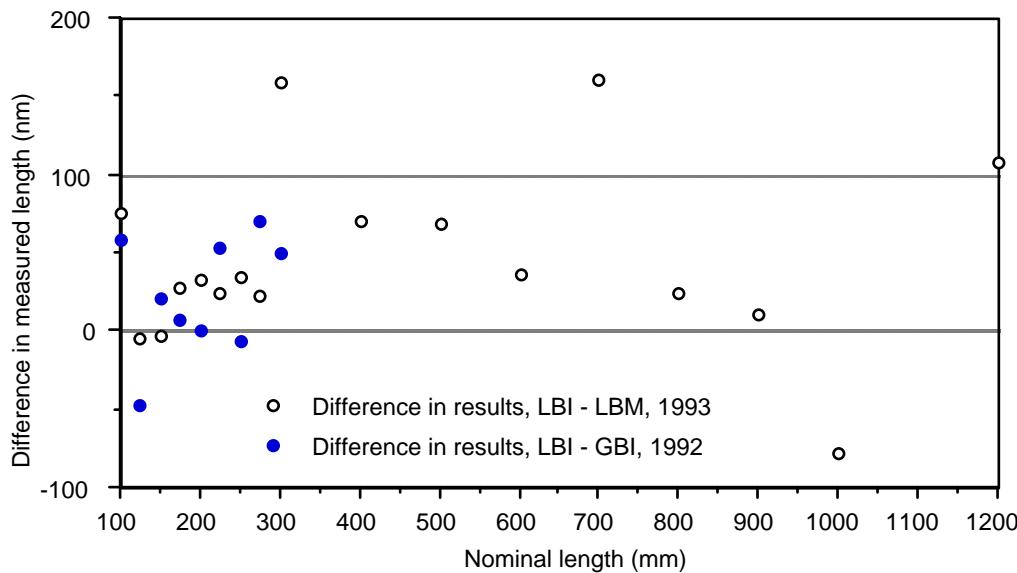


Figure 9.1 - Comparison of length bar measurements from three instruments

A full uncertainty budget for the Length Bar Interferometer is given in chapter 10. The uncertainty for central length measurements is approximately $\pm 30 \text{ nm} \pm 64 L \text{ nm}$, where L is the length of the bar, in metres. The uncertainty budget for the GBI, adapted for length bars, is approximately $\pm 54 \text{ nm} \pm 480 L \text{ nm}$. All the differences between the results of the GBI and the LBI fall within the uncertainty budget of the GBI alone. The uncertainty budget for the LBM is approximately $\pm 68 \text{ nm} \pm 350 L \text{ nm}$. All of the differences between the results of the LBM and the LBI fall within the uncertainty budget of the LBM alone, except for the 300 mm bar. Both the LBM and LBI give repeatable results for this bar which are different by approximately 160 nm. The bar is however not particularly flat and is out of tolerance on parallelism (variation in length) and so some of this discrepancy could be due to the LBM probing the surface in a slightly different place from the LBI (possibly 1 mm separation). The LBI result is a mean of 81 pixels at the centre of the optical face of the bar, whereas the LBM result is a single point contact with the mechanical surface. The poor flatness and parallelism could cause distortion of the platen when wringing, resulting in difference between the two instruments.

The results of the LBM are all consistently longer than the results for the same bars measured one year ago, except for the 700 mm bar, which appears to be shorter by 122 nm. This bar also displays a larger than expected difference between the LBM and LBI results indicating a possible error in the LBM result. The results of the LBI are all consistently shorter than the results for the same bars measured one year ago, except for the 100 mm, 125 mm and 300 mm bars.

Although the result for the 100 mm bar is within the uncertainty budget, there is a significant difference between the LBM and LBI results. This is thought to be due to lifting of the bar when contacted by the LBM probes - the weight of the bar is insufficient to keep it in place on the supports when probed. This effect, which has also been observed with other short bars in the LBM, would constitute a cosine error in the length measurement.

9.5 FLATNESS & PARALLELISM MEASUREMENTS

Some example results of flatness and parallelism measurements were given in figure 6.16. Additionally, 100 repeated measurements have been made of length bars sized 125 mm and 900 mm.

Results in fringes (633 nm)		Mean	Spread	Std Deviation
125 mm	Flatness	0.359	0.04	0.0086
	Parallelism	0.539	0.04	0.0090
900 mm	Flatness	0.224	0.05	0.0106
	Parallelism	0.375	0.05	0.0096

Table 9.4 - Flatness and parallelism results after repeated measurement

Repeatability of flatness and parallelism measurements is within 0.05 fringe (16 nm). The repeatability after the bar is re-wrung and re-measured depends on the quality of the wringing and the positioning of the software cursors, but is similar to the figures in table 9.4.

Figure 9.2 shows a typical screen display (this photograph was taken when the software was still in development and so some of the details have since changed). The screen displays information about the bar down the left side of the image. The red fraction and green fraction (also now the orange fraction) are the measured fringe fractions used in the multiple-wavelength analysis. The Peak and Valley results (now combined as Variation) are the maximum and minimum values of the fringe fraction across the surface of the bar. The calculated length is the length of the bar calculated from the two (now 3) fringe fractions, corrected to 20 °C. The departure is the difference between the measured length and the nominal length input by the user. Down the right side of the display are (top) the phase map ($\lambda = 633$ nm) in a colour representation with scale to the right hand side, and (below) the phase map in a pseudo-three-dimensional display. Note the area which has been masked off shown as dark blue in the upper display and seen as the flat area surrounding the bar in the lower display.

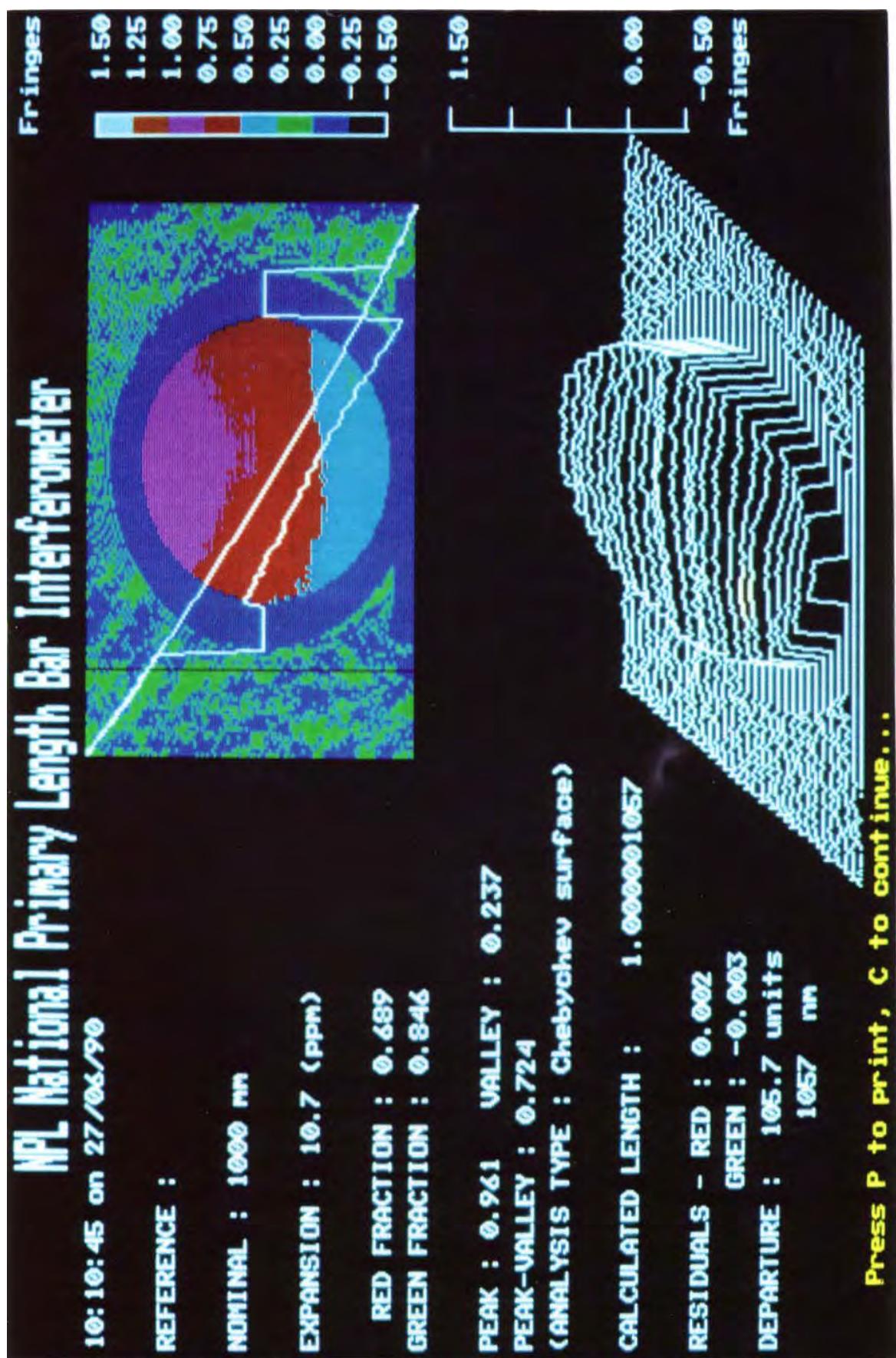


Figure 9.2 - Photograph of screen showing results for a 1000 mm length bar (using a previous version of the software)

9.6 THERMAL EXPANSION MEASUREMENTS

Measurements of thermal expansion have been performed for 6 length bars over the range 100 mm to 1m. The bars were measured at a minimum of 5 different temperatures over the range 20 °C - 30 °C. At each temperature, the alignment of the interferometer was checked using the return-spot technique and the interferometer allowed to stabilise at the correct temperature. Each bar was measured in one orientation only. The bars were selected from NPL set 1455. The data of the measured lengths and temperatures were entered into *Mathematica*, and a least squares quadratic fit was determined. This was of the form $L_{20} (1 + \alpha (T-20) + \beta (T-20)^2)$, where L_{20} is the length at 20 °C, α is the linear coefficient of thermal expansion, β is the second order coefficient of thermal expansion, and T is the temperature of the bar in °C.

100 mm bar

Bar temperature (°C)	Bar Length (mm)
20.006	100.001 689
22.061	100.004 039
23.927	100.006 194
26.035	100.008 630
28.117	100.011 048
29.996	100.013 237

$$L_{20} = 100.001\ 680 \text{ mm}, \alpha = 11.442 \times 10^{-6} \text{ K}^{-1} \text{ and } \beta = 12.0 \times 10^{-9} \text{ K}^{-2}$$

125 mm bar

Bar temperature (°C)	Bar Length (mm)
20.006	124.999 718
22.067	125.002 471
23.941	125.004 996
26.046	125.007 843
28.130	125.010 664
10.167	125.013 437

$$L_{20} = 124.999\ 707 \text{ mm}, \alpha = 10.694 \times 10^{-6} \text{ K}^{-1} \text{ and } \beta = 10.8 \times 10^{-9} \text{ K}^{-2}$$

225 mm bar

Bar temperature (°C)	Bar Length (mm)
20.006	225.000 620
21.887	225.005 130
24.711	225.011 897
27.734	225.019 188
30.037	225.024 743

$$L_{20} = 225.000\ 607 \text{ mm}, \alpha = 10.630 \times 10^{-6} \text{ K}^{-1} \text{ and } \beta = 5.8 \times 10^{-9} \text{ K}^{-2}$$

500 mm bar

Bar temperature (°C)	Bar Length (mm)
20.008	500.001 270
21.884	500.011 279
24.715	500.026 408
27.702	500.042 463
30.008	500.054 879

$$L_{20} = 500.001\ 229 \text{ mm}, \alpha = 10.650 \times 10^{-6} \text{ K}^{-1} \text{ and } \beta = 7.2 \times 10^{-9} \text{ K}^{-2}$$

900 mm bar

Bar temperature (°C)	Bar Length (mm)
20.006	124.999 718
22.067	125.002 471
23.941	125.004 996
26.046	125.007 843
28.130	125.010 664
30.051	900.097 523

$$L_{20} = 900.000\ 570 \text{ mm}, \alpha = 10.633 \times 10^{-6} \text{ K}^{-1} \text{ and } \beta = 8.6 \times 10^{-9} \text{ K}^{-2}$$

1000 mm bar

Bar temperature (°C)	Bar Length (mm)
20.009	1000.003 679
21.896	1000.023 850
24.719	1000.054 058
27.727	1000.086 357
30.034	1000.111 151

$$L_{20} = 1000.003\ 580 \text{ mm}, \alpha = 10.678 \times 10^{-6} \text{ K}^{-1} \text{ and } \beta = 4.2 \times 10^{-9} \text{ K}^{-2}$$

Whilst there is no obvious length dependency of the α and β coefficients, it is the two shortest bars which have the largest β coefficients and the 100 mm bar has a markedly larger α coefficient than the other bars, indicating that this bar has perhaps had a different hardening treatment than the others. Whether the β coefficient as measured is a real second-order non-linearity is difficult to say conclusively, since the calculated uncertainties of the α and β values are of the order of $\pm 10^{-8}$. Thus the apparent non-linearity could be due to temperature dependent errors in the length measurement.

All these bars were measured using a nominal coefficient of $\alpha_{nom} = 10.7 \times 10^{-6} \text{ K}^{-1}$. The above results show that this value was not exactly correct for each bar, however the scanning range of the interferometer was sufficient to cope with this and was able to select the correct solution in the multiple-wavelength analysis.

To confirm the accuracy of these results for α and β , the 36 inch standard length bar previously measured at PTB was measured in the PLBI. The PTB result was:

$$L = 914.4 \text{ mm} - 1.04 \mu\text{m} + 9.887 (T-20)\mu\text{m} + 0.005 (T-20)^2\mu\text{m} \quad (10.1)$$

or $\alpha = 10.813 \times 10^{-6} \text{ K}^{-1}, \beta = 5.5 \times 10^{-9} \text{ K}^{-2}$ (10.2)
 $(\pm 0.0065 \times 10^{-6} \text{ K}^{-1}, \pm 2 \times 10^{-9} \text{ K}^{-2})$

The result from the PLBI was:

$$\alpha = 10.798 \times 10^{-6} \text{ K}^{-1}, \beta = 6.5 \times 10^{-9} \text{ K}^{-2} \quad (10.3)$$

$$(\pm 0.012 \times 10^{-6} \text{ K}^{-1})$$

The results for α agree within the measurement uncertainties of the two instruments. The β results are also in close agreement. The length of the bar is not quoted for the PLBI because the bar was re-lapped in 1992 to remove surface irregularities, and was found to be approximately 2 μm shorter than in 1988.

The 100 mm and 125 mm length bars were independently measured in the NPL Gauge Block Dilatometer (GBD) [1,2]. This is a Fizeau interferometer under development at NPL for the measurement of the thermal expansion coefficient of gauge blocks. Unfortunately the results obtained by the GBD for the 100 mm and 125 mm bars were inconclusive because the GBD software had difficulty measuring accurate fringe

fractions - the optics were designed for gauge blocks, not length bars which leave very little of the platen visible for accurate interpolation.

9.7 GAUGE BLOCK MEASUREMENTS

Although the principles of operation are the same for gauge block and length bar measurement, an example of a gauge block measurement is included here for comparison. The main difference is the shape of the mask is now rectangular rather than circular and the image magnification has been decreased to fit the image of the end of the gauge into the imaging plane (see § 3.2.4).

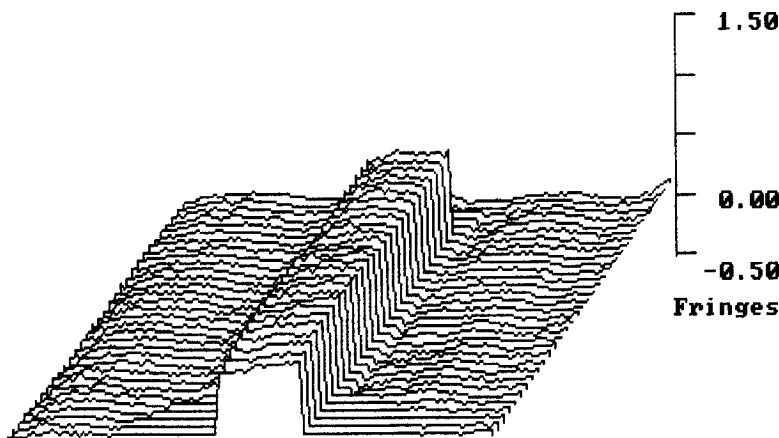


Figure 9.3 - Example measurement of a gauge block

9.8 DOUBLE-ENDED MEASUREMENTS

When the optics are adjusted for double-ended imaging (smaller magnification, carriage displaced laterally, length bar supported at exactly the Airy points), the image digitised into the framestore is similar to that shown in figure 9.4. The right image is the front face of the bar, the left image is the other end of the bar, which would normally be wrung to the platen. After the phase-stepping, discontinuity-removal and surface-fitting have been performed, the resultant 3 phase maps are as shown in figure 9.5.

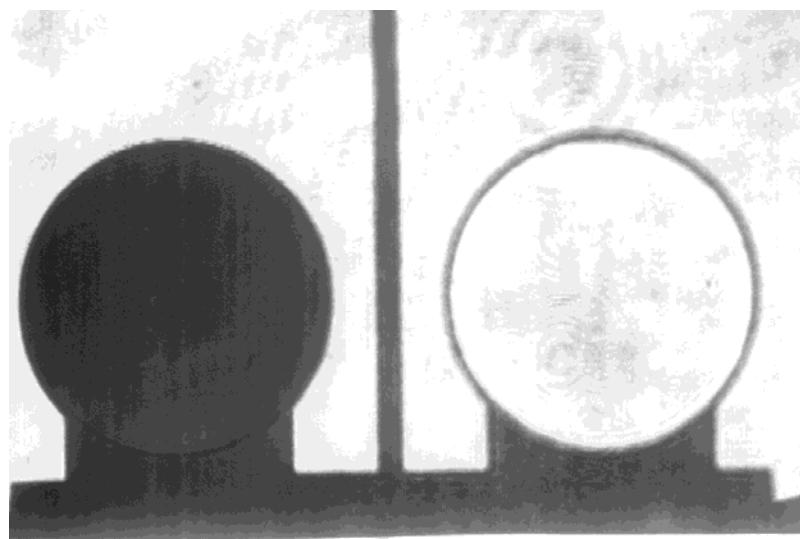


Figure 9.4 - Double-ended image, stored in framestore during measurement

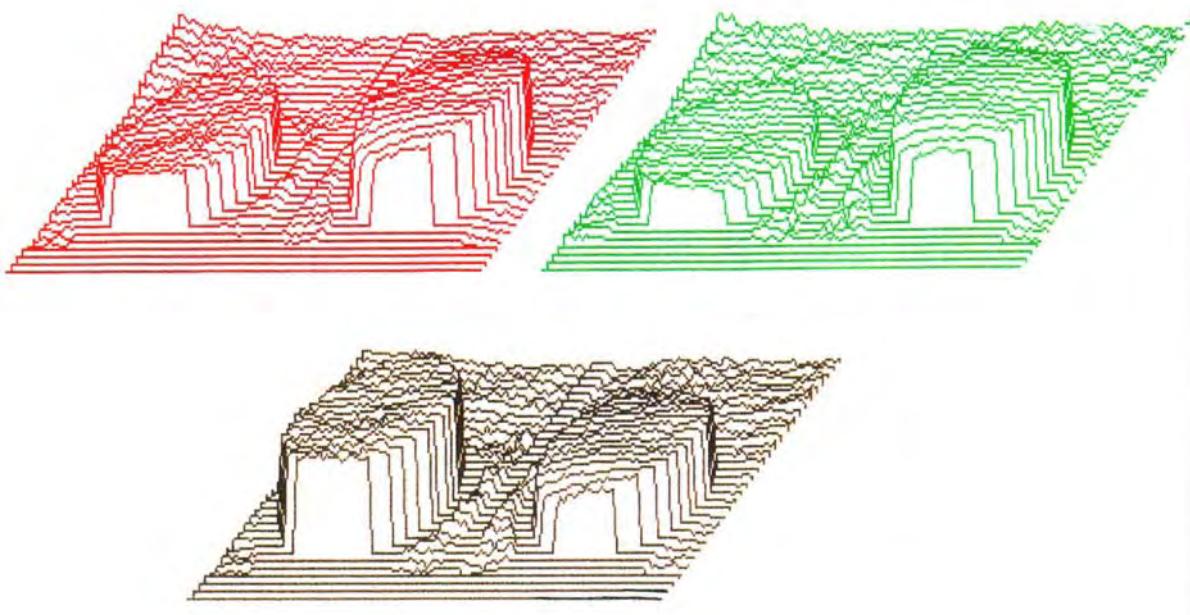


Figure 9.5 - Three phase maps obtained during a double-ended measurement

As shown in figure 9.5, the phase data in the background area of the image is very noisy due to the reduced fringe contrast, visible in figure 9.4 (and more apparent in figure 4.23). Also, the amount of background data available for use in the surface-fitting is much less than in the analysis used with wrung bars, so the fitting of a least-squares plane is less accurate. Any mis-match between the two orthogonal mirrors will appear as extra tilt in either of the 2 sides of the background data, leading to further reduced accuracy in the least-squares plane fitting. For these reasons it has not been possible to measure fringe fractions with sufficient accuracy to use the double-ended multiple-wavelength calculation.

9.9 CONCLUSIONS ON PERFORMANCE

The differences between the measurements of the fringe fraction samples on the LBI and GBI are within 0.016 fringes at $\lambda = 633$ nm. This is similar to the mean of the zero-length measurements implying that the uncertainty in the processed fringe fraction measurements is similar to this, hence a value of ± 0.016 fringes or ± 5 nm will be assumed for the accuracy of the fringe fractions measured by the interferometer in the overall uncertainty budget in chapter 10. In § 5.4 it was shown that the errors due to the phase-stepping are smaller than this so the limiting factor must be the data analysis surface-fitting. Hopefully this will be improved when the least-squares best fit surface is programmed or by using flatter platens.

The flatness and parallelism results are repeatable to within 0.05 fringe (16 nm). It is expected that the parallelism (variation) results are better than the flatness results as the latter requires further data fitting of a least-squares plane, whereas the former uses the data directly.

The intercomparison between the 3 instruments shows that the results for all 3 instruments are within their respective uncertainty budgets. This does not confirm that the uncertainty budgets are exactly correct, but that they are not too small. With more measurements one would expect that 5% of the readings would be outside the uncertainty budgets because they are at the 95% confidence level.

The new interferometer performs as expected and the variation in results is within the uncertainty budget derived in chapter 10.

REFERENCES FOR CHAPTER 9

- [1] Hughes E B NPL Gauge Block Dilatometer - *MSc dissertation* Brunel University (1992)
- [2] Hughes E B Measurement of the thermal expansion coefficient of gauge blocks by interferometry *Proc. SPIE* **2088** (1993) 179-189

CHAPTER 10

UNCERTAINTY OF MEASUREMENTS

"It is much easier to recognise error than to find truth; error is superficial and may be corrected; truth lies hidden in the depths"

Goethe

10.1 THE NATURE OF ERRORS

When a length bar is measured in the interferometer, the result of the length calculation will be subject to an uncertainty due to the design and operation of the instrument. The total uncertainty will be the sum of many contributing uncertainties. These may be due to uncertainties in measured physical quantities, imperfections in the theory describing the interferometer operation, or departure from the theory in the real world. It is important when using the interferometer to measure a bar, to be aware of the uncertainty in the measurement.

10.1.1 The ‘orthodox’ theory of errors

According to orthodox views of error theory [1], there are 2 basic types of error: random and systematic. Random errors can be seen when the measured value of a physical quantity is different under nominally identical circumstances. Systematic errors can arise when a derived correction is applied to measured data, e.g. the refractive index correction.

The two types of error are very different in their effects on the measurement of length in the interferometer. If one makes sufficient measurements, the random uncertainties will be symmetrically distributed about a mean value, which, in the absence of systematic errors, will be the correct value. However, even when many measurements are made with systematic errors present, the calculated mean may be biased away from the true mean, especially if many of the systematic errors add with the same sign, and hence do not cancel each other.

There will also be unknown sources of error, whose nature is unknown. These may cause the cautious experimenter to overestimate the effects of one or other of the types of uncertainty when trying to make allowance for these errors.

Another distinction in the sources of error can be made for a length measuring interferometer. There will be some sources of error which are inherent in the basic design of the instrument, which will contribute an error, even if a ‘zero-length’ object were measured. Other sources of error will depend on the length of the object being measured, *i.e.* they are length dependent. It is useful to quote the total uncertainty of the instrument in a form which separates these two types of error:

$$U = a + bL \quad (10.1)$$

where U is the total uncertainty, a is the inherent uncertainty (random and systematic), b is the length dependent uncertainty (random and systematic), and L is the length being measured. In order to be able to compare random and systematic errors in this way, a common form of reference must be established.

As measurements made by the interferometer will be used at the top of the UK’s hierarchy of traceable length measurements, the calculation and expression of the uncertainty of the result must be made with reference to standard statistical treatments of uncertainty. The basis of the following error analysis is NAMAS document NIS 3003 [2]. This is similar to the draft WECC document 19-1990 [3].

10.1.2 Combination of errors

In the orthodox view, uncertainties or errors are usually combined in quadrature [2,4]

$$U_{tot} = \sqrt{\sum_i U_i^2} \quad (10.2)$$

This is only correct if the estimates of the errors, U_i , are equally weighted, *i.e.* they have the same confidence intervals. For random errors which are normally distributed, this method is correct, as the representative uncertainty of a set of observations is the variance, σ , which always corresponds to a confidence interval of 0.68, or 68% for a normal (Gaussian) distribution. However, the confidence interval of a distribution of errors of a systematic nature is not always the same.

For high accuracy calibrations, such as those offered by the interferometer, it is usual to take a confidence interval of 0.95 (95%) to standardise the uncertainty of measurement when comparing measurements made using different instruments.

10.1.3 Random errors

For the purposes of this error analysis it is assumed that the random uncertainties in a set of N observations or measurements are from a larger distribution, which is itself assumed to be Gaussian. In the absence of sufficient data, the standard deviation can be estimated from the range, R , of the measured values by

$$\sigma = \pm \kappa R \quad (10.3)$$

where κ is approximated by $\kappa \approx \frac{1}{\sqrt{N}}$ (10.4)

The standard error of the mean of the N observations is given by

$$SEOM = \pm \frac{\sigma}{\sqrt{N}} \quad (10.5)$$

To convert this to a confidence interval, the *SEOM* is multiplied by a factor t , the *student t factor*, which depends on the required confidence interval and the number of measurements made. Values of t are tabulated in the literature [5,2,6]. When the behaviour of an instrument or uncertainty is well known, either by having made a large number of measurements, or by assuming an uncertainty from the specifications of the instrument, it is then correct to take a value of t corresponding to an infinite number of measurements. At a confidence interval of 95%, this value is $t = 1.96$ (sometimes referred to as k).

Thus the confidence interval for random uncertainties is given by

$$C_R = \pm \frac{t\sigma}{\sqrt{N}} \quad (10.6)$$

and the total random uncertainty is given by

$$U_R = \sqrt{\sum_R C_R^2} \quad (10.7)$$

10.1.4 Systematic errors

When assessing the effect of systematic errors, an estimate of the standard deviation of a systematic effect on the mean value of the quantity being measured should be used. If this is not possible, then realistic limits for the systematic contribution should be estimated. When a number of error distributions are combined, the Central Limit Theorem states that the overall combined distribution will tend towards a Gaussian. The accuracy of the approximation will depend on the form of the individual distributions and their standard deviations. If it is assumed that a systematic error lies within the bounds $-R/2$ to $+R/2$, then an approximate standard deviation for this distribution will be

$$\sigma = \frac{R}{2\sqrt{3}} \quad (10.8)$$

To convert this to a confidence interval, it is multiplied by a factor k_s , which is dependent on the required confidence level. For a 95% confidence level, $k_s = 1.96$. Thus

$$C_s = \pm \frac{k_s R}{2\sqrt{3}} \quad (10.9)$$

and the overall systematic uncertainty is given by

$$U_s = \sqrt{\sum_s C_s^2} \quad (10.10)$$

According to the NAMAS guidelines, provided that $k_s > 1.8$, the probability of the error falling within $\pm C_s$ will always be greater than for a truly Gaussian distribution of the same standard deviation.

10.2 BIPM RECOMMENDATIONS ON ERROR ASSESSMENT

Many scientific and industrial activities require only rough-and-ready ‘uncertainty’ estimates using simple techniques. However metrologists and others making fundamental physical measurements require a rigorous and objective (*i.e.* demonstrably realistic) theory of errors on which to base accurate estimates of uncertainty. The BIPM has issued recommendations for the estimation of experimental uncertainty [7].

A summary of their recommendations follows.

1. The uncertainty in the result of a measurement generally consists of several components which may be grouped into two categories according to the way in which their numerical value is estimated:

- A - those which are evaluated by statistical methods,
- B - those which are evaluated by other means.

There is not always a simple correspondence between the classification into categories A or B and the previously used classification into “random” and “systematic” uncertainties. The term “systematic uncertainty” can be misleading and should be avoided.

Any detailed report of the uncertainty should consist of a complete list of the components, specifying for each the method used to obtain its numerical value.

2. The components in category A are characterised by the estimates s_i^2 , (or the estimated “standard deviations” s_j) and the number of degrees of freedom v_i . Where appropriate, the estimated covariances should be given.
3. The components in category B should be characterised by quantities u_j^2 , which may be considered as approximations of the corresponding variances, the existence of which is assumed. The quantities u_j^2 may be treated like variances and the quantities u_j like standard deviations. Where appropriate, the covariances should be treated in a similar way.
4. The combined uncertainty should be characterised by the numerical value obtained by applying the usual method for combination of variances. The combined uncertainty and its components should be expressed in the form of “standard deviations”.
5. If, for particular applications, it is necessary to multiply the combined uncertainty by a factor to obtain an overall uncertainty, the multiplying factor must always be stated.

10.3 COMPARISON OF 3 THEORIES OF ERROR AND RECOMMENDATIONS

Colclough [2] compared the orthodox and BIPM recommendations on errors and considered a third theory, the “Randomotic Theory of Errors” in which all errors are treated in the same way as random errors in the orthodox theory. In his analysis, he stated that all errors could be divided into 4 classes, with each error belonging to one class and one class only.

The four classes (illustrated in figure 10.1) illustrate the way in which the observed results of an experiment behave when the experiment is repeated several times:

Class 1 - each result may differ from the true value by the same amount and with the same sign, *i.e.* the error is constant,

Class 2 - each error may vary randomly realising a stable distribution with a non-zero mean,

Class 3 - each error may vary randomly realising a stable distribution with a zero mean,

Class 4 - each error may vary non-randomly (*e.g.* cyclically or by failing to produce convergent distributions, sometimes referred to as a ‘locally systematic error’)

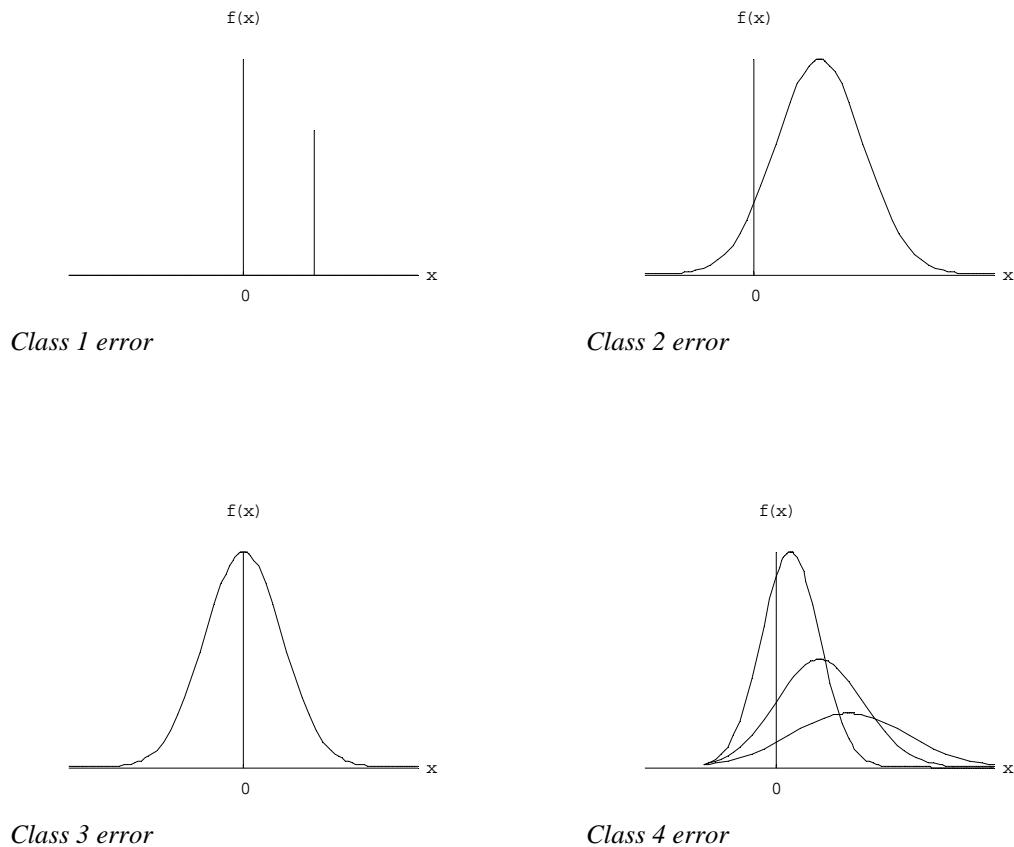


Figure 10.1 - Four classes of experimental error

Colclough showed that all three theories of errors were flawed: the orthodox theory is not rigorous enough in the combination of errors and there is uncertainty as to which results contain random errors; the BIPM technique uses approximations of variances and is still controversial; the Randomotic theory uses unrealistic distributions and raises controversial questions in terms of the law of error propagation. The subject of error

theory still raises controversy particularly since the experimenter has to assess probabilities in the absence of both statistical data and real data.

It is thus difficult to choose a particular technique for calculating the uncertainty budget for the new interferometer. However all the above theories make recommendations which are of use in this situation.

Recommended analysis

Firstly, the whole of the experimental procedure should be defined, and all sources of error identified. A confidence level is chosen, beyond which errors will be regarded as improbable. This confidence level must be clearly stated. Each error is then attributed to a class: random/systematic or class 1 to class 4. This decision is often taken in the absence of trial data by careful consideration of the conditions.

In the case of class 4 errors, either they should be reduced by modification of the experimental technique, or maximum errors of the quantity concerned are computed - these should be treated as systematic errors.

Next, the maximum and minimum possible or likely values of the class 1 errors and the constant components of class 2 errors are estimated, either by reference to assumed specifications or by examining error distributions. These errors are propagated through to the final measurement uncertainty. These components are added arithmetically to give an overall systematic uncertainty in the final result.

All the class 2 and class 3 sources of random errors are identified and propagated through to the final measurement uncertainty. These components are combined in quadrature to obtain a standard deviation for the random error component.

The systematic uncertainties are then used to define upper and lower limits for the mean of the overall random uncertainty giving two worst-case distributions. The upper and lower confidence limits of these two distributions are used to arrive at a final estimate of the uncertainty.

10.4 SOURCES OF UNCERTAINTY

The individual sources of error which affect an individual length measurement made by the interferometer will now be examined. These include errors in the measurement of physical variables in which there may be several contributing uncertainties and also errors due to the design and operation of the interferometer. For each source of error, its magnitude will be estimated along with its effect (random or systematic) including whether or not it is length dependent. The uncertainties are quoted as uncertainties in physical units followed by the corresponding uncertainties converted to length units, where L is the length of the bar, in metres. The class of error is also identified for both the random/systematic and class 1...class 4 schemes, labelled as *e.g.* R 3 for a pseudo-random class 3 uncertainty.

Where error sources relate to manufacturer-specified accuracies or for calibrations of equipment, these are for a confidence level not less than 95%. Thus the effect of these error may be overestimated by a factor of 1.96 in the final calculation - this is tolerable, since in many cases these errors are small and an over-estimation of the final error is better than under-estimation.

10.4.1 Air pressure measurement

The pressure is measured by a Druck DPI140 pressure transducer (see § 7.3.1). The instrument is calibrated at yearly intervals against NPL primary standards. The measurement is performed with dry air over 3 pressure cycles. The deviation of the measured pressure from the accurately known supplied pressure is noted at 9 points during both rising and falling pressure conditions. The calibration is performed at approximately 20 °C.

The DPI140 measures the pressure inside the chamber via a sample pipe. The pipe is at approximately the same height as the length bar being measured. The optical beam diameter at the length bar is 80 mm. The interferometer chamber contains moist air from the room at relative humidity (RH) 50% ($\pm 5\%$).

The following sources of uncertainty have been identified:

Accuracy of NPL working Standard	R 3	± 0.05 mbar	$\pm 1.34 \times 10^{-8} L$
Maximum departure of DPI readings from mean during up/down cycle	R 3	± 0.06 mbar	$\pm 1.61 \times 10^{-8} L$
Error in reading at 50 % RH ($\pm 5\%$), 20 °C due to water vapour	S 1 R 3	+ 0.057 mbar ± 0.0057 mbar	$\pm 1.53 \times 10^{-8} L$ $\pm 1.53 \times 10^{-9} L$
Pressure gradient due to gravity, across beam diameter	S 1	+ 0.0034 mbar	$\pm 9 \times 10^{-10} L$
Resolution of DPI140 instrument	R 3	± 0.01 mbar	$\pm 2.68 \times 10^{-9} L$
TOTAL		± 0.0789 mbar	$\pm 2.12 \times 10^{-8} L$
TOTAL		+ 0.0604 mbar	$\pm 1.62 \times 10^{-8} L$

10.4.2 Air temperature measurement

The temperature of the air in the chamber is measured using a PRT. The PRT is placed in a heatsink and is positioned near to the measurement beam, usually behind the bar being measured. The temperature is measured by measuring the resistance of the PRT using a resistance bridge. The PRT is calibrated at 2 yearly intervals by Temperature Section, NPL, against the water triple point and gallium melting point. Equations conforming to the ITS-90 specification [8] are used to interpolate between these two standard temperatures. The bridge is calibrated monthly by using it to measure the resistance of a standard $100\ \Omega$ resistor, which is itself calibrated yearly. The PRTs are checked every 6 months by using them to measure the temperature of a water triple point cell.

The following sources of uncertainty have been identified:

Resolution of resistance bridge	R 3	$\pm 10\ \mu\Omega = \pm 0.03\ \text{mK}$	$\pm 2.78 \times 10^{-11}\ L$
Resistance bridge accuracy:	R 3	$\pm 1\ \text{ppm} \pm 10\ \mu\Omega = \pm 101\ \mu\Omega = \pm 0.3\ \text{mK}$	$\pm 2.78 \times 10^{-10}\ L$
Accuracy of external standard resistor	R 3	$\pm 8\ \mu\Omega = \pm 0.024\ \text{mK}$	$\pm 2.22 \times 10^{-11}\ L$
<u>PRT calibration</u>			
Water triple point accuracy	R 3	$\pm 0.5\ \text{mK}$	$\pm 4.65 \times 10^{-10}\ L$
Gallium melting point accuracy	R 3	$\pm 0.5\ \text{mK}$	$\pm 4.65 \times 10^{-10}\ L$
Interpolating equations	R 3	$\pm 0.13\ \text{mK}$	$\pm 1.21 \times 10^{-10}\ L$
Drift between calibrations	R 3	$< \pm 0.5\ \text{mK}$	$\pm 9.3 \times 10^{-10}\ L$
TOTAL	R 3	$\pm 0.926\ \text{mK}$	$\pm 8.58 \times 10^{-10}\ L$

10.4.3 Air humidity measurement

The humidity of the air inside the chamber is measured by extracting a sample of the air through a Michell S3000 dewpoint hygrometer. The S3000 is calibrated by a NAMAS accredited laboratory against standard humidity gases at a flow rate of 0.5 l min^{-1} . The voltage output of the S3000 is read by an IEEE voltmeter. The voltmeter is calibrated at the 0 V and 999.9 mV points using a standard voltage generator. The agreement at interpolated voltages is within $\pm 0.2 \text{ mV}$. Magnus' equation [9] is used to convert dewpoint into partial pressure. This has been compared with other techniques, such as Goff-Gratch [10] and found to be in agreement to within 2% RMS over the range 0 to 30 °C.

The following sources of uncertainty have been identified:

Accuracy of dewpoint of standard humidity gases	R 3	$\pm 0.25 \text{ }^{\circ}\text{C DP} = \pm 0.207 \text{ mbar}$	$\pm 5.65 \times 10^{-9} L$
Resolution of S3000	R 3	$\pm 0.1 \text{ }^{\circ}\text{C DP} = \pm 0.083 \text{ mbar}$	$\pm 2.26 \times 10^{-9} L$
Resolution of IEEE voltmeter	R 3	$\pm 0.1 \text{ mV} = \pm 0.01 \text{ }^{\circ}\text{C DP}$	$\pm 2.26 \times 10^{-10} L$
Accuracy of IEEE voltmeter calibration	R 3	$\pm 0.2 \text{ mV} = \pm 0.02 \text{ }^{\circ}\text{C DP}$	$\pm 4.52 \times 10^{-10} L$
Accuracy of standard voltage source	R 3	$\pm 0.2 \text{ mV} = \pm 0.02 \text{ }^{\circ}\text{C DP}$	$\pm 4.52 \times 10^{-10} L$
Humidity gradient between sample point and beam	R 3	$\pm < 0.05 \text{ }^{\circ}\text{C DP}$	$\pm 1.13 \times 10^{-9} L$
Accuracy of Magnus' eqn	R 3	$\pm 0.2 \text{ }^{\circ}\text{C DP}$	$\pm 4.52 \times 10^{-9} L$
TOTAL	R 3	$\pm 0.340 \text{ }^{\circ}\text{C DP}$	$\pm 7.70 \times 10^{-9} L$

10.4.4 Air CO₂ measurement & Edlén's equations

The CO₂ content of the air inside the chamber is measured by extracting a sample of the air (the same as used for the humidity measurement) through an Edinburgh Instruments GASCARD CO₂ meter. The GASCARD meter is calibrated at two points against standard gases with CO₂ concentrations of 0 ppm and 370 ppm CO₂ by volume. This calibration is performed yearly.

The following sources of uncertainty have been identified:

Resolution of GASCARD meter	R 3	$\pm 18 \text{ ppm}$	$\pm 2.65 \times 10^{-9} L$
Accuracy of 0 ppm standard gas	R 3	$\pm 1 \text{ ppm}$	$\pm 1.47 \times 10^{-10} L$
Accuracy of 370 ppm standard gas	R 3	$\pm 30 \text{ ppm}$	$\pm 4.41 \times 10^{-9} L$
Interpolation between calibration points	R 3	$\pm 5 \text{ ppm}$	$\pm 7.35 \times 10^{-10} L$
Variation in concentration between sample point and measurement beam	S 1	- 10 ppm	- 1.47 x 10 ⁻⁹ L
TOTAL		$\pm 35.4 \text{ ppm}$	$\pm 5.20 \times 10^{-9} L$
		S 1	-10 ppm
Accuracy of modified Edlén equation with CO ₂	R 3	$\pm 1 \times 10^{-8}$	$\pm 1 \times 10^{-8} L$

10.4.5 Laser wavelength

The lasers are all frequency-stabilised helium-neon continuous wave lasers operating at 632.990876 nm (red), 543.516364 nm (green) and 611.970617 nm (orange). They are calibrated by direct comparison with NPL Primary lasers, one of which (at approximately 633 nm) represents the UK's realisation of the metre. The calibration is a beat frequency comparison so there is no correction for the refractive index of the air. The measured length of the bar is the length measured by the red wavelength as this has a lower overall uncertainty than the mean of the lengths measured by three wavelengths with equal weighting. The green and orange laser wavelength uncertainties are given here for comparison. The lasers are calibrated by direct frequency comparison against primary reference lasers at NPL. The primary lasers are stabilised by saturated absorption in molecular iodine at the following transitions:

632.991 398 22 nm	$(\pm 2.5 \times 10^{-11})$	11-5 R(127) a ₁₃
611.970 770 0 nm	$(\pm 3 \times 10^{-10})$	9-2 R(47) a ₇
543.516 333 1 nm	$(\pm 2.5 \times 10^{-10})$	26-0 R(127) a ₉

The uncertainties quoted for the wavelengths are the “estimated relative standard uncertainties”, which are similar to 1σ values.

The following sources of uncertainty have been identified:

RED			
Uncertainty of primary standard frequency	R 3	$\pm 2.5 \times 10^{-11}$	$\pm 2.5 \times 10^{-11} L$
Accuracy of calibration	R 3	$\pm 1 \times 10^{-9}$	$\pm 1 \times 10^{-9} L$
Variability (short-term) in stabilised test laser	R 3	$\pm 1.6 \times 10^{-9}$	$\pm 1.6 \times 10^{-9} L$
GREEN			
Uncertainty of primary standard frequency	R 3	$\pm 2.5 \times 10^{-10}$	$\pm 2.5 \times 10^{-10} L$
Accuracy of calibration	R 3	$\pm 1 \times 10^{-9}$	$\pm 1 \times 10^{-9} L$
Variability (short-term) in stabilised test laser	R 3	$\pm 9 \times 10^{-9}$	$\pm 9 \times 10^{-9} L$
ORANGE			
Uncertainty of primary standard frequency	R 3	$\pm 3 \times 10^{-10}$	$\pm 3 \times 10^{-10} L$
Accuracy of calibration	R 3	$\pm 1 \times 10^{-9}$	$\pm 1 \times 10^{-9} L$
Variability (short-term) in stabilised test laser	R 3	$\pm 3.3 \times 10^{-9}$	$\pm 3.3 \times 10^{-9} L$
RED WAVELENGTH TOTAL	R 3	$\pm 1.89 \times 10^{-9}$	$\pm 1.89 \times 10^{-9} L$

10.4.6 Mechanical - optical effects

No correction is made for the thickness of the wringing film since it is included in the definition of the length of the bar when measured interferometrically. However its variability can lead to a measurement uncertainty.

The following sources of uncertainty have been identified:

Effect of the source size (see § 4.1.3)	S 1	+ 4 µm diameter	- 4 x 10 ⁻¹³ L
Source off axis (see § 4.1.3)	S 2	+ 50 µm	- 5.6 x 10 ⁻¹⁰ L
	R 2	± 50 µm	± 5.6 x 10 ⁻¹⁰ L
Chromatic aberration - focal length error	S 1	+ 0.47 mm	- 4.4 x 10 ⁻¹¹ L
Laser beam diffraction	S 1	+ 80 mm diameter	- 2 x 10 ⁻¹¹ L
Spherical aberration in collimation	S 1	- 1 x 10 ⁻⁹ L	- 1 x 10 ⁻⁹ L
Spherical aberration in de-collimator	S 1	- 1 x 10 ⁻⁹ L	- 1 x 10 ⁻⁹ L
Prismatic tilt at beamsplitter	S 1	+ 4.5 fringes	- 5.1 x 10 ⁻¹⁰ L
Bar - beam alignment	R 3	± 2 fringes tilt	± 1.62 x 10 ⁻⁹ L
Shortening due to support points	S 1	bar slope < 8 x 10 ⁻⁶	- 6.4 x 10 ⁻¹¹ L
Reference beam alignment	R 3	± 60 µm off axis	± 8.0 x 10 ⁻¹⁰ L
Phase difference, dispersion and surface roughness difference	S 2	-14 nm	-14 nm
	R 2	± 27 nm	± 27 nm
Wringing film thickness	R 3	± 5 nm	± 5 nm
Accuracy of fringe fraction result and data analysis	R 3	± 0.016 fringe	± 5 nm
	R 3	TOTAL	± 28 nm ± 1.89 x 10 ⁻⁹ L
	S 1	TOTAL	-14 nm - 3.25 x 10 ⁻⁹ L

10.4.7 Bar expansivity at 20 °C

Because it is not possible to make all measurements at exactly 20 °C, the measured length of the bar is corrected to 20 °C. This requires both a measurement of the temperature of the bar and also an estimate of its coefficient of thermal expansion. The temperature is measured using two PRTs which are in small copper blocks in thermal contact with the bar. The temperature of these PRTs is measured using a resistance bridge. The bridge is calibrated monthly by using it to measure an external 100 Ω standard resistor. The PRTs are calibrated at 2-yearly intervals and are checked every 6 months against a water triple point cell. The nominal coefficient of thermal expansion used for length bars (and also for gauge blocks over 100 mm in length) is $\alpha = 10.7 \text{ ppm K}^{-1}$. Variation in the value of α from bar to bar is estimated to be within $\pm 0.5 \text{ ppm K}^{-1}$. The temperature of the bar inside the chamber is $20 \text{ }^\circ\text{C} \pm 0.03 \text{ }^\circ\text{C}$.

The following sources of uncertainty have been identified:

Resolution of resistance bridge	R 3	$\pm 10 \mu\Omega = \pm 0.03 \text{ mK}$	$\pm 3.21 \times 10^{-10} L$
Resistance bridge accuracy	R 3	$1 \text{ ppm} \pm 10 \mu\Omega = \pm 101 \mu\Omega = \pm 0.3 \text{ mK}$	$\pm 3.21 \times 10^{-9} L$
Accuracy of standard resistor <u>PRT calibration</u>	R 3	$\pm 8 \mu\Omega = \pm 0.024 \text{ mK}$	$\pm 2.57 \times 10^{-10} L$
Water triple point accuracy	R 3	$\pm 0.5 \text{ mK}$	$\pm 5.35 \times 10^{-9} L$
Gallium melting point accuracy	R 3	$\pm 0.5 \text{ mK}$	$\pm 5.35 \times 10^{-9} L$
Interpolating equations	R 3	$\pm 0.13 \text{ mK}$	$\pm 1.39 \times 10^{-9} L$
Drift between calibrations	R 3	$< \pm 0.5 \text{ mK}$	$\pm 5.35 \times 10^{-9} L$
Contact of PRT with bar	R 3	$\pm 0.5 \text{ mK}$	$\pm 5.35 \times 10^{-9} L$
Non-linear gradient at 20 °C	R 3	$\pm 0.1 \text{ mK}$	$\pm 1.07 \times 10^{-9} L$
TOTAL		$\pm 1.05 \text{ mK}$	$\pm 1.13 \times 10^{-8} L$
Accuracy of nominal α	R 3	$\pm 0.5 \text{ ppm K}^{-1} (@ 20.03 \text{ }^\circ\text{C})$	$\pm 1.5 \times 10^{-8} L$

10.5 SUMMATION OF UNCERTAINTIES

In accordance with the guidelines, the random and systematic (class 3 and class 1) uncertainties are summed individually. The length dependent and length independent contributions are also treated separately. There are thus four separate error contributions:

- S systematic, length independent
- S_L systematic, length dependent
- R random, length independent
- R_L random, length dependent

The contributions to S and S_L are summed arithmetically, whereas the contributions to R and R_L are summed in quadrature. The random (class 3) uncertainties are then multiplied by a factor of 1.96 to obtain results at a confidence level of 95%. The final totals are:

$$\begin{aligned} S &= -14 \text{ nm} \\ S_L &= +1.15 \times 10^{-8} L \\ R &= \pm 28 \text{ nm} \\ R_L &= \pm 6.22 \times 10^{-8} L \end{aligned}$$

where L is the length of the bar in metres.

Thus a full uncertainty statement for the interferometer is

**The central length measurement uncertainty for the
Primary Length Bar Interferometer is
 $-14 \text{ nm} \pm 28 \text{ nm} + 1.15 \times 10^{-8} L \pm 6.22 \times 10^{-8} L$
at a confidence level of 95%, for a bar of length L metres.**

Depending on how the errors are combined, it is possible to obtain different estimates of the error for a particular measured length.

Firstly, the maximum and minimum possible values can be calculated as per the guidelines: $(S + S_L + R + R_L)$ and $(S + S_L - R - R_L)$ respectively. This will be referred to as the RECOMMENDED uncertainty estimate.

Secondly, the quadrature sum of the random uncertainties can be either added or subtracted from the systematic error total: $(S + S_L \pm \sqrt{R^2 + R_L^2})$. This will be referred to as the STANDARD uncertainty estimate.

The final method of combining the errors is that recommended by the BIPM where the systematic errors are combined in quadrature with the random errors to produce two figures, one length dependent, the other length independent, which are then added in quadrature: $S^2 + S_L^2 + R^2 + R_L^2$. This results in a figure of $\pm 30 \text{ nm} \pm 64 L \text{ nm}$. For comparison, the NPL Length Bar Machine has an uncertainty of length measurement of $\pm 68 \pm 350 L \text{ nm}$. This will be referred to as the BIPM uncertainty estimate (this is the most common technique of quoting uncertainties for metrological purposes).

These different combinations are plotted in figure 10.2.

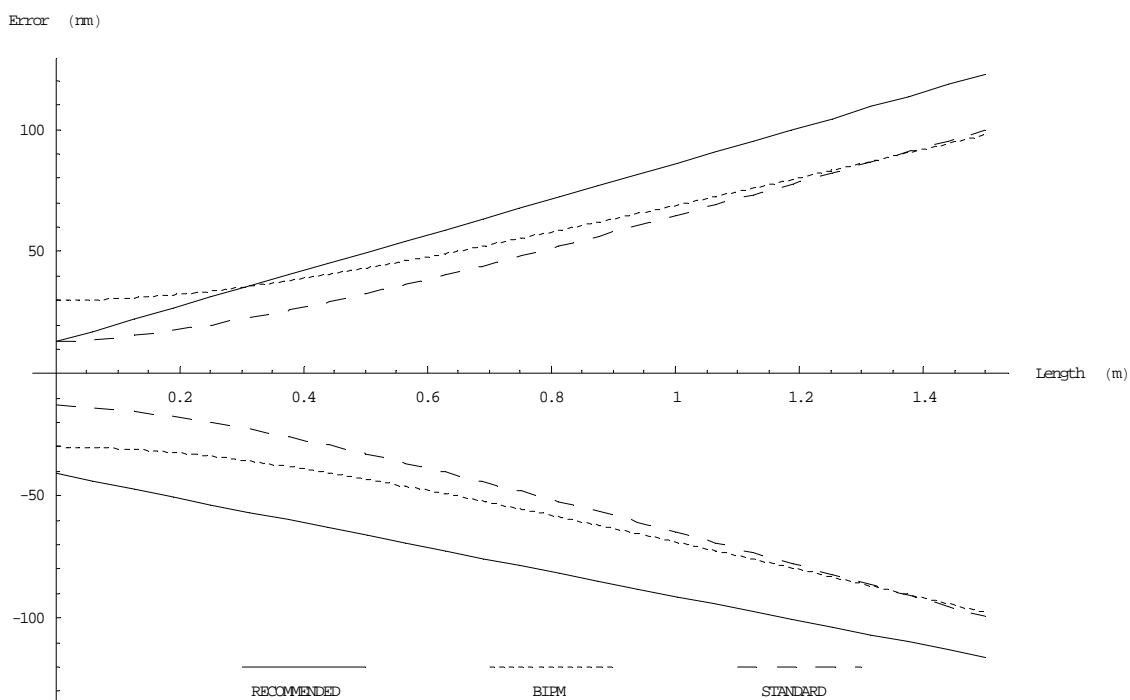


Figure 10.2 - Plot of total uncertainty in length measurement over length range 0.1 - 1.5 m

The differences between the techniques are due to whether they sum the components in quadrature (sign symmetric) or arithmetically (sign asymmetric). The difference between the three techniques is approximately 20 nm, though this depends on the length of the bar. Except for bars of length 300 mm and below, the RECOMMENDED uncertainty is larger than the other techniques and is thus more ‘safe’ to quote if a simple analysis is required. The BIPM and STANDARD estimates are in good agreement for longer bars. Thus the importance of quoting the result in the most comprehensive form, where all the terms are listed, can be seen.

10.6 POSSIBLE STEPS TO IMPROVE THE ACCURACY

This accuracy can be improved significantly by reducing the uncertainty associated with the thermal expansion coefficient of the bar. As detailed in chapter 8, the interferometer was also designed to measure the coefficient of linear thermal expansion of length bars (and long gauge blocks). The contribution of the uncertainty in thermal expansion coefficient is $\pm 1.5 \times 10^{-8} L$ for an uncertainty of $\pm 0.5 \times 10^{-6} \text{ K}^{-1}$ in α . From § 8.6 it is seen that by measuring the expansion coefficient in the interferometer, this can be reduced to an uncertainty of between ± 0.2 and $\pm 0.05 \times 10^{-6} \text{ K}^{-1}$, which corresponds to a length measurement uncertainty of between $\pm 6 \times 10^{-9} L$ and $\pm 1.5 \times 10^{-9} L$.

10.7 COMBINED UNCERTAINTY BUDGETS OF INSTRUMENTS

As stated in § 9.4 the differences between the measurements of set 1455 in the LBM and the LBI all fall within the uncertainty budget of the LBM alone, except for the 300 mm bar which has been explained. It was thus not necessary to consider the combination of the uncertainty budgets of the two instruments. For reference, this will now be discussed briefly. When comparing two results from different instruments it should be remembered that the results are given as single values with confidence limits. To a good approximation, the errors of the two instruments are randomly distributed and can be combined statistically. Standard statistical tests [11] can be used to ascertain a confidence level for whether or not the two sets of results share a common overlap of any statistical significance. In the case of the results given in chapter 9, the differences between the two instruments' results are all within the 95% confidence limits of the LBM uncertainty budget alone, and so there is 95% confidence that the results agree, within the stated uncertainties of the instruments.

[2002 re-release note: Since the thesis was completed, the *Guide to the Expression of Uncertainty in Measurement (GUM)*, published by ISO, has become the *de facto* standard for uncertainty budget preparation. The style set out in the GUM is quite different to that presented in this thesis.]

REFERENCES FOR CHAPTER 10

- [1] Colclough A R Two theories of experimental error *J. Res. NBS* **92** (1987) 167-185 (and references therein)
- [2] NAMAS document **NIS 3003** Edition 7 (1991) *The expression of uncertainty and confidence in measurements*
- [3] WECC Draft **19-1990** (1990) *Guidelines for the expression of the uncertainty of measurement in calibrations*
- [4] Campion P J, Burns J E, Williams A *A code of practice for the detailed statement of accuracy* (HMSO London) (1980) 23
- [5] Kaye G W C & Laby T H *Tables of Physical and Chemical Constants* 15th edn (1989) (Harlow, Essex: Longmans) 456-462
- [6] Fisher R A *Statistical Methods for Research Workers* (1958) Oliver & Boyd
- [7] Giacomo P News from the BIPM *Metrologia* **17** (1981) 69-74
- [8] Rusby R L, Hudson R P & Durieux M, Schooley J F, Steur P P M & Swenson C A Thermodynamic basis of the ITS-90 *Metrologia* **28** (1991) 9-18
- [9] British Standard BS 1339 (1965) (London: British Standards Institution)
- [10] Gibbins C J A survey and comparison of relationships for the determination of the saturated vapour pressure over plane surfaces of pure water and of pure ice *Annales Geophysicae* **8** (1990) 859-886, table 1
- [11] Kaye G W C & Laby T H *Tables of Physical and Chemical Constants* 15th edn (1989) (Harlow, Essex: Longmans) 456-462

CHAPTER 11

CONCLUSIONS

“I never think of the future, it comes soon enough”
A Einstein

11.1 CONCLUSIONS

- 1 The first conclusion must be that the instrument works, and meets its design criteria, namely to offer: improved accuracy calibrations of length bars over 100 mm with an easy to use interface, measurement of flatness and variation of measurement faces, and measurement of thermal expansion.
- 2 The spread in the measurements and deviations from the results of other instruments are within the uncertainty budgets of the instruments concerned.
- 3 The extra option for double-ended measurement offers potential for measurement of length bars without wringing.
- 4 The automation of the instrument not only reduces the required skill level of potential operators, but has the advantages of an objective measurement compared to the subjective measurement of instruments requiring manual operation.
- 5 It is therefore hoped that this instrument may be commercialised. This will complete the cycle of development of measurement techniques noted in § 1.3.7, bringing a new level of accuracy to commercial measurements and initiating the search for the next level of measurement accuracy which will be required from national standards laboratories in the future.
- 6 A particularly useful feature of the instrument is the ability to measure not only length bars, but also gauge blocks, allowing comparison with other instruments such as the Gauge Block Interferometer and the Gauge Block Dilatometer which is being developed.

- 7 Apart from the lower accuracy measurements of bar diameter and straightness which are measured using calibrated micrometers, the interferometer can be used to measure all the dimensional aspects of length bars, according to the relevant standards.
 - 8 A single measurement on the interferometer takes under 2 minutes to complete, and gives not only length measurement, but also flatness and parallelism. This compares favourably with the 30 minutes to perform a measurement on the Kösters-Zeiss (length of one bar per loading), or approximately 3 minutes for the NPL Length Bar Machine (central length only, 5 bars per loading, with larger uncertainty).
 - 9 Thermal expansion measurements can be performed in one week, with measurements at 5 temperatures over the range 20 °C - 30 °C.
 - 10 The ability to leave the interferometer running, making repeated measurements gives increased confidence in the results obtained, because of the small spread in the results.
 - 11 One important factor throughout the whole of the interferometer design has been the small source size, and accuracy of placing the source on the axis of the interferometer. The alignment inaccuracy is one of the drawbacks of the Length Bar Machine, which has a smaller laser beam which is not so easy to align, since no large field interference pattern can be viewed, nor a return spot smaller than 2 - 3 mm diameter. The small source size of the interferometer is also important for the coherence exhibited by the interferometer, particularly in double-ended mode, and for the good fringe contrast at large path differences and the good depth of focus.
 - 12 This work has shown that it is possible to combine multiple-wavelength interferometry with phase-stepping interferometry over long path lengths to achieve accurate length measurement.
 - 13 As shown in chapter 10, much of the measurement uncertainty is due to factors other than the multiple-wavelength phase-stepping interferometry which is a powerful technique which is theoretically capable of high accuracy measurement. With careful control and accurately known chosen wavelengths, it should be possible to build a three-wavelength interferometer with a multiple-wavelength repeat distance of up to 0.5 mm using the three wavelengths used in this work, accurate refractive index determination and fringe fraction measurement of better than ± 0.015 fringes over long path lengths.
-

11.2 PROPOSALS FOR IMPROVEMENTS

The measurement or calculation of refractive index is of vital importance to measurements of length made in the interferometer. There is no doubt that by operating in a vacuum, the problem of refractive index determination would be removed. A vacuum though poses other problems (as mentioned in chapters 3 & 7) and may not prove more accurate overall. Perhaps the best technique would be to incorporate a refractometer inside the interferometer, either simply inside the chamber, or directly in the measurement beam (as is the case in the Kösters-Zeiss). This would require further piping to allow evacuation of the refractometer cell inside the chamber, and possible reduction in the beam area available for measurements. One way of making sure of having a cell of well-determined length would be to use accurately measured length bars as dimensional structures in the refractometer, with accurately measured expansion coefficients. Rather than use quadrature fringe counting, which is prone to offset and gain errors, the fringe order of the refractometer could be determined approximately by Edlén calculations, and the fractional fringe order measured by phase-stepping interferometry.

For the ultimate accuracy, the Zeeman stabilised lasers could be replaced directly with iodine-stabilised lasers, with much better frequency stability, allowing a larger uncertainty in the initial estimate of the length of the bar. Adding another wavelength would also increase the allowable initial length uncertainty by extending the range of the multiple-wavelength technique. Alternatively, selecting a different wavelength may increase the range of the three-wavelength technique, though the laser would still have to be frequency-stabilised.

On a professional note, some of the optical mounts used in the interferometer are not fully kinematic. Before the instrument is commercialised it would be useful to design some proper kinematic mounts for the optics such as the beamsplitter and the collimator mirror which have repeatable positioning and are not over-constrained.

The choice of computer system was rather limited at the time of purchase, since only the IBM PC-compatible market had interface cards that were needed for the instrumentation, and a seemingly-suitable language for the programming. Given the choice at the present time, perhaps a more powerful computer would be chosen, either a Hewlett-Packard workstation or a top range Macintosh, such as a Quadra. These machines have no imposed 640 K memory restrictions and have very good programming languages and interfaces. For programming language, any well-structured, easy to read language will suffice - Pascal, C, or one of the better

implementations of BASIC, such as HP Basic, so long as it can be used to drive the interfaces and access large arrays.

On the subject of thermal control, recent work at NPL on the Gauge Block Dilatometer has shown that individually controlled heating panels, using electric resistive heating can be used to control a small volume of air very uniformly. Perhaps this technique could be expanded to the scale of the PLBI thus decreasing temperature gradients even further. A disadvantage is the lack of cooling, so this would require a room at a temperature below 20 °C in order to be able to make measurements at temperatures below 20 °C or to stabilise under active control at 20 °C.

It would be useful to include more PRTs inside the chamber, particularly when measuring thermal expansions, to check temperature gradients along the bars. The data could enable finer adjustments to be made to the heating system (particularly in the case of individual panels, as described above) to decrease thermal gradients in the bars. This would require further channels on the Tinsley resistance bridge selector switch and longer measurement time, though once proper temperature uniformity had been achieved, only one or two PRTs would be needed for the measurement of temperature during a ‘real’ measurement.

Another possibility, which was considered at the design stage, would be to fill the chamber with a gas other than air. Nitrogen or helium are good candidates as they are both inert. Helium offers two advantages: it is less dense than air or nitrogen so its refractive index is less sensitive to pressure changes and it has approximately 6 times the thermal conductivity of air, which would result in lower temperature gradients and decreased stabilisation times. However accurate knowledge of the refractive index of helium would be required and it is difficult to use with many pressure transducers because its density is different to air (some transducers work by measuring density) and it would leak into any reference vacuum compartments found inside other transducers. Venting a chamber of helium to atmosphere after an experiment would be interesting for any observers in the room at the time!

In § 9.8 it was stated that the accuracy of the fringe fraction measurement is limited by the data fitting of the phase data on the platen surface in the case where the platen is not flat. This can be improved in two ways. Firstly, a 2nd order polynomial surface fitting should give a better result than the best fit plane and Chebychev techniques. Secondly, platens with flatter surfaces would allow more accurate data fitting, although there is the limitation that when bars are wrung to the platens, the wringing forces can distort the platen surface in attempting to reach closer contact with the bar. These effects have been observed when wringing gauge blocks, at PTB, Germany.

Although this thesis has presented some work on the use of double-ended interferometry, it must be remembered that this was designed as an add-on to the main interferometer and so could be improved. For instance, the use of a compensator plate would be of a great advantage to both the double-ended work and the mainstream use of the interferometer by removing the dispersive effects of the wedged beamsplitter and allowing almost perfect alignment of the reference beam to be maintained for different wavelengths. The collimated beam diameter was designed to be sufficient for single-ended measurement but is not really large enough for double-ended use. The double-ended analysis requires more data in the background region for more accurate data fitting. At present there is only a limited amount of data and the results are inaccurate. Increased spatial coherence is also needed to improve the fringe contrast in the background area (and on the rear face of the bar) to further decrease fringe fraction measurement errors. With better data, proper results for flatness and parallelism can be calculated.

Least-squares planes would be fitted to the phase data for each of the faces, to determine individual flatness and variation results. The images of the bar faces are both horizontally inverted in the camera image, *i.e.* each is the view one would obtain by eye, looking at each end of the bar separately, but flipped left-to-right. However, to compare the results so that mutual parallelism or variation may be measured, one phase map must be horizontally inverted. The parallelism would then be obtained as the sum of the two phase maps - the parallelism of the rear face with respect to the front face is already inverted because of the double-ended configuration, so the phase maps would be added rather than subtracted, to get mutual parallelism.

11.3 ACKNOWLEDGEMENTS

Firstly I would like to thank the Director of the National Physical Laboratory (where this work was performed) for allowing me to submit this work for my Doctorate thesis and for letting me study at Imperial College.

I am very grateful to my supervisors for their precious time and effort in seeing this work through to its conclusion. At NPL: David Pugh for his enthusiasm, knowledge and interest in this work; Keith Jackson for his knowledge, experience and humour. At Imperial College: Jonathan Maxwell for his time, observations and comments.

I would also like to acknowledge the craftsmen, designers and engineers of NPL's Engineering Services and Optical Workshop for turning my (somewhat unconventional) designs into reality - much of what the interferometer contains has been manufactured by them.

I am grateful to the staff of End Standards Section, NPL, for putting up with me 'going virtual' on Friday afternoons (and other times !) and laying claim to the Macintosh for hours at a time. I am also grateful to Ben Hughes and Richard Leach for the many discussions on aspects of this work and for the encouragement and enthusiasm of other colleagues at NPL. Thanks also to Dr Graham Peggs for the use of his scanner.

I would also like to put in a brief thank you to my friends from Cambridge and Leicester for their help and advice during my 3 years as a 'real' student.

Finally, the biggest thank you, for everything else, to my parents.

“Original ideas are exceedingly rare and the most that philosophers have done in the course of time is to erect a new combination of them”

G Sarton

APPENDIX A

OPTO-MECHANICAL EQUIPMENT LIST

Optical table

Newport Scientific series, custom made, 1.2 m x 2.4 m, custom sealed hole pattern, non-magnetic 300-series stainless steel surface, custom lapped for sealing against o-ring. Damped construction, compliance $< 3 \times 10^{-5}$ mm N $^{-1}$ for f > 90 Hz, and $< 3 \times 10^{-4}$ mm N $^{-1}$ for f > 20 Hz. Impulse decay time approximately 50 ms.

Vibration isolation: 4 x XL-A pneumatic legs

Optics baseplate

Material - PERALUMAN (similar to AA 5083 aluminium) from Swiss Aluminium Ltd, precision rolled, thickness tolerance 0.1 mm over surface, single sheet. Thermal expansivity 24×10^{-6} K $^{-1}$, RA 0.4 μm , density 2660 kg m $^{-3}$, Youngs modulus 70 000 N mm $^{-2}$, thermal conductivity 1.2 W cm $^{-1}$ K $^{-1}$.

Dimensions: 2020 x 800 x 20 mm

Optical fibres

3M EOTEC speciality fibres - single mode for visible wavelengths

FS-VS-2211 (543 nm) - 2.8 μm core, 80 μm fibre, 200 μm coating, 0.12 NA

FS-SN-3221 (633, 612 nm) - 3.6 μm core, 125 μm fibre, 250 μm coating, 0.12 NA

Custom PTFE sheathing, 6 mm brass ferules manufactured by NPL Optical Workshop

Optical fibre z-motion stage

Photon Control TS 75-25H, micrometer drive, 3 μm sensitivity, 25 mm range

Optical Fibre X-Y stage

Micro Controle SB 18 YZ 133 202, micrometer drive, 0.1 μm sensitivity, ± 2 mm range

Collimator Lens

Spindler & Hoyer 32 2315 achromatic doublet, f 1500 mm, diameter 100 mm, AR coated centred and corrected for $\lambda = 656$ nm, 486 nm, measured spherical aberration 0.15 λ

Collimator lens mount

Custom made, machined by NPL Engineering Services

Cast iron 'L'-section, WDS tooling aids, L section 913-407, 200x70x200 mm - 150

Collimator mirror

Custom made, ZERODUR flat, protected Al overcoat, manufactured by NPL Optical Workshop - 150 mm diameter, 28 mm thickness, measured wavefront aberration 0.05λ

Collimator mirror mount

Custom made, cast iron ‘L’-section, machined by NPL Engineering Services from WDS Tooling Aids, L section 913-407, 200x70x200 mm - 150

Beamsplitter

Custom made, fused silica, 70 mm x 120 mm x 20 mm, 0.5° wedge along major axis coating 50/50 transmission/reflection, broadband dielectric, wavefront aberration $\lambda/40$
Manufactured by NPL Optical Workshop

Beamsplitter mount

Custom made, cast iron ‘L’-section, machined by NPL Engineering Services from WDS Tooling Aids, L section 913-407, 200x70x200 mm - 150

Reference mirror

Custom made, fused silica, manufactured by NPL Optical Workshop - diameter 90 mm, 20 mm thick, overcoated with 100 nm chromium, measured wavefront aberration 0.02λ

Reference mirror mount

Custom made, aluminium alloy, manufactured by NPL Engineering Services

Reference mirror PZT driver

Queensgate Instruments AX100, Capacitive feedback digital positioning translator (DPT), Resolution < 1 nm, range 10 μm

Reference arm adjustable mirror

Newport 40D20AL.2 Borosilicate, 101.6 mm diameter (4 inch), 18 mm thickness, overcoated with protected aluminium, wavefront aberration $\lambda/20$

Reference arm adjustable mount

Photon control GM100D gimbal mount - 0.25 arcsec resolution

Reference arm PZT adjusters

Physik Instrumente P854.00 Piezomikes - Manual : ± 6 mm range, PZT: $\pm 30 \mu\text{m}$, -10 to +120 V

Measurement arm adjustable mirror

Newport 40D20AL.2 Borosilicate, 101.6 mm diameter (4 inch), 18 mm thickness, overcoated with protected aluminium, wavefront aberration $\lambda/20$

Measurement arm adjustable mount

Oriel 17741 precision mirror mount, 4", 0.05 arcsec resolution with NPL-designed 2-axis flexure mount, manufactured by NPL Engineering Services

Measurement arm PZT adjusters

Burleigh PZ40, 15 μm travel, 1000 V DC

De-collimator mirror

Comar 160 ME 100, 160 mm x 100 mm x 10 mm, enhanced reflectivity

De-collimator mirror mount

Custom made, aluminium alloy, WDS Tooling Aids - L section 913-406 160x160x70 mm - 150

Machined by NPL Engineering Services

De-collimating lens

Spindler & Hoyer 32 2313 achromatic doublet - f 1000 mm, diameter 100 mm, AR coated, centred and corrected for $\lambda = 656 \text{ nm}$, 486 nm , measured spherical aberration 0.1λ

De-collimator lens mount

Custom made, machined by NPL Engineering Services

Cast iron 'L'-section, WDS tooling aids, L section 913-407, 200x70x200 mm - 150

45° mirror

Daedel 2850 $\lambda/4$, 25 x 35 x 6 mm

Imaging lens

Comar 200 DQ 32, aplanatic doublet, f 200 mm, 32 mm diameter, AR coated

Imaging lens mount

Custom lens holder in X-Y stage

Chamber window

Comar 63 GH 60, Borosilicate, multi-layer AR coating

CCD camera

Sony AVC-D5CE with CMA-D5/D5CE PSU

Length bar support carriage

Material - PERALUMAN (similar to AA 5083 aluminium) from Swiss Aluminium Ltd, precision rolled, thickness tolerance 0.1 mm over surface, single sheet. Thermal expansivity $24 \times 10^{-6} \text{ K}^{-1}$, RA 0.4 μm , density 2660 kg m^{-3} , Youngs modulus $70\,000 \text{ N mm}^{-2}$, thermal conductivity $1.2 \text{ W cm}^{-1} \text{ K}^{-1}$.

Dimensions: 1500 x 260 x 20 mm

Length bar supports

Custom made in TUFNOL with brass inserts for PRTs

Martock Design MD255 slides with MD208 mini adjusters

Platens for length bars

Custom made by Tesa Reference Standards, Leicester

High quality tool steel, 75 mm diameter, 15 mm thick, precision ground and lapped to $\lambda/20$, surface finish similar to length bars and gauge blocks

APPENDIX B

QUALITY OF OPTICAL SURFACES

Manufacturing tolerances and imperfections in materials and polishing will lead to errors in the surface shapes and homogeneity of the optical components of the interferometer. This will lead to wavefront distortions, the effects of which will vary with the path length difference in the interferometer. At zero path difference only non-collimation errors will be seen, *e.g.* flatness errors of the reference mirror, inhomogeneity of the beamsplitter refractive index. At non-zero path differences, collimation errors will also be seen (even with perfectly flat reference mirror and homogeneous beamsplitter). These collimation errors include collimator spherical aberration and astigmatism, astigmatism in the collimator mirror and any defocus of the collimator.

The optics of the interferometer have been individually checked on a ZYGO Mk 4 interferometer - this is a phase-shifting Fizeau interferometer operating at $\lambda = 633$ nm, which has had its reference surface calibrated using absolute flatness techniques [1,2]. This reference surface shows a P-V variation of 0.02λ .

Astigmatism in mirrors

A mirror, used at an oblique angle, will contribute a wavefront astigmatism of maximum amplitude $A\lambda$ if its surface has a minimum radius of curvature, R_{min} given by

$$R_{min} = \frac{nD^2 \cos \theta_i}{4A\lambda} \left(\frac{1}{\cos^2 \theta_i} - 1 \right) \quad (\text{B.1})$$

where n is the refractive index of the surrounding medium (air), θ is the angle of incidence and D is the diameter of the beam.

Collimator mirror

The beam diameter is 80 mm, the angle of incidence 60° . Thus at a wavelength $\lambda = 633$ nm, R_{min} is $3791/A$ metres. The collimator mirror was found to be flat to 0.06λ

over its 150 mm diameter. By simple trigonometry, this is equivalent to a radius of curvature of approximately 6×10^5 metres, giving a value of the astigmatism of $A = 1/160$, *i.e.* $\lambda/160$ wavefront astigmatism.

Reference mirror

The angle of incidence at the reference mirrors is 0° (straight on). The reference mirror was found to be flat to 0.015λ and was smoothly varying.

Path folding mirrors

In the reference arm, the angle of incidence is 45° leading to $R_{min} = 1787/A$, and in the measurement arm the angle is 30° , leading to $R_{min} = 730/A$. These two mirrors are flat to 0.025λ leading to astigmatism of $\lambda/139$ in the measurement arm and $\lambda/57$ in the reference arm.

Roof mirrors

Roof mirror 1 was found to be flat to 0.025λ and roof mirror 2 was found to be flat to 0.03λ , both with slight roll-off at the edges. The angle of incidence at both mirrors is 45° , leading to $R_{min} = 1787/A$. Roof mirror 1 thus contributes $\lambda/57$ of astigmatism and roof mirror 2 contributes $\lambda/48$ of astigmatism.

Collimator lens

The collimator lens was tested in a double-pass arrangement and found to have a double pass P-V wavefront distortion of 0.3λ , giving a single pass distortion of less than 0.15λ . This is mostly spherical aberration and its effect on measured length was shown in § 4.1.7.2 to be of magnitude $1 \times 10^{-9} L$.

De-collimating lens

The de-collimating lens was checked in the same manner as the collimating lens and found to have a P-V wavefront distortion of 0.1λ , which was mostly spherical aberration. The effect on measured length is of magnitude $1 \times 10^{-9} L$.

REFERENCES FOR APPENDIX B

- [1] Schultz G & Grzanna J Absolute flatness testing by the rotation method with optimal measuring-error compensation *Appl. Opt.* **31** (1992) 3767-3780
 - [2] Schwider J Fizeau and Michelson type interferograms and their relation to the absolute testing of optical surfaces *Optik* **89** (1992) 113-117
-

APPENDIX C

FLEXING OF LENGTH BARS

C.1 FLEXING OF A LENGTH BAR DUE TO ITS OWN WEIGHT

Any object lying in a horizontal plane will sag under its own weight unless it is infinitely stiff or is supported at many points along its length. For length bars this causes two problems. Firstly, if there is any sagging in the vicinity of the ends of the bar, this will cause the two end faces to tilt with respect to one another causing a bar with otherwise parallel faces to appear out of parallel. Secondly, since the material of the bar no longer lies in a straight line between the two end faces, the extra bending may cause the length of the bar, measured as the separation between the end faces, to become shorter than in its free state.

One solution is to measure the bars vertically, though this is not possible because of three reasons. Firstly, the relevant standards state that the bars should be measured in a horizontal plane, supported at two points termed the “Airy points” (see later), since this is how they will be used in practice. Secondly, a bar standing vertically will contract under its own weight, see Appendix D. Thirdly, the variation of refractive index between the top and bottom of the bar due to (i) the air pressure gradient due to the Earth’s gravitational field and (ii) the variation in the air temperature, contributes a significant measurement uncertainty.

Historical solutions such as floating the bar in mercury or supporting it on a system of 8 rollers or supports [1] have been rejected as hazardous or impractical. They also do not conform to the relevant standards. The chosen solution is to support the bar on two points whose positions are chosen to make the ends of the bar vertical and parallel with each other. These are termed the “Airy points” of the bar and their positions are usually engraved on the bar’s surface. The position of these points will now be derived.

C.2 DERIVATION OF POSITIONS OF AIRY POINTS

Consider a uniform solid bar of length L , cross-sectional moment of inertia I , and total weight W . This bar is supported at 2 points, symmetrically placed about its middle, separated by a distance S . Let the reactions at the two supports be R_1 & R_2 as shown in figure C.1.

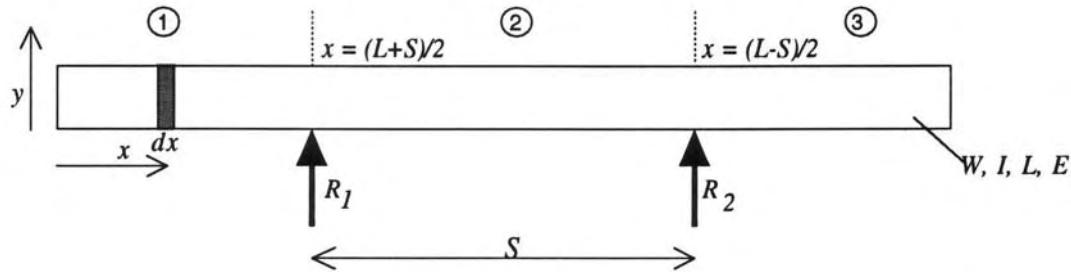


Figure C.1 - Bar supported at two points

Resolving vertically,

$$R_1 + R_2 = W, \quad R_1 = R_2$$

$$\therefore R_1 = R_2 = \frac{W}{2}$$

Now, split the bar into three sections (1) to (3) as shown in figure C.1, for the following analysis. In each section,

$$\sum \text{bending moments} = EI \frac{d^2y}{dx^2} \quad (\text{Bernoulli-Euler theory})$$

Since the bar is uniform, EI is a constant, and as such will be removed from the following equations for simplicity.

In section (1)

$$\frac{d^2y}{dx^2} = \frac{Wx^2}{2L} \quad (\text{C.1})$$

In section (2)

$$\frac{d^2y}{dx^2} = \frac{Wx^2}{2L} - R_1 \left[x - \frac{(L-S)}{2} \right] \quad (\text{C.2})$$

In section (3)

$$\frac{d^2y}{dx^2} = \frac{Wx^2}{2L} - R_1 \left[x - \frac{(L-S)}{2} \right] - R_2 \left[x - \frac{(L+S)}{2} \right] \quad (\text{C.3})$$

Integrating equations (C.1) (C.2) and (C.3) gives, respectively,

$$\frac{dy}{dx} = \frac{Wx^3}{6L} + C_1 \quad (\text{C.4})$$

$$\frac{dy}{dx} = \frac{Wx^3}{6L} - R_1 \left[\frac{x^2}{2} - \frac{(L-S)x}{2} \right] + C_2 \quad (\text{C.5})$$

$$\frac{dy}{dx} = \frac{Wx^3}{6L} - R_1 \left[\frac{x^2}{2} - \frac{(L-S)x}{2} \right] - R_2 \left[\frac{x^2}{2} - \frac{(L+S)x}{2} \right] + C_3 \quad (\text{C.6})$$

The slope of the bar, $\frac{dy}{dx}$ must be continuous at the supports therefore equating (C.4) and (C.5), and substituting $x = \frac{L-S}{2}$ gives

$$C_1 = C_2 - R_1 \left[\frac{(L-S)^2}{8} - \frac{(L-S)^2}{4} \right]$$

$$ie \quad C_1 = C_2 + R_1 \left[\frac{(L-S)^2}{8} \right] \quad (\text{C.7})$$

One constraint is that we require vertical end faces, ie $\left. \frac{dy}{dx} \right|_{x=0} = 0$

This implies that $C_1 = 0$ Substituting this result into (C.7) and using the fact that $R_1 = \frac{W}{2}$ gives

$$C_2 = -\frac{W}{2} \left[\frac{(L-S)^2}{8} \right] \quad (\text{C.8})$$

Now, matching $\frac{dy}{dx}$ at $x = L$ gives

$$C_3 = -\frac{WL^2}{6} \quad (\text{C.9})$$

With C_1, C_2, C_3 determined, equations (C.3) (C.4) and (C.5) completely describe the bending of the bar, once S is known. To find S , $\frac{dy}{dx}$ is matched at the boundary between regions (2) and (3).

In region (2)

$$\left. \frac{dy}{dx} \right|_{x=\frac{L+S}{2}} = \frac{W(L+S)^3}{48L} - \frac{W}{2} \left[\frac{(L+S)^2}{8} - \frac{(L-S)(L+S)}{4} + \frac{(L-S)^2}{8} \right] \quad (\text{C.10})$$

and in region (3)

$$\left. \frac{dy}{dx} \right|_{x=\frac{L+S}{2}} = \frac{W(L+S)^3}{48L} - \frac{W}{2} \left[\frac{(L+S)^2}{8} - \frac{(L-S)(L+S)}{4} + \frac{(L+S)^2}{8} - \frac{(L+S)^2}{4} \right] - \frac{WL^2}{6}$$

(C.11)

Equating (10) and (11) gives

$$-\frac{W(L-S)^2}{16} = -\frac{W}{2} \left[-\frac{(L+S)^2}{8} \right] - \frac{WL^2}{6}$$

which with reduction gives

$$S^2 = \frac{L^2}{3}$$

i.e.

$$S = \frac{L}{\sqrt{3}}$$

This is the symmetrical spacing of the Airy points, i.e. approximately 0.577 of the length of the bar. This is only valid for a bar supported at two points with no additional reference flats or other masses attached to it. Even when a bar is supported at the Airy points, its central length will be different to the case where it is unsupported due to the extra curvature of the bar. Figure C.2 shows the difference dL in length between a bar which is unsupported and one which rests on supports positioned a distance a away from the end faces ($L - S = 2a$). Note that supporting at the Airy positions ($a = 0.211$) causes a change in length of $dL = -0.4$ nm, which is negligible. The support positions corresponding to $a = 0.185$ for which there is the minimum change in length are termed the Bessel points.

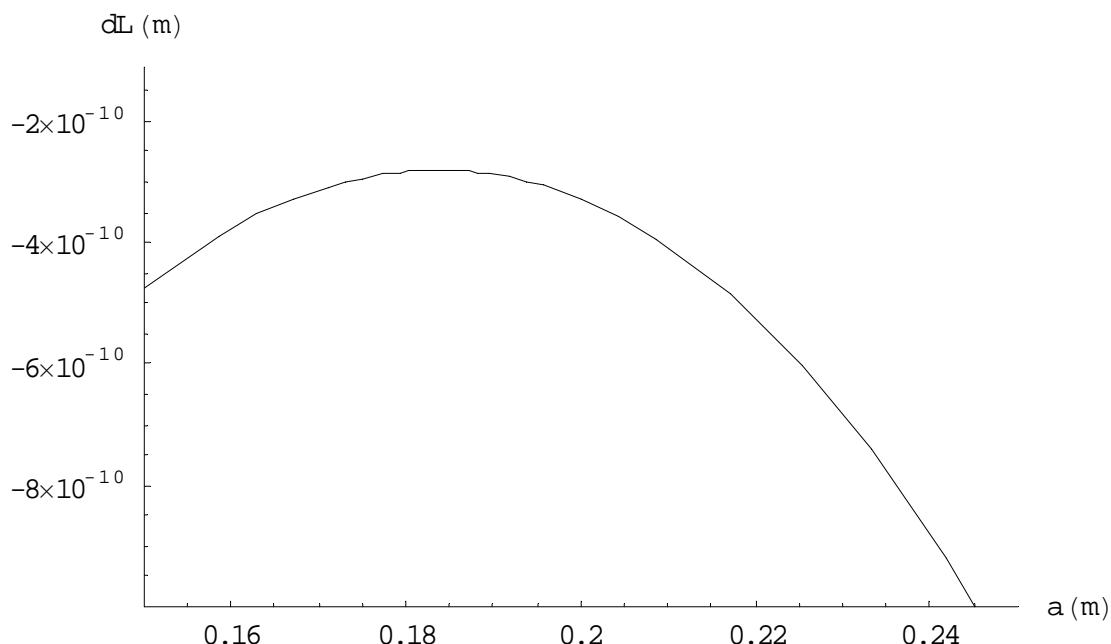


Figure C.2 - Effect of support point position, a , on change in length, dL , of bar from unsupported state for a 1 m bar.

C.3 COMPENSATION FOR MASS OF WRUNG FLAT

When a reference flat is wrung to one end face of a bar, this adds additional bending and will cause the bar supported at the Airy points to exhibit a parallelism error. Techniques for compensating for the extra mass of the flat include supplying an additional lifting force by means of weights or levers which effectively cancels out the weight of the flat [2] or by moving the support points towards the ends of the bar [3].

The latter solution has been adopted as being easier to implement and is detailed below.

Consider the bar and reference flat (platen) shown in figure C.3.

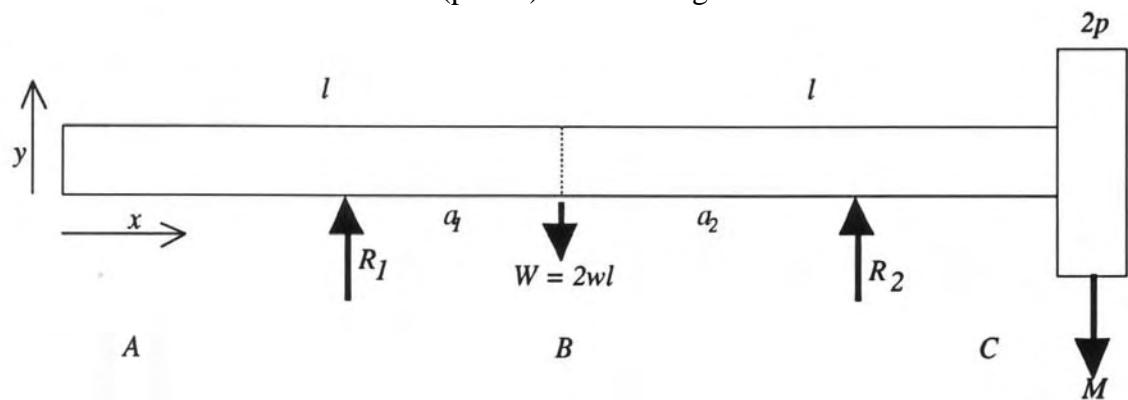


Figure C.3 - Bar supported at new support points with flat attached to one face

The supports are positioned at $x = l - a_1$ and $x = l + a_2$, with l being the half-length of the bar. As before, applying Bernoulli-Euler bending theory to the three regions gives three equations

$$EI \frac{d^2y}{dx^2} \Big|_A = \frac{wx^2}{2} \quad \text{for } 0 < x \leq l - a_1 \quad (\text{C.13})$$

$$EI \frac{d^2y}{dx^2} \Big|_B = \frac{wx^2}{2} - R_1(x - l + a_1) \quad \text{for } l - a_1 < x \leq l + a_2 \quad (\text{C.14})$$

$$EI \frac{d^2y}{dx^2} \Big|_C = \frac{wx^2}{2} - R_1(x - l + a_1) - R_2(x - l - a_2) \quad \text{for } l + a_2 < x \leq 2l \quad (\text{C.15})$$

Integrating (C.13) (C.14) and (C.15) and determining arbitrary constants by continuity at support points, gives

$$EI \frac{dy}{dx} \Big|_A = \frac{wx^3}{6}$$

$$EI \frac{dy}{dx} \Big|_B = \frac{wx^3}{6} - \frac{R_1}{2}(x - l + a_1)^2$$

$$EI \frac{dy}{dx} \Big|_C = \frac{wx^3}{6} - \frac{R_1}{2}(x - l + a_1)^2 - \frac{R_2}{2}(x - l - a_2)^2$$

This means that the angle between the end faces, α , is given simply by

$$\frac{dy}{dx} \Big|_{C, x=2l}$$

Thus

$$EI\alpha = \frac{4}{3}wl^3 - \frac{1}{2}(R_1 + R_2)(l^2 + a_1a_2) - \frac{1}{2}(R_1a_1 - R_2a_2)(2l + a_1 - a_2) \quad (\text{C.16})$$

Now, resolving vertically, $R_1 + R_2 = W + M$ and taking moments about the centre of the bar gives $R_1a_1 - R_2a_2 = -M(l + p)$, substituting into (C.16) gives

$$2EI\alpha = W \left(\frac{l^2}{3} - a_1a_2 \right) + M \left\{ (l + p + a_1)(l + p - a_2) - p^2 \right\} \quad (\text{C.17})$$

To check the previous derivation for the Airy points, setting $M = 0$, $a_1 = a_2$ does indeed give the same solution for the positions of the supports.

To see the effect of supporting the bar and flat at the unmodified Airy points, the excess tilt of the ends of the bar can be calculated from

$$\alpha = \frac{Ml^2}{3EI} \left(1 + \frac{3p}{l} \right)$$

The flats are 70 mm diameter, 15 mm thick and have a density of 7800 kg m⁻³. This gives values of $M = 0.4503$ kg, $p = 7.5 \times 10^{-3}$ m, $I = 1.1923 \times 10^{-8}$ m⁴, and for steel, Youngs modulus, $E = 203$ GPa. For a 1 m bar, $l = 1$ m, this gives a value for α of 6.34×10^{-5} radians. Converting this to a change of length across the face of the bar gives a value of 1.4 μm, or over 4 fringes. To correct this, the two supports must be moved either symmetrically, or by moving just one support.

Let

$$\frac{M}{W} = \frac{np}{l}$$

where n is the ratio of the cross-section of the flat to the cross-section of the bar, assuming that the bar and flat are made of the same material, as required to minimise the phase correction.

From (C.17), setting $\alpha = 0$, dividing by W and substituting $\frac{M}{W} = \frac{np}{l}$ gives

$$\left(\frac{l^2}{3} - a_1 a_2\right) + \frac{np}{l} \left\{ (l + p + a_1)(l + p - a_2) - p^2 \right\} = 0$$

There are 4 solutions for the positions of the support points: the first two being non-symmetrical and the remaining two being symmetrical and identical except for a change of sign. The non-symmetrical solutions leave one of the supports at its Airy point, and the solution of the above equation gives the position of the other support. For the symmetrical solution, both of the supports are moved outwards from their Airy points and retain their symmetrical placing about the centre of the bar.

Case (i), support 2 is unmoved, substituting $a_2 = \frac{l}{\sqrt{3}}$ in (C.17)

$$\left(\frac{l^2}{3} - \frac{a_1 l}{\sqrt{3}}\right) + \frac{np}{l} \left\{ (l + p + a_1) \left(l + p - \frac{l}{\sqrt{3}} \right) - p^2 \right\} = 0$$

Separating terms in a_1

$$\frac{l^2}{3} - a_1 \frac{l}{\sqrt{3}} + \frac{np}{l} \left\{ (l + p) \left(l + p - \frac{l}{\sqrt{3}} \right) - p^2 \right\} + a_1 \frac{np}{l} \left(l + p - \frac{l}{\sqrt{3}} \right) = 0$$

$$\frac{l^2}{3} + \frac{np}{l} \left\{ (l + p) \left(l + p - \frac{l}{\sqrt{3}} \right) - p^2 \right\} = a_1 \left(\frac{l}{\sqrt{3}} - \frac{np}{l} \left(l + p - \frac{l}{\sqrt{3}} \right) \right)$$

$$a_1 = \frac{\frac{l^2}{3} + \frac{np}{l} \left\{ (l + p) \left(l + p - \frac{l}{\sqrt{3}} \right) - p^2 \right\}}{\frac{l}{\sqrt{3}} - \frac{np}{l} \left(l + p - \frac{l}{\sqrt{3}} \right)}$$

Removing a common factor of $\frac{l}{\sqrt{3}}$ gives

$$a_1 = \frac{l}{\sqrt{3}} \left\{ 1 + \frac{\left[\frac{-\sqrt{3}}{l}(l+p)-1 \right] \left[-\frac{np}{l} \left(l+p - \frac{l}{\sqrt{3}} \right) \right] - \frac{p^2 \sqrt{3} np}{l^2}}{\frac{l}{\sqrt{3}} - \frac{np}{l} \left(l+p - \frac{l}{\sqrt{3}} \right)} \right\}$$

Multiplying top and bottom by $\sqrt{3}$, separating factors and rearranging gives

$$a_1 = \frac{l}{\sqrt{3}} \left\{ 1 + \frac{\frac{np}{l} \left(l+p - \frac{l}{\sqrt{3}} \right) \left(\sqrt{3} + 3 + \frac{3p}{l} \right) - \frac{3p^2}{l}}{l - \frac{np}{l} \left(\sqrt{3}l + \sqrt{3}p - l \right)} \right\}$$

Multiplying and collecting terms, dividing by l gives

$$a_1 = \frac{l}{\sqrt{3}} \left\{ \frac{\frac{np}{l} \left[2 + \frac{6p}{l} \right]}{1 - \frac{np}{l} \left(\sqrt{3} + \frac{\sqrt{3}p}{l} - 1 \right)} \right\}$$

$$a_1 = \frac{l}{\sqrt{3}} \left\{ \frac{\frac{2np}{l} \left[1 + \frac{3p}{l} \right]}{1 + \frac{np}{l} \left(1 - \sqrt{3} - \frac{\sqrt{3}p}{l} \right)} \right\}$$

Thus with

$$f(\chi) \equiv 1 + \frac{2np/l(1+3p/l)}{1+np/l(1+\chi+\chi p/l)}$$

$$a_1 = \frac{l}{\sqrt{3}} f(-\sqrt{3})$$

For **case (ii)**, support 1 is unmoved, substituting $a_1 = \frac{l}{\sqrt{3}}$ in (C.17) gives a similar solution to case (i), though because the signs of a_1 and a_2 are reversed, the sign of the radical is also reversed in the solution, *i.e.*

$$a_1 = \frac{l}{\sqrt{3}} f(\sqrt{3})$$

For **case (iii)**, both supports are moved symmetrically, substituting $a_1 = a_2 = a$ in (C.17) gives

$$\frac{l^2}{3} - a^2 + \frac{np}{l} \{ (l+p+a)(l+p-a) - p^2 \} = 0$$

Separating terms in a

$$\frac{l^2}{3} - a^2 + \frac{np}{l} \{ (l+p)(l+p) - a^2 \} = 0$$

$$a = \frac{l}{\sqrt{3}} \sqrt{\frac{1 + \frac{3np}{l^3} (l^2 + 2pl)}{1 + \frac{np}{l}}}$$

Dividing

$$a = \frac{l}{\sqrt{3}} \sqrt{1 + \frac{\left[3\left(1 + \frac{2p}{l}\right) - 1 \right] \frac{np}{l}}{1 + \frac{np}{l}}}$$

$$a = \frac{l}{\sqrt{3}} \sqrt{1 + \frac{\frac{2np}{l} \left(1 + \frac{3p}{l}\right)}{1 + \frac{np}{l}}}$$

$$a = \frac{l}{\sqrt{3}} \sqrt{f(0)}$$

Strictly, $a = \pm \frac{l}{\sqrt{3}} \sqrt{f(0)}$ though these two solutions correspond to the two choices of labelling the supports, *i.e.* they are the same physical solution.

In summary, setting α to zero in (C.17) allows for three solutions:

(i) Support 2 remains at the Airy position, and support 1 moves to a new position

$$a_2 = \frac{l}{\sqrt{3}} \quad , \quad a_1 = \frac{l}{\sqrt{3}} f(-\sqrt{3})$$

(ii) Support 1 remains at the Airy position and support 2 moves to a new position

$$a_1 = \frac{l}{\sqrt{3}} \quad , \quad a_2 = \frac{l}{\sqrt{3}} f(\sqrt{3})$$

(iii) Both supports move by equal amounts to new symmetrical positions

$$a_1 = a_2 = a = \frac{l}{\sqrt{3}} \sqrt{f(0)}$$

where $f(\chi) \equiv 1 + \frac{2np/l(1+3p/l)}{1+np/l(1+\chi+\chi p/l)}$

Suitable tolerances on the positioning of the supports may be calculated by differentiating (C.17) with respect to a , this will be performed for the symmetrical solution (case (iii)).

Substituting $a_1 = a_2 = a$ in (C.17) gives

$$2EI\alpha = W\left(\frac{l^2}{3} - a^2\right) + M\{(l+p+a)(l+p-a) - p^2\}$$

$$2EI\alpha = \frac{Wl^2}{3} - Wa^2 + M(l^2 + 2lp - a^2)$$

Differentiating with respect to a gives

$$2EI\delta\alpha = -2a(W + M)\delta a$$

Hence $\delta a = \frac{EI}{a(W + M)} \delta\alpha$ (18)

For a 1 m bar, for a maximum value of $\delta\alpha$ of 1.126×10^{-6} which corresponds to the value of $1 \mu\text{in}$ ($0.025 \mu\text{m}$) error chosen by Williams, $\delta a = 2.4 \times 10^{-3}$, or 2.4 mm. This is better than the tolerance for the general case for which Williams calculated a value of 0.7 mm. Thus the use of symmetrical support positions is preferable, for which positioning within 2.4 mm is required.

Thus by accurate positioning of the support positions, the additional bending may be altered in such a way that the end faces of the bar remain vertical and parallel. The effect of this additional bending on the length of the bar will now be examined.

C.4 EFFECT OF FLEXURE OF BARS ON THEIR LENGTH

The effect on the measured length of the bar is measured on the neutral axis of the bar which runs through the centre of the bar. For a section of the bar, length dx , with

gradient θ the change in length compared to the free state is given by $\frac{\theta^2}{2} dx$, and $\theta = \frac{dy}{dx}$. Thus the total change in length along the whole bar is given by

$$\int_0^{2l} \frac{1}{2} \left(\frac{dy}{dx} \right)^2 dx$$

It is possible to perform this integral, substituting for $\frac{dy}{dx}$ from equations derived earlier, but a simple order of magnitude estimate shows that this is not required as the overall change in length is negligible. Since

$$\int_0^{2l} \frac{1}{2} \left(\frac{dy}{dx} \right)^2 dx \leq l \left(\frac{dy}{dx} \Big|_{\max} \right)^2$$

a maximum value for the change in length due to bending may be calculated. Figure C.4 shows the variation in the vertical position of the neutral axis of a 1m bar with a flat wrung on, supported at the modified symmetrical Airy points and the slope of the bar. The maximum slope is seen to be 8×10^{-6} at approximately 0.7 m from the free end of the bar. Thus the maximum change in length of the bar is 6.4×10^{-11} m (0.002 fringe), *i.e.* negligible.

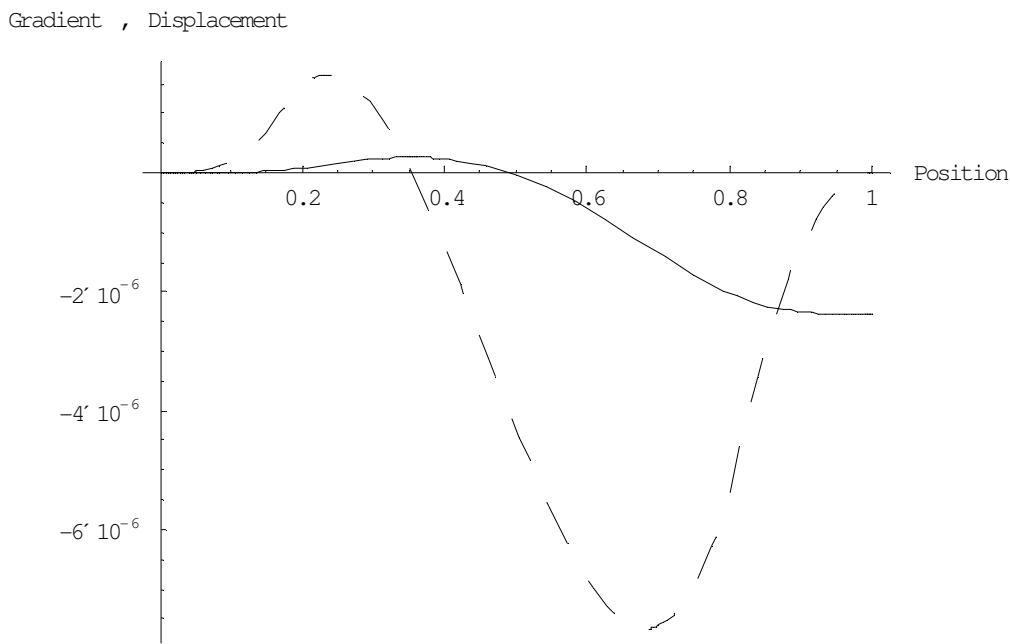


Figure C.4 - Variation in vertical position and gradient (dashed line) of the neutral plane of a 1 m bar, supported at modified Airy points

REFERENCES FOR APPENDIX C

- [1] Rolt F H *Gauges and Fine Measurement* (MacMillan & Co.: London) (1929) 15-16
- [2] Bayer-Helms F Beigung von Endmaßen bei horizontaler Lagerung auf Scheiden *PTB Mitteilungen* (1967) **1/67** 25-30 **2/67** 124-130],
- [3] Williams D C The parallelism of a length bar with an end load *J. Sci. Instrum.* (1962) **39** 608-610

APPENDIX D

PRISMATIC COMPRESSION OF LENGTH BARS

(These results were derived by mathematicians in NPL's Division of Information Technology and Computing).

With reference to Timoshenko & Goodier [1], the forces in the body are

$$X = Y = 0, \quad Z = -\rho g$$

The differential equations of equilibrium ((127) of Timoshenko & Goodier) are satisfied by

$$\sigma_z = \rho g(z - l), \quad \sigma_x = \sigma_y = \tau_{xy} = \tau_{yz} = \tau_{xz} = 0$$

i.e. by assuming that on each cross section we have a uniform compression produced by the upper portion of the bar (see figure D.1).

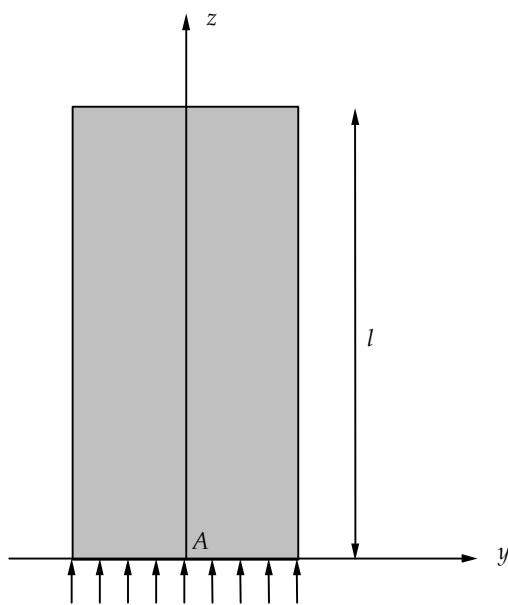


Figure D.1 - Compression on cross section of bar

Hooke's law gives

$$\varepsilon_z = \frac{\partial w}{\partial z} = \frac{\sigma_z}{E} = \frac{\rho g}{E}(z - l) \quad (\text{D.1})$$

$$\varepsilon_x = \varepsilon_y = \frac{\partial u}{\partial x} = \frac{\partial v}{\partial y} = -v \frac{\rho g}{E} (z - l) \quad (\text{D.2})$$

$$\gamma_{xy} = \gamma_{xz} = \gamma_{yz} = \frac{\partial u}{\partial y} + \frac{\partial v}{\partial x} = \frac{\partial u}{\partial z} + \frac{\partial w}{\partial x} = \frac{\partial v}{\partial z} + \frac{\partial w}{\partial y} = 0 \quad (\text{D.3})$$

Integrating (D.1) gives

$$w = \frac{\rho g z^2}{2E} - \frac{\rho g l z}{E} + w_0 \quad (\text{D.4})$$

where w_0 does not depend on z , i.e. $w_0 = w_0(x, y)$. Substituting (D.4) into the second and third of equations (D.3) we find

$$\frac{\partial w_0}{\partial x} + \frac{\partial u}{\partial z} = 0, \quad \frac{\partial w_0}{\partial y} + \frac{\partial v}{\partial z} = 0$$

from which

$$u = -z \frac{\partial w_0}{\partial x} + u_0, \quad v = -z \frac{\partial w_0}{\partial y} + v_0 \quad (\text{D.5})$$

where u_0 and v_0 are functions of x and y only. Substituting these expressions into (D.2) we find

$$-z \frac{\partial^2 w_0}{\partial x^2} + \frac{\partial u_0}{\partial x} = -\frac{\nu \rho g}{E} (z - l), \quad -z \frac{\partial^2 w_0}{\partial y^2} + \frac{\partial v_0}{\partial y} = -\frac{\nu \rho g}{E} (z - l)$$

Equating polynomial terms in z , we have

$$\frac{\partial^2 w_0}{\partial x^2} = \frac{\partial^2 w_0}{\partial y^2} = \frac{\nu \rho g}{E}, \quad \frac{\partial u_0}{\partial x} = \frac{\partial v_0}{\partial y} = \frac{\nu \rho g l}{E} \quad (\text{D.6})$$

Substituting expressions (D.5) into the first of equations (D.3), we find

$$-2z \frac{\partial^2 w_0}{\partial x \partial y} + \frac{\partial u_0}{\partial y} + \frac{\partial v_0}{\partial x} = 0$$

and since u_0 and v_0 do not depend on z , we have

$$\frac{\partial^2 w_0}{\partial x \partial y} = 0, \quad \frac{\partial u_0}{\partial y} + \frac{\partial v_0}{\partial x} = 0 \quad (\text{D.7})$$

From the six equations in (D.6) and (D.7) we can write general expressions for u_0 , v_0 and w_0 . All these equations are satisfied by

$$u_0 = \frac{\nu \rho g l x}{E} + \delta y + \delta_1$$

$$v_0 = \frac{\nu \rho g l y}{E} - \delta x + \gamma_1$$

$$w_0 = \frac{\nu \rho g}{2E} (x^2 + y^2) + \alpha x + \beta y + \gamma$$

Now substituting these expressions into equations (D.4) and (D.5), the general expressions for the displacements are

$$u = -\frac{\nu \rho g x z}{E} - \alpha z + \frac{\nu \rho g l x}{E} + \delta y + \delta_1$$

$$v = -\frac{\nu \rho g y z}{E} - \beta z + \frac{\nu \rho g l y}{E} - \delta x + \gamma_1$$

$$w = \frac{\rho g z^2}{2E} - \frac{\rho g l z}{E} + \frac{\nu \rho g}{2E} (x^2 + y^2) + \alpha x + \beta y + \gamma$$

The six arbitrary constants are determined from the conditions at the support. We prevent translatory movement of the bar by fixing the centroid A of the lower end of the bar so that for $x = y = z = 0$, we have $u = v = w = 0$. We eliminate rotation of the bar about axes through A parallel to the x and y axes by fixing an element of the z axis at A . Then $\partial u / \partial z = \partial v / \partial z = 0$ at A . We avoid the possibility of rotation about the z axis by ensuring that $\partial v / \partial x = 0$ at the point A . From these six conditions at A , we find that all the arbitrary constants (α , β , γ , γ_1 and δ_1) vanish. Thus we are left with

$$u = \frac{\nu \rho g x}{E} (l - z)$$

$$v = \frac{\nu \rho g y}{E} (l - z)$$

$$w = \frac{\rho g z^2}{2E} + \frac{\nu \rho g}{2E} (x^2 + y^2) - \frac{\rho g l z}{E}$$

Hence for the point $(x,y,z) = (0,0,l)$ initially, we have, after deformation:

$$z = l + w(0,0,l) = l - \frac{\rho gl^2}{2E}$$

Hence the change in length of a bar when standing vertically, due to its own weight, is given by

$$-\frac{\rho gl^2}{2E}$$

Using this equation and average values for steel of $\rho = 7800 \text{ kg m}^{-3}$, $E = 208 \text{ GPa}$, $g = 9.8 \text{ m s}^{-1}$ produces the graph shown in figure D.2.

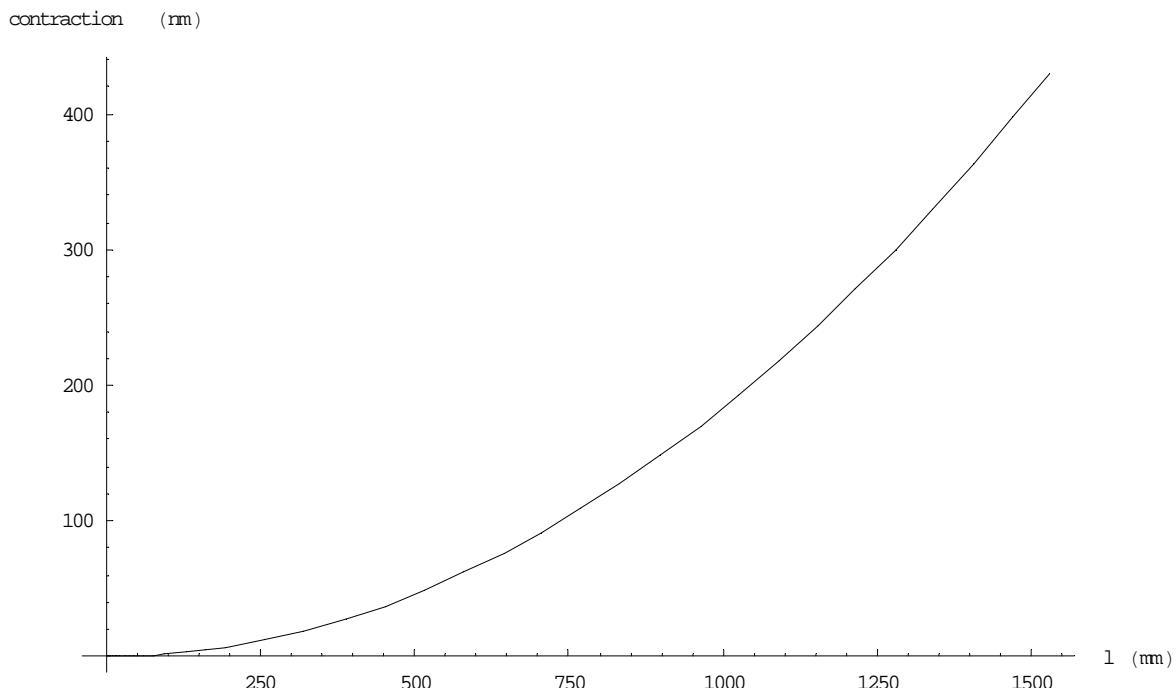


Figure D.2 - Contraction of a steel length bar, standing vertically

REFERENCES FOR APPENDIX D

- [1] Timoshenko S & Goodier J N *Theory of Elasticity* (McGraw-Hill, New York) (1951) 246

APPENDIX E

CONNECTORS & CONNECTIONS

(This appendix is provided mainly as a reference for future modifications of the instrument and for maintenance.)

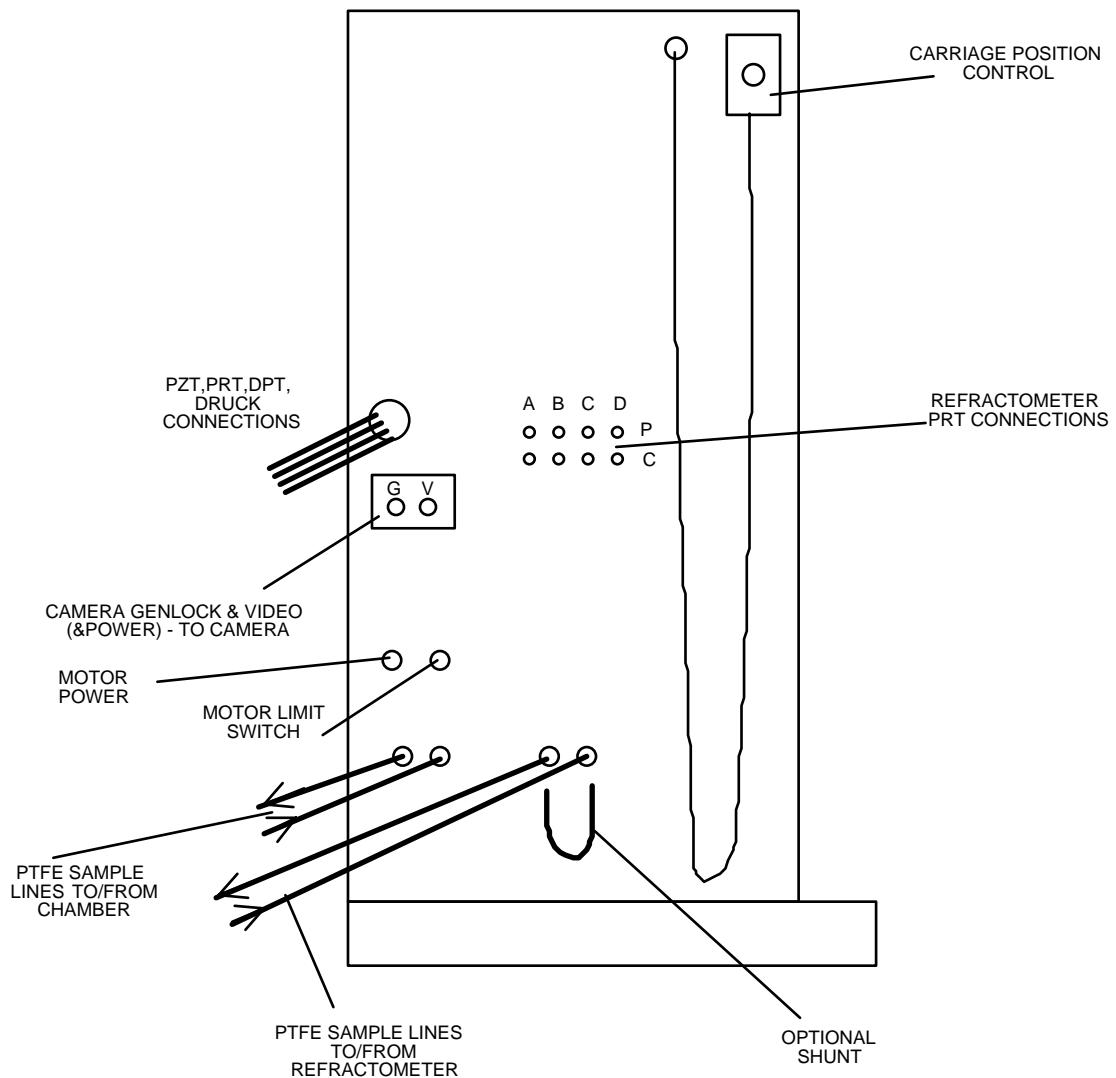


Figure E.1 - Electronics rack, left side view of connectors

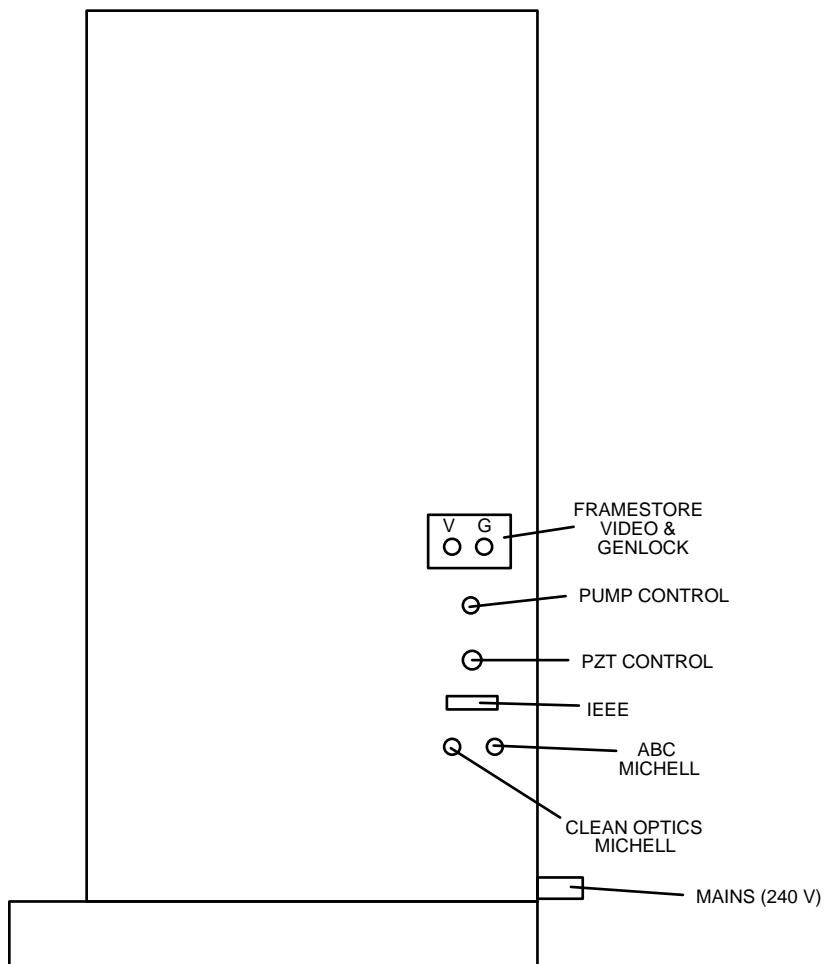


Figure E.2 - Electronics rack, right side view of connectors

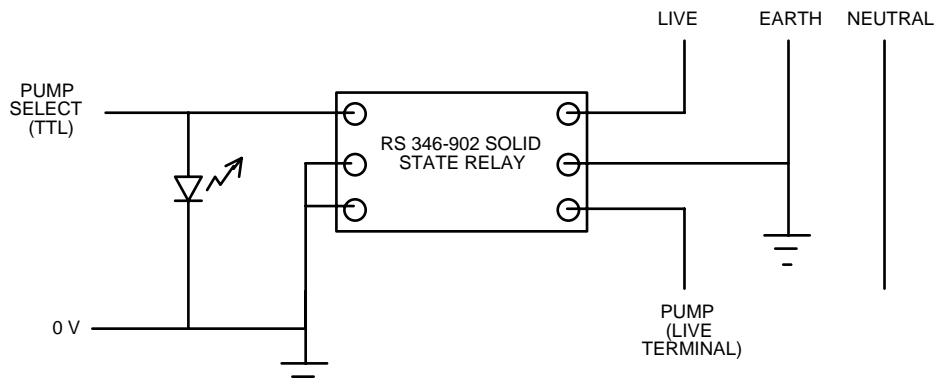


Figure E.3 - Pump control using solid state relay

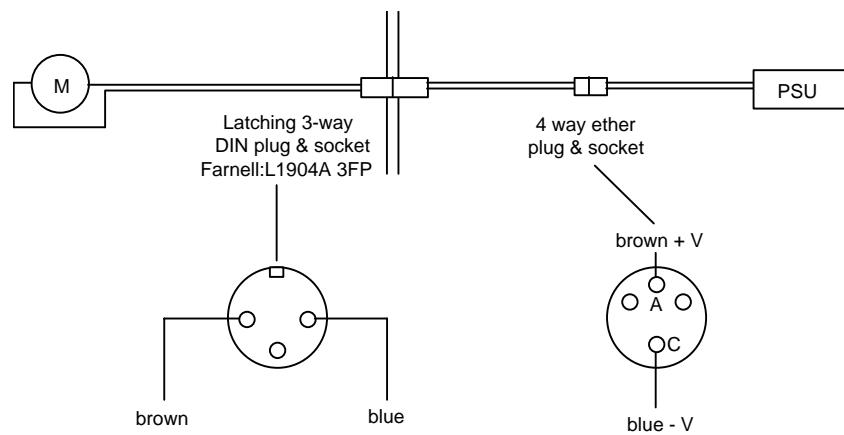


Figure E.4 - Motor power connections from motor to PSU

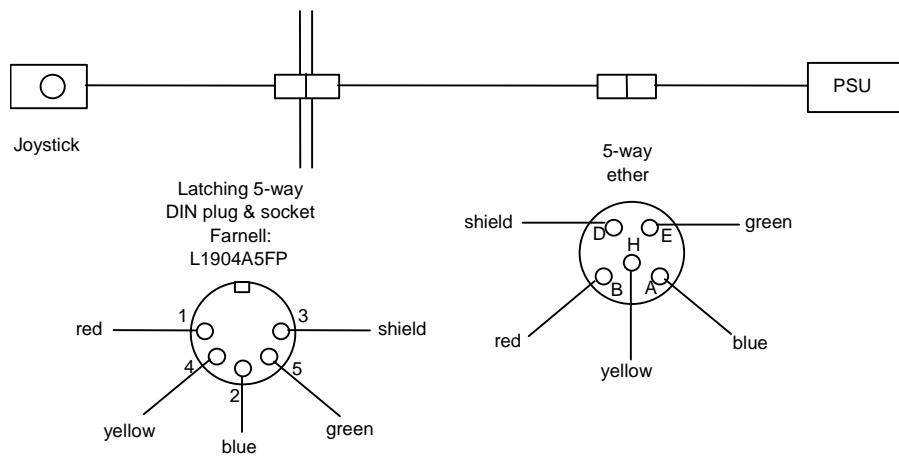


Figure E.5 - Motor direction control, from joystick microswitches to PSU board

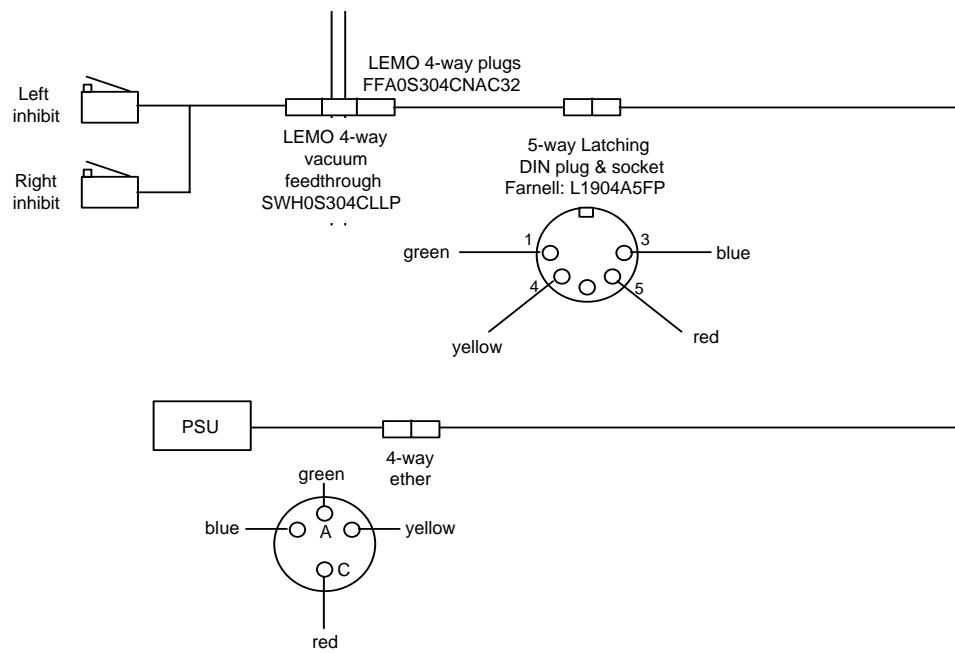


Figure E.6 - Motor inhibit connections, microswitches on stage to PSU board

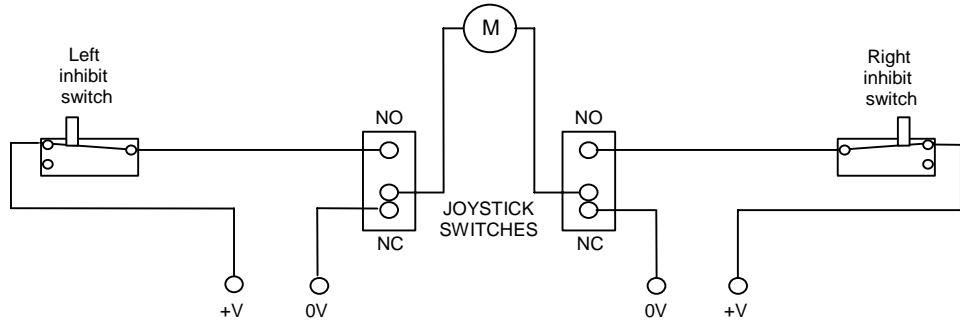


Figure E.7 - Motor control from joystick switches to motor

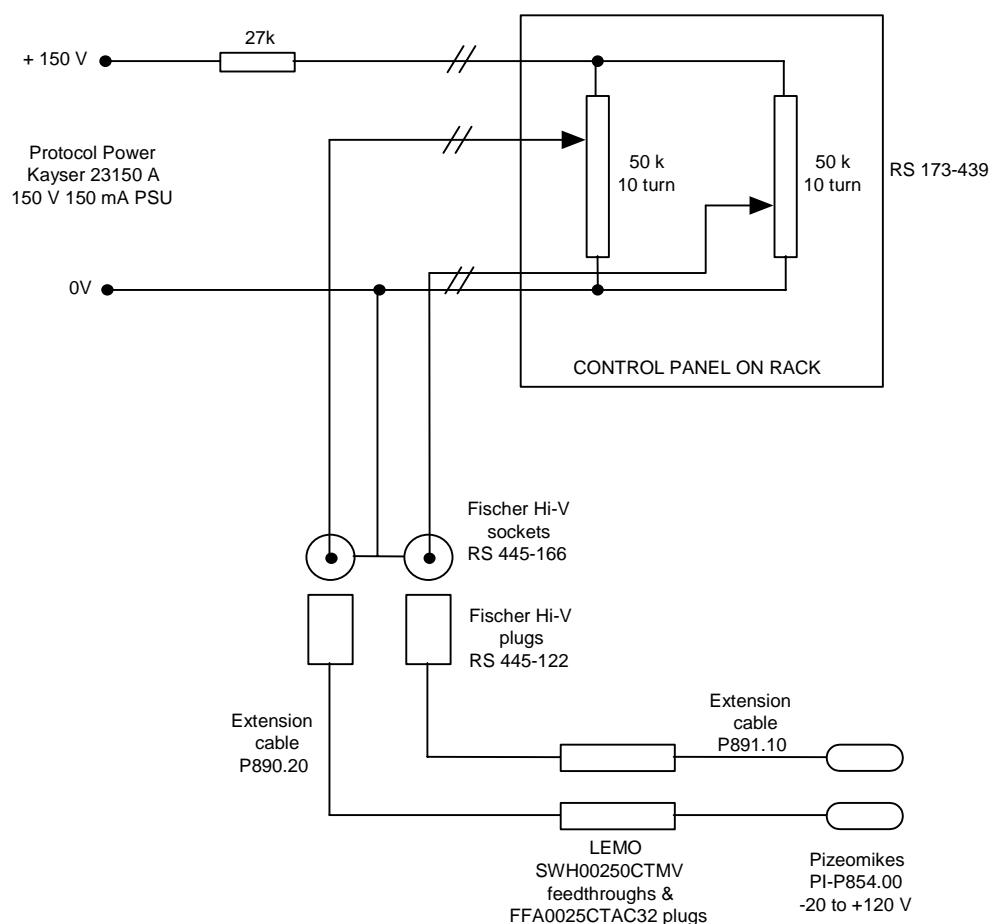


Figure E.8 - Low voltage PZT power supply connections

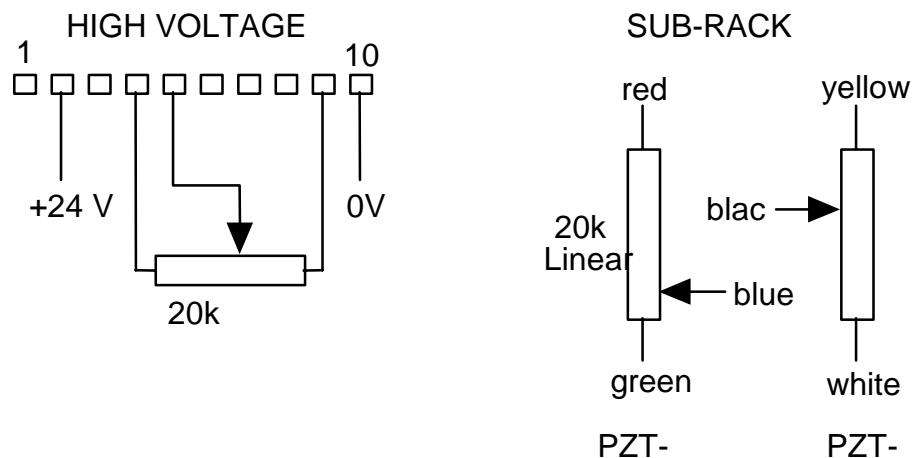


Figure E.9 - High voltage PZT power supply connections

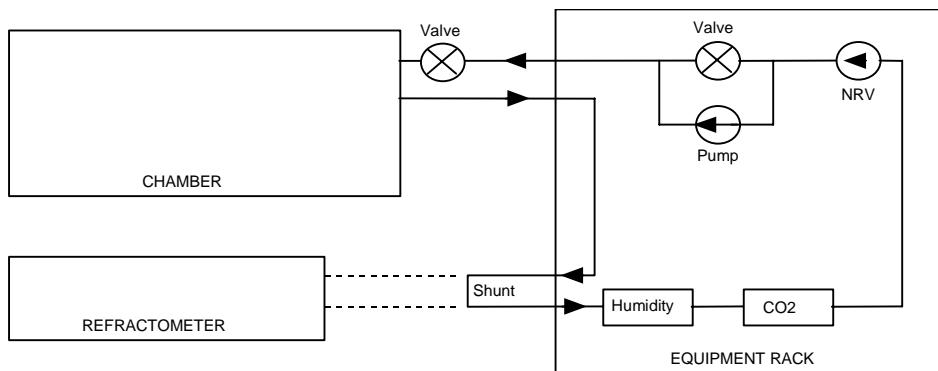


Figure E.10 - Gas sample flow connections

LEMO connectors

PRTs

Feedthroughs:	SWH0S304CLLP
Plug internal:	FFA0S304CNAC32
Plug external:	FFA0S304CNLC42

High voltage PZTs

Feedthroughs:	SWH1S405FTMSV
Plugs:	FFB1S405CTAC57

Low voltage PZTs

Feedthroughs:	SWH00250CTMV
Plugs:	FFA0025CTAC32

Queensgate DPT

PZT voltage feedthrough:	SWH1S405FTMSV
Sense/drive feedthrough:	SWH00250CTMV
Sense/drive plugs:	FFA0025CTAC

Motor limit switches

Feedthroughs:	SWH0S304CLLP
Plugs:	FFA0S304CNAC32

Fan

Feedthrough:	SWH0S304CLLP
Plugs:	FFA0S304CNAC32

APPENDIX F

PUBLICATIONS BY THE AUTHOR

This appendix contains copies of papers published by the author, relevant to the thesis.
These are:

Hughes E B, Jackson K, Lewis A J & Pugh D J

Recent advances in length measurement at the National Physical Laboratory, England

Metrology and Total Quality - Proceedings of the National Conference of Standard Laboratories, Washington DC (1990) 285-294

Lewis A J

Two-wavelength phase-stepping interferometry for absolute length measurement

Applied Optics Digest - Proceedings of Applied Optics & Opto-Electronics Conference, Nottingham (1990) 269-270

Lewis A J & Pugh D J

Design Note: Interferometer light source and alignment aid using single-mode optical fibres

Meas. Sci. Technol. **3** (1992) 929-930

Lewis A J

Three-wavelength phase-stepping interferometer for length measurement up to 1.5 m in a controlled environment

Proceedings of Applied Optics & Opto-Electronics Conference, Leeds (1992) 170-172

A paper entitled "Measurement of length, surface form and thermal expansion coefficient of length bars up to 1.5 m using multiple-wavelength phase-stepping interferometry" has been submitted for publication in a special issue of Measurement Science and Technology on Optical Techniques in Measurement, and is due for publication in June 1994.

RECENT ADVANCES IN LENGTH MEASUREMENT AT THE NATIONAL PHYSICAL LABORATORY, ENGLAND

E B Hughes, K Jackson, A J Lewis and D J Pugh

Division of Mechanical and Optical Metrology

National Physical Laboratory, Teddington, England.

ABSTRACT

Gauge blocks and length bars provide industry with practical secondary length standards. To provide traceability of these material standards to international standards, they are measured in terms of known wavelengths of frequency-stabilised lasers.

In the last five years NPL has been active in updating its facilities for the measurement of end standards. The main aims of the programme have been to simplify the process of measurement by interferometry so that relatively unskilled staff can operate the equipment and to reduce the time required for measurement.

The use of modern instrumentation and stabilised lasers has also meant that higher accuracy measurements can be made and that more information relating to the variations in length over the gauge measurement surfaces can be obtained.

The NPL programme has concentrated on three main areas of work:

- the development of an automatic gauge block interferometer for the measurement of gauges between 0.1 and 100 mm in length,
- a white light interferometer for gauge block measurement,
- a phase stepping interferometer for length bar measurement in the range **100 mm to 1500 mm**.

INTRODUCTION – GAUGE BLOCK INTERFEROMETRY

The method of gauge block measurement using multiple wavelength interferometry is well established 1,2. Gauges to be measured are wrung vertically on to a reference flat and placed in an interferometer producing a set of interference fringes across the gauge and flat; see Figure 2. The fringe spacing corresponds to a height difference of half the wavelength of light used, approximately 300 nm. Having previously determined the approximate length of the gauge, an accurate value can be deduced by measuring the displacement of the fringes on the gauge relative to those on the flat at several wavelengths.

Accurate measurements of the air pressure, temperature and humidity are needed to correct for changes in the measurement wavelengths due to changes in the refractive index of air. Also because of the thermal expansion of the gauge it is necessary to measure accurately its temperature, so that the gauge length can be corrected for the specified operating temperature.

Until recently the interference fringe displacements were estimated visually by skilled operators who were also required to measure manually the pressure, temperatures and humidity. The process was tiring, slow and subject to errors. A new instrument was designed to overcome these deficiencies and to provide greater accuracy.

AUTOMATIC GAUGE BLOCK INTERFEROMETER

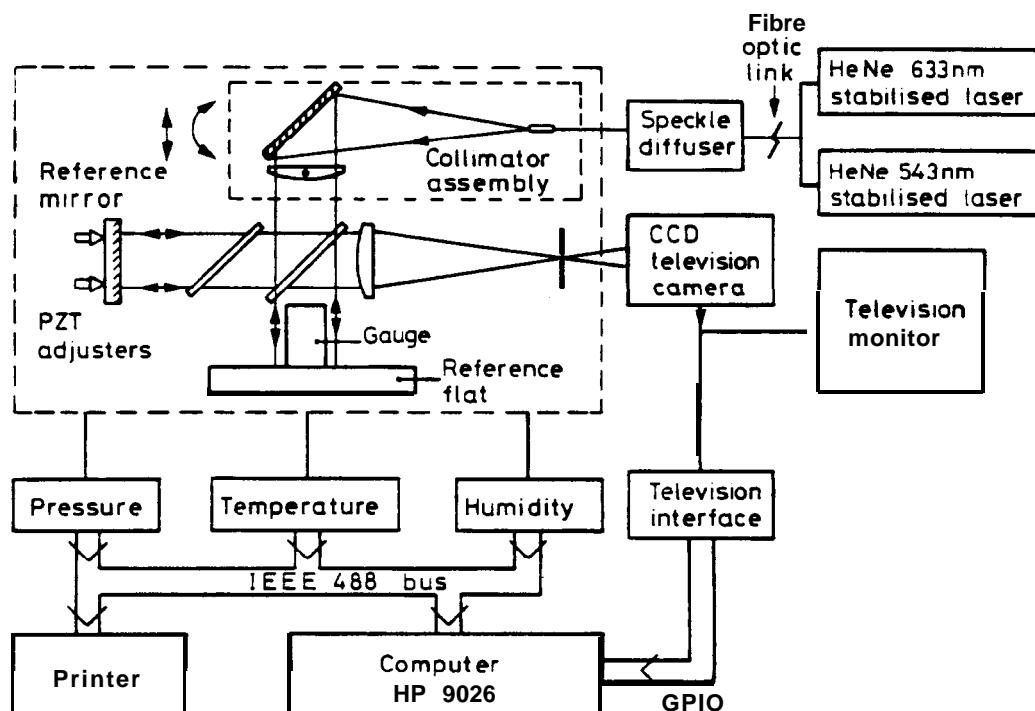


Figure 1. Schematic diagram of automatic gauge block interferometer

A Twyman-Green interferometer forms the basis of the new instrument ³, the high contrast sinusoidal interference **fringes** being well suited to computer analysis. Two He-Ne frequency stabilised lasers (633 and 543 nm) developed at NPL provide the measurement wavelengths. These lasers wavelengths are very stable and are traceable to international length standards with an uncertainty less than 1 part in 10^8 . The brightness and spectral purity of these lasers allow the measurement of lengths up to several metres without loss of fringe contrast.

Automatic measurement of the fringe shifts at the two wavelengths is achieved by examining the fringes with a CCD television camera and calculating the fringe positions with a computer. A photograph of the fringes detected by the television camera is shown in Figure 2; the shift between the fringes on the gauge and reference flat can be seen clearly.

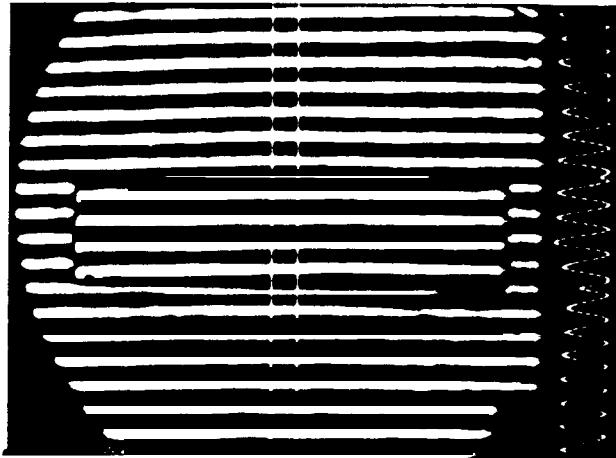


Figure 2. Interferometer output

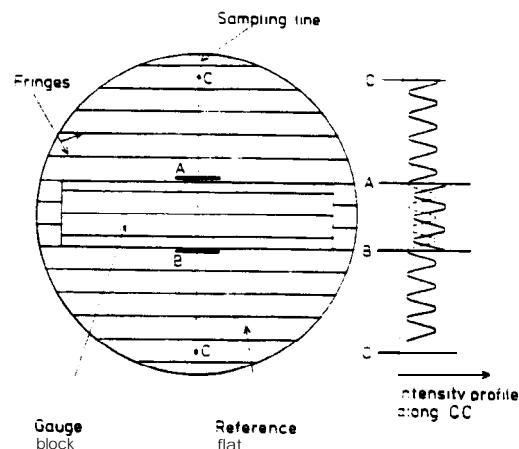


Figure 3. Fringe intensity profile

Figure 3 shows a drawing of the **fringes** on the gauge and flat with a plot of the intensity profile of the fringes along the line CC, on the right. By entering this profile into the computer and calculating the positions of the fringe minima, the fringe shift can be determined automatically. Because of the influence of pressure, temperature and humidity on the wavelength of light, frequent and accurate monitoring of these parameters is very important. This is achieved using an automatic resistance bridge with platinum resistance thermometers, a vibrating cylinder pressure transducer and a dew-point hydrometer, all interfaced to the computer. Wavelength corrections are applied automatically by the computer. The total uncertainty in the wavelength correction due to errors in this instrumentation and in the computer correction algorithm is typically 1 part in 10^7 (± 10 nm for a 100 mm gauge).

Use of the instrument

In practice the instrument is very easy to use. Up to fourteen gauges are wrung on to a large circular reference flat, placed in the instrument, and allowed to reach thermal equilibrium over several hours.

To measure the length of a gauge, the operator views it in the interferometer on a television screen, positions the gauge between two bright line markers (see Figure 2) and adjusts the fringes by remote control of the interferometer reference mirror.

From this point on, the measurement is controlled by the computer. The operator is prompted to type in the gauge reference number, the thermal expansion coefficient and the nominal gauge length. The gauge temperature, air temperature, pressure and humidity are now read automatically by the computer. The two laser wavelengths are selected sequentially and the fringe shift for each wavelength measured. The gauge length is then calculated with all corrections applied automatically. The whole process takes about two minutes for each gauge. The length of each gauge measured is held in a data file and at the end of the measurement sequence a calibration certificate can be printed.

Flatness and surface topography

By measuring the gauge length at a number of points on the measurement surface the gauge flatness and topography can be calculated. This is achieved by moving the sampling line (Figure 3) to eleven positions along the gauge surface under computer control. Typical results of a gauge measured by this technique, over the whole of a gauge measurement face, are shown in Figure 4. Flatness, parallelism, and surface topography are calculated from the variations in length over the gauge surface.

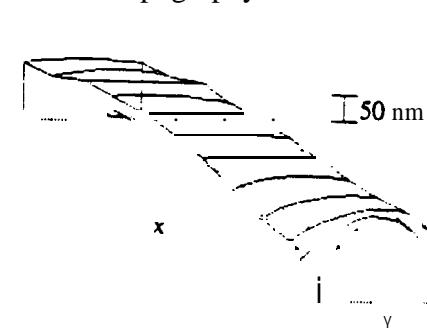
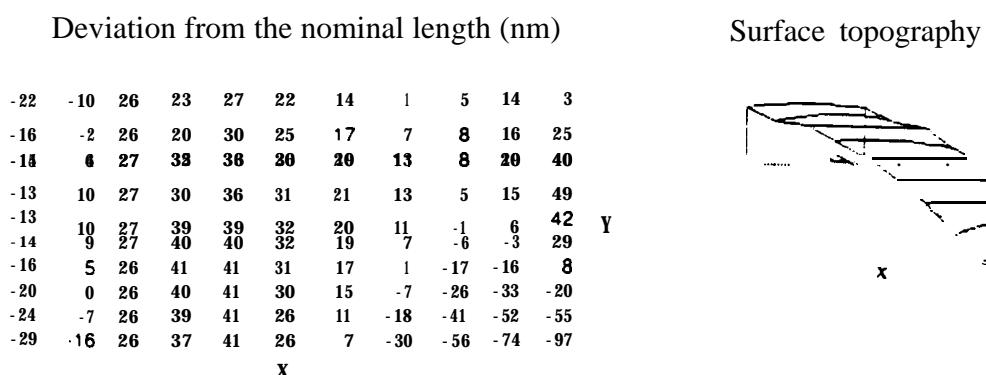


Figure 4. Flatness and topography results



Figure 5. Interferometer and control console

Performance

As with all absolute measurement instruments the overall measurement uncertainty can only be determined by considering the errors in each part of the measurement process, and combining them by calculation. Major contributions to the overall uncertainty arise due to errors in the measurement of the gauge temperature, fringe displacement, and uncertainty in the gauge thermal expansivity. For this instrument the uncertainty of measurement has been calculated to be $\pm (0.02 + 0.2L) \mu\text{m}$, where L is in metres, with a confidence level of 99% (eg $\pm 40 \text{ nm}$ for a 100 mm gauge). Recent intercomparisons with other laboratories suggest that this is a conservative estimate.

Because of the demand from other international laboratories for an instrument of this type, the design for manufacture has been licensed to a company specialising in this type of equipment. (Tesa Metrology, Telford)⁴

A photograph of the interferometer and control console is shown in Figure 5.

WHITE LIGHT INTERFEROMETER

Introduction

To prevent ambiguity in the measurements made in the automatic gauge block interferometer described above, it is necessary to have prior knowledge of the gauge block length to better than $\pm 1 \mu\text{m}$. This is usually achieved by separate and time consuming mechanical comparison techniques. To reduce the measurement time and to avoid the necessity to store a large number of calibrated gauges for measurement by mechanical comparison, the existing automatic gauge block interferometer at NPL has been modified.

The new instrument now uses two independent techniques to measure the length: white light interferometry to determine the approximate value, and two wavelength laser interferometry to give a more accurate measurement of the length of the gauge.

Principle of operation

The existing gauge block interferometer has been modified to allow the technique of white light interferometry to be used. A **slideway** to carry the reference mirror and a commercial plane mirror interferometer, to monitor its position, have been added. Photodetectors have been introduced in the image plane to detect the fringes and an optical compensation plate positioned between the reference mirror and beamsplitter. These are illustrated in Figure 6.

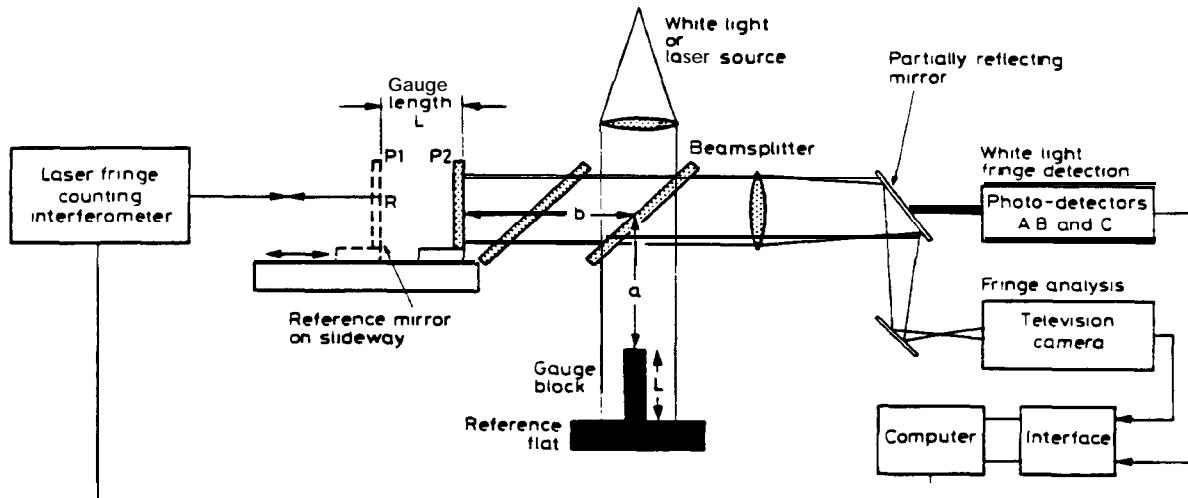


Figure 6. Schematic diagram of white light interferometer

When the laser source in the interferometer is replaced with a white light source, interference fringes will only be observed when the two optical path lengths of the interferometer are nearly equal, that is when $b = a$ and when $b = a + L$. With a gauge wrung to a flat, fringes will only be visible for two positions of the reference mirror, **P1** and **P2**, corresponding to the top of the gauge and reference flat respectively. The gauge length can be determined by moving the reference mirror and measuring the distance between the positions at which the fringes are detected.

In operation the reference mirror is driven along a slideway, the distance continuously monitored by a fringe counting interferometer and the data entered into a computer. At the same time photodetectors in the image plane, also coupled to the computer, record the intensity of the fringes. By computing the mirror positions at which the central dark fringes occur, at the gauge and either side of reference flat, the gauge length can be calculated. Figure 7 shows the position of the photodetectors A, B and C in the image plane and their outputs as the reference mirror is moved along the slideway. The position of the reference flat is measured at either side of the gauge block by the photodetectors A and C; the mean of these two results provides the reference flat position in the central region of the gauge.

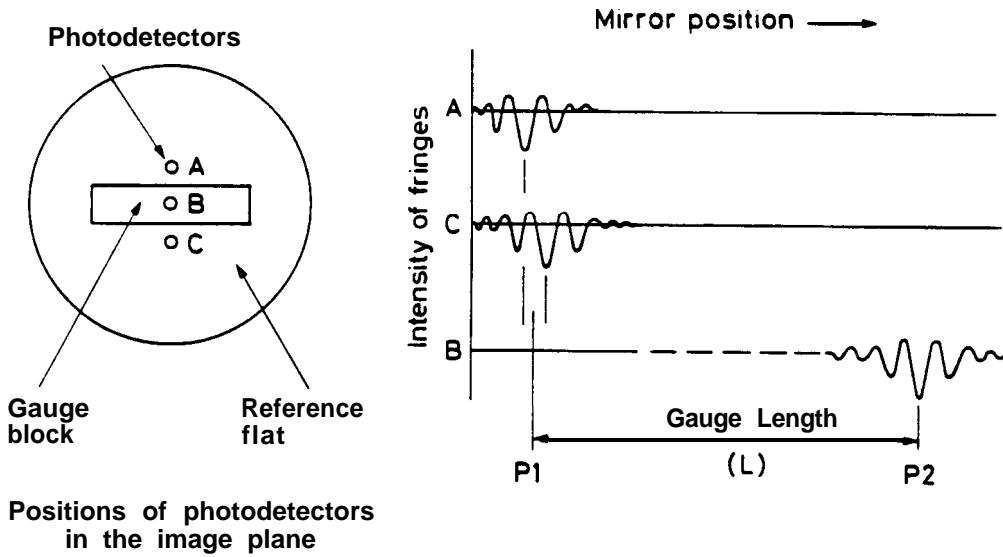


Figure 7.

Operation of white light interferometer

Performance

In operation, the combined measurement of white light and two wavelength interferometry is totally automatic. Measurement and calculation of the data is treated separately so that two completely independent results are obtained. In practice the difference in the two measured lengths does not exceed ± 50 nm. The time taken for a dual measurement varies between 3 and 5 minutes, depending on the gauge length.

LENGTH BAR INTERFEROMETER

As machines for mechanical measurements improve in accuracy there is a need for higher accuracy length standards to validate their performance. To fulfil this requirement and a need for higher accuracy measurements in NPL, the development of a new facility has begun, which will enable length bars from 0.1 to 1.5 metres to be measured with a total uncertainty off 0.1 μm .

The interferometric technique which has been adopted is similar to that used for gauge block measurement but the need for control and measurement of the ambient conditions is much more stringent due to the long interference path length.

The most critical areas affecting the performance of the new instrument are the measurement of the refractive index of air, length bar temperature and the fringe shifts at the two wavelengths.

Length uncertainties due to errors in reading the bar temperature can be reduced to an acceptable level using high quality platinum resistance thermometers and a precision a.c. resistance bridge; an accuracy of $\pm 2 \text{ mK}$ is achievable (equivalent to a length uncertainty of ± 30 nm in 1.5 metres).

The error in determining the fringe shifts using phase stepping interferometry can be reduced to a level equivalent to a length uncertainty of only ± 3 nm.

Because of the effect of the refractive index of air on the laser wavelengths, its accurate measurement is **essential**. This measurement provides the major problem in high accuracy length interferometry. The indirect determination of the refractive index of air using 'state-of-the-art' instrumentation to measure air pressure, temperature and humidity in combination with a correction formulae⁵ would not be sufficiently accurate for this application.

Our initial approach to the problem will be to operate the interferometer in a sealed chamber, filled with a gas of known composition. The refractive index of the gas will be determined in a separate experiment or with an independent refractometer linked to the chamber. With both of these options the uncertainties in the measured refractive index should be in the region of ± 2 parts in 108 (± 20 nm in 1 metre). Frequency stabilised lasers similar to those in the gauge block instrument are used as a light source and these are **sufficiently** accurate for the measurement of bars up to 1.5 metres.

Description of the interferometer

The interferometer, which is similar to that used for gauge block measurement, is of a Twyman-Green design, but the optical components have been rearranged so that length bars can be measured in the horizontal position. Path folding mirrors are included to reduce the physical size of the instrument. This is illustrated in Figure 8.

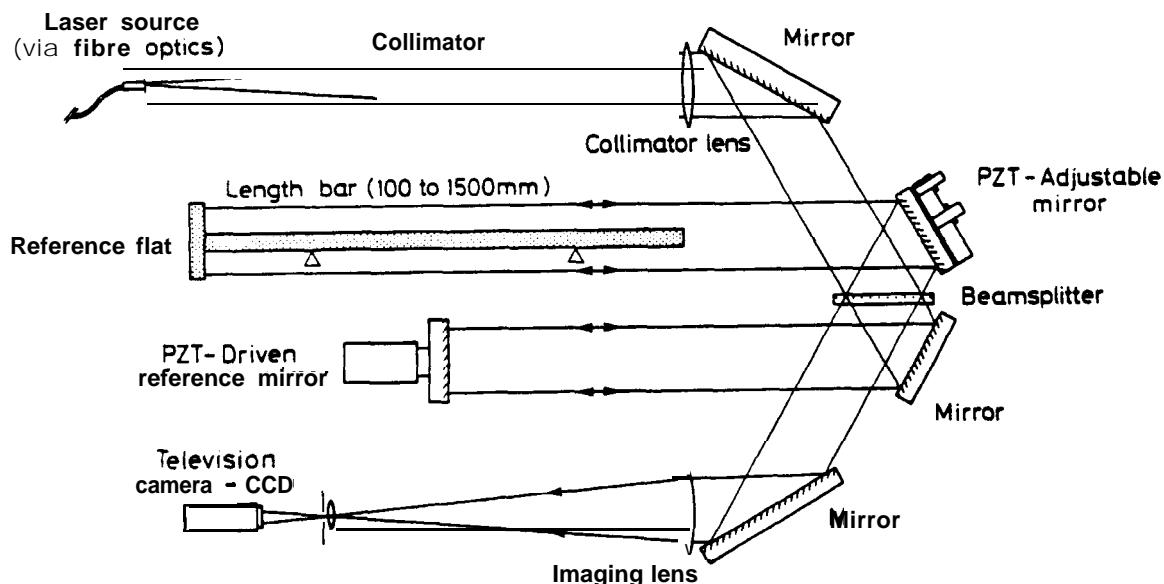


Figure 8. Optical layout of length bar interferometer

Bars are measured in a horizontal position, with a small reference flat wrung on to one end. They are supported near the Airy^{6,7} points which provide the minimum deflection of the two end faces, from the **free** state. An adjustable mirror with remote **PZT** control is included to ensure that the interferometer beam runs parallel to the bar to avoid obliquity errors and also to allow adjustment of the fringe orientation.

Phase stepping interferometry ⁸ is being used to determine the phase shifts at the two wavelengths. In this technique the interferometer reference mirror is accurately moved by a PZT actuator in five nominally equal steps and at each step an image of the fringes is captured and stored in a TV frame store.

From the stored data of the five images the phase at any point in the image can be computed and a phase map of the whole image built up. Phase shifts between the length bar and reference flat can be determined to better than ± 3 nm using this technique and data relating to the surface flatness can be accurately determined. An example of the surface topography of a measurement face of a typical length bar is shown in Figure 9.

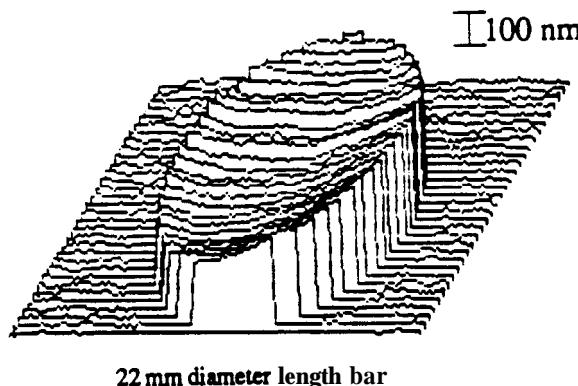


Figure 9. Measurement face surface topography

CONCLUSIONS

The NPL programme for updating its facilities for the absolute measurement of end standards is well under way. A fully automatic gauge block interferometer has been in regular service for four years and a modified version, for the measurement of gauge blocks without prior knowledge of the gauge length, is nearing completion. Our experience with these automatic techniques, is not only that they are faster and more accurate, but that the overall quality of the results has improved. This is mainly due to the ease of use of the instrument, encouraging the metrologist to repeat measurements whenever there is the slightest inconsistency in the results.

Specifications regarding the flatness of the gauge surfaces can be validated objectively rather than relying on subjective visual estimations, as currently used in manual interferometers.

The length bar interferometer programme has been running for one year and the initial results from a prototype instrument, using the phase stepping method, are being obtained. It is planned that this instrument will be fully operational by 1993.

REFERENCES :

- 1 Hume, K.J., Engineering Metrology, Macdonald, London, 1963.
- 2 Candler, C., Modem Interferometry, Hilger and Watts Ltd., 195 1.
- 3 Pugh, D.J., Jackson, K., "Automatic Gauge Block Measurement Using Multiple Wavelength Interferometry", SPIE Proc. Vol. 656.1986, pp 244 - 250.
- 4 Tesa Metrology, P.O. Box 418, Halesfield, Telford, Shropshire, **TF7 4QN**, England.
- 5 Edlén, B., "The refractive Index of Air", Metrologia Vol. 2.1966, pp 7 1.
- 6 Airy, G.B., Phil. Trans. Vol. 147, 1847, pp 629.
- 7 Williams, D.C., "The Parallelism of a Length Bar with an End Load", J. Sci. Instrum. Vol. 39, pp 608 - 610.
- 8 Creath, K., "Phase Measurement Interferometry Techniques", Progress in Optics XXVI, Elsevier Science Publishers, 1988.

Two-Wavelength Phase-Stepping Interferometry For Absolute Length Measurement

A. J. Lewis,

Division of Mechanical and Optical Metrology,
National Physical Laboratory, Teddington, Middlesex. TW11 0LW, UK.

Introduction

Traceability of engineering measurements in the UK to national length standards requires the calibration of the highest grade of mechanical length standards at NPL. These standards, known as length bars, take the form of solid metal rods with polished ends ranging in size from 25 mm to 1500 mm in length, and are measured in a Twyman-Green, phase-stepping interferometer using light of known wavelengths. Phase-stepping interferometry has two distinct advantages over fringe interpolation techniques: reduced uncertainty in the measurement of fringe displacements and more detailed information about the measured surface.

1. Interferometry of Length Bars

The length of a bar is measured in terms of the wavelength of light emitted by a frequency-stabilised laser. A small reference flat is wrung[†] to one end of the bar, then the bar and flat are positioned in one arm of a Twyman-Green interferometer. After alignment, interference tilt fringes of a sinusoidal profile are visible on the bar. This image is viewed by a TV camera and stored in a computer framestore.

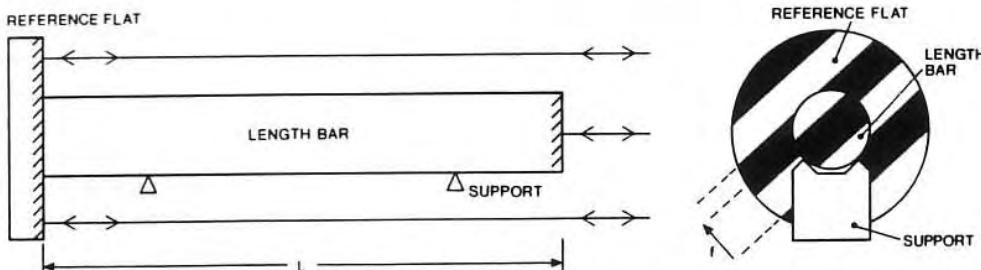


Figure 1. Length bar supported in interferometer, with image of fringes.

The displacement of the fringes on the bar with respect to those on the flat is due to the difference in the optical path length, $2L$, between the light paths which strike the front of the bar and those which are reflected from the surface of the reference flat. The path difference can be expressed in terms of the wavelength of the light, λ :

$$2L = (n + f) \lambda \quad 1.1$$

where n is an integer, or fringe order and f is the fringe fraction. It is the fractional shift of the fringes, f , which is visible and is measured.

2. Phase-Stepping

The 5-position technique, proposed by Hariharan (2), provides accurate measurement of these fringe fractions in terms of phase. In operation, the reference mirror in the reference arm of the interferometer is moved from an arbitrary start position, in 4 equal steps of 1/4 of a fringe. The images of the fringes at all 5 positions of the mirror are digitised into a framestore. The phase, ϕ , of the light reflected from the surfaces, with respect to some arbitrary value, taken to be zero, is given by:

$$\phi(x,y) = \arctan \left\{ 2 \left(\frac{I_2 - I_4}{2I_3 - I_5 - I_1} \right) \right\} \quad 2.1$$

where I_n represents the digitised intensity at a point (x,y) in image number n . This expression is able to return values of ϕ in the range 0 to 2π by using the signs of the numerator and denominator to determine the quadrant of the phase.

Discontinuities in the resultant phase map are removed by sequentially scanning through pixels, looking for phase differences larger than π between neighbouring pixels. Due to the discontinuity present around the circumference of the bar, a new 3-directional phase-unwrapping technique is used, the phase being corrected firstly in a downwards direction, then from left to right, and finally from right to left. A mathematical function is fitted to the now smooth phase map to remove tilt.

An initial estimate of the length of the bar is required before the technique of multiple-wavelength interferometry can be used. The length of the bar is measured initially in another instrument, to within the required uncertainty of $\pm 0.9 \mu\text{m}$. The bar is then measured in the interferometer using two wavelengths: 633 nm and 543 nm. Fringe fractions measured at the centre of the bar at these two wavelengths and the initial length measurement are used in the method of exact fractions (3) to calculate the accurate length of the bar.

3. Conclusions

The techniques of phase-stepping and multiple-wavelength interferometry have been successfully combined in a length measuring interferometer. Results from the prototype length bar interferometer indicate that this technique can be used to measure the lengths of mechanical length standards up to 1500 mm in length.

Two standards of lengths 4.5 mm and 400 mm have been measured; the spreads of measured lengths were $\pm 1 \text{ nm}$ and $\pm 5 \text{ nm}$, and the standard errors of the means were $\pm 0.3 \text{ nm}$ and $\pm 2.4 \text{ nm}$, respectively. These uncertainties represent the errors in the calculated length due to the spread in the fringe fractions, and do not include other errors associated with the measurement of temperature, humidity, and pressure of the air, as well as the temperature of the bar. Surface profiles of these objects were also determined as part of the length measuring process.

References

1. K. Creath, *Progress In Optics*, XXVI, 351 - 393, (1988).
2. P. Hariharan, B. F. Oreb, T. Eiju, *Appl. Opt.*, 26, 2504 - 2505, (1987).
3. C. Candler, *Modern Interferometers*, Hilger & Watts, 218 - 220, (1951).

[†] 'Wringing' takes place when two flat lapped surfaces are brought into contact in a sliding movement. It is generally accepted that a molecular adhesion results between the surfaces.

DESIGN NOTE

Interferometer light source and alignment aid using single-mode optical fibres

A J Lewis and D J Pugh*

Division of Mechanical and Optical Metrology, National Physical Laboratory,
Teddington TW11 OLW, UK

Received 9 April 1992, accepted for publication 8 June 1992

Abstract. A system has been developed using optical fibres to allow accurate positioning of a light source in the collimator of an interferometer. The design reduces the obliquity effect of the source by using single-mode fibres as the light source. Alignment of the interferometer is achieved using an autocollimation technique where one fibre detects the return spot. Up to three lasers can be launched into the collimator with no speckle in the image field

1. Introduction

When launching light into the collimator of a Twyman-Green length-measuring interferometer (Twyman and Green 1916, Dyson 1970) it is important to ensure that the light source is accurately positioned at the focus of the collimating lens and on the axis of the interferometer. An error in either the focus or off-axis adjustment will cause the wavefront to travel at an angle to the measurement axis. This causes an error in the length measured by the interferometer, termed an 'obliquity error', which is proportional to $s^2/2f^2$ where f is the focal length of the collimator and s is the distance by which the source is positioned off-axis (Bruce 1955). As an example, the obliquity error from a source positioned 1 μm off axis in a 1 m focal length collimator is approximately 5×10^{-7} or a length measurement error of 0.5 μm in 1 m. A similar error arises from the finite size of the source which also results in the wavefront travelling obliquely to the measurement axis. This error is proportional to $r^2/4f^2$ where r is the radius of the source (Bruce 1955).

To overcome these effects, a simple three-fibre system has been developed to launch up to three lasers into a collimator and to position them on axis and at the focus of the collimating lens.

2. Design

The system uses three single-mode optical fibres which have had the buffer coating removed from both ends. At one end the fibres are cemented into a tight bundle. The other end of each of the fibres is individually

mounted and polished (see figure 1). Each fibre in the bundle can serve two functions: it can act as the light source for the collimator when light from a laser is focused into the fibre core and, secondly, the fibre can be used to detect the return spot, when used in an autocollimation arrangement, for which another fibre is used as the light source (see figure 2). The numerical aperture of the fibres can be matched with the focal length and diameter of the collimator lens to achieve efficient illumination of the collimator.

There are two advantages of single-mode over multi-mode fibre. Firstly, the effective source diameter is much smaller, typically 10 μm contributing to a smaller

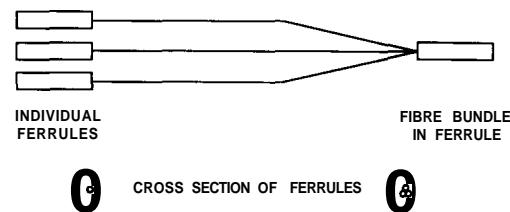


Figure 1. Three-fibre system

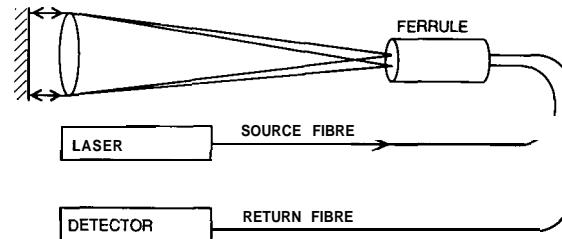


Figure 2. Autocollimation arrangement

* D J Pugh has now retired from NPL.

Design note

obliquity error, and secondly the problem of speckle in the image due to mode mixing in multimode fibre is eliminated since only one mode is transmitted by single-mode fibre.

In principle it is possible to use this technique with just one fibre acting as both source and detector though the extra optical components required, such as beam-splitters or couplers, could introduce losses which would make the detection of the return spot more difficult. Three fibres were used in this evaluation rather than two because a bundle of three fibres was easier to manufacture and mount, and there was a requirement to launch three lasers into the collimator.

3. Autocollimation

When used in the autocollimation arrangement of figure 2, the reference mirror of the interferometer is used to reflect the beam back to the source where one of the fibres in the bundle is used to detect the return spot. The fibre bundle is moved in three orthogonal directions and the intensity of the light incident on the detector fibre is monitored using a photodetector. When the detected intensity is maximized, the source and detector fibres are symmetrically positioned on either side of the axis of the interferometer, and at the principal focus. The off-axis position of the source is then half the separation of the fibres, which is typically less than 100 µm. This technique aligns the collimator with the reference arm of the interferometer. The expected obliquity error from this system is less than 5×10^{-9} or 5 nm in 1 m. After alignment, the detector fibre may be used to launch a third laser source.

4. Measurements

Measurements were made using a collimator of focal length 1500 mm, illuminated by a laser operating at 633 nm. Figure 3 shows the peak in the detected intensity as the fibre bundle was positioned radially and axially. These results were repeatable after coarse adjustment over several millimetres of travel. Correct collimation was checked using a shearing interferometer (Melles Griot 09SPM003) placed in the collimated beam at various points along its length. Assuming the achromatic collimator lens to be diffraction limited, the expected central maximum (Airy disc) of the return spot diffraction pattern should be ~25 µm in diameter (Hecht 1987), and should result in a peak of width ~20 µm when a 10 µm diameter fibre is scanned across the moving diffraction pattern, as occurs when the fibre bundle undergoes radial motion. This can be seen in figure 3(a). It is thought that the non-symmetrical peaks in the observed data are due to cross-talk from the adjacent fibre which becomes partially illuminated. When diffrac-

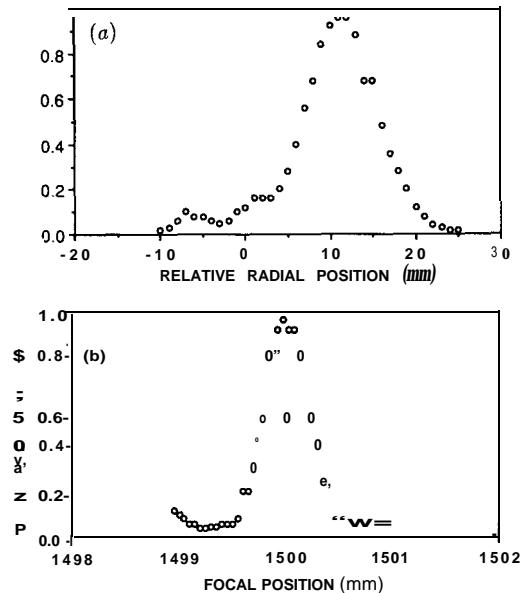


Figure 3. Detected intensity during (a) radial positioning and (b) axial positioning of the fibre bundle in normalized units.

tion theory is applied to an unaberrated circular pupil with defocus it predicts minima in the diffraction pattern, spaced at 1.1 mm along the focal axis. The results shown in figure 3(b) are consistent with the theory.

5. Conclusion

The single-mode fibre system provides a simple, efficient, speckle-free light source for an interferometer. The autocollimation arrangement using one fibre as a detector allows accurate repositioning of the light source, allowing the collimator beam to be aligned with the interferometer axis, whilst minimizing the obliquity effect due to the source.

Acknowledgments

The advice of Mr M Virdee on the design of the single-fibre ferrules and the polishing of the ferrules by the NPL Optical Workshop are gratefully acknowledged.

References

- Bruce C F 1955 The effects of collimation and oblique incidence in length interferometers Aust. J. Phys. **8** 224-40.
- Dyson J 1970 *Interferometry as a Measuring Tool* (London: The Machinery Publishing Company) pp 114-5
- Hecht E 1987 *Optics* 2nd edn (Reading, MA: Addison-Wesley) pp 416-21
- Twymann F and Green A 1916 *British Patent 103832*

Three wavelength phase-stepping interferometer for length measurement up to 1.5 m in a controlled environment

A. J. Lewis

Division of Mechanical and Optical Metrology
National Physical Laboratory, Teddington, Middlesex, TW11 0LW, UK.

1. Introduction

Length bars are widely used in industrial calibration laboratories as secondary length standards, and are becoming increasingly important as reference artefacts used to validate the performance of coordinate measuring machines. The highest grade length bars have their lengths calibrated at NPL in terms of the wavelengths of light emitted by calibrated frequency-stabilised lasers. A new interferometer is under construction at NPL to measure length bars up to 1.5 m in length to a target accuracy of 5 parts in 10^8 (50 nm in 1 m). The interferometer uses the techniques of three-wavelength phase-stepping interferometry in a sealed, temperature-controlled environment, see figure 1.

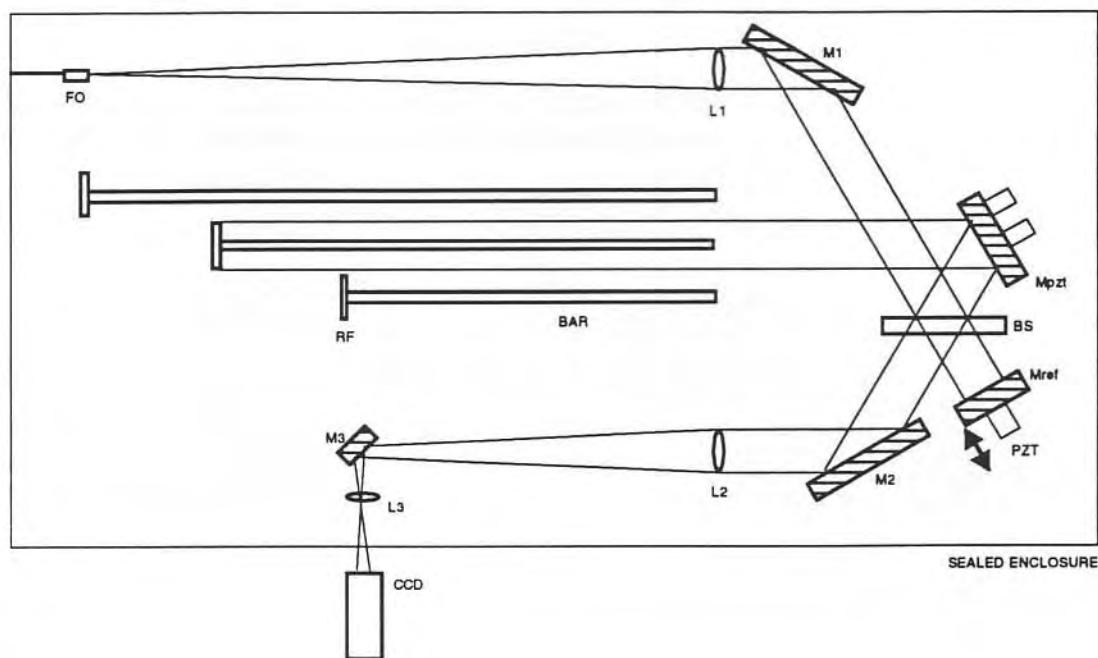


Figure 1 - optical layout of interferometer inside chamber
(L₁, L₂, L₃ - lenses; M₁, M₂, M₃, M_{ref}, Mpzt - mirrors; RF - reference flat;
BS - beamsplitter; FO - fibre optic source)

2. Refractive index

The refractive index along the optical path directly determines the laser wavelength. To minimise turbulence and drifts in refractive index, the interferometer has been constructed inside a sealed, temperature-controlled chamber. This ensures a uniform refractive index along the optical path which can be measured by sampling air from the chamber. The air is sampled using a PTFE pump which re-circulates the air back to the chamber. Measurements are made of the air temperature, pressure, humidity and carbon-dioxide content at the time each length measurement is made. The effect of each of these parameters is shown in table 1 for typical variations encountered in a temperature and humidity-controlled calibration laboratory.

Parameter	Typical value	Daily variation	Effect of variation on refractive index
temperature	20 °C	± 0.2 °C	± 2 × 10 ⁻⁷
pressure	1013.25 mbar	± 20 mbar	± 5 × 10 ⁻⁶
humidity	13 mbar	± 2 mbar	± 1 × 10 ⁻⁷
CO ₂ content	300 ppm	± 100 ppm	± 2 × 10 ⁻⁸

Table 1 - effect of air parameters on refractive index

3. Thermal control

The temperature of the chamber is stabilised by controlling the temperature of water circulating in pipes in the lid and baseplate of the chamber. All heat sources such as lasers and CCD camera are external to the chamber, with the light brought into the interferometer using single-mode optical fibres. The temperature of the air inside the chamber is stabilised to $20\text{ }^{\circ}\text{C} \pm 0.05\text{ }^{\circ}\text{C}$ and has a drift rate of less than $0.001\text{ }^{\circ}\text{C}$ per hour. The temperature drift of the length bar is similar and the temperature gradient along the bar is less than $0.01\text{ }^{\circ}\text{C}/\text{m}$, at $20\text{ }^{\circ}\text{C}$.

4. Thermal expansion

Although each bar has a nominal thermal expansion coefficient, α , of $10.7\text{ ppm}/^{\circ}\text{C}$, in practice this may vary by up to $1\text{ ppm}/^{\circ}\text{C}$ from bar to bar. The resulting uncertainty can cause an error if the bar is not measured or used at the standard temperature of $20\text{ }^{\circ}\text{C}$. To overcome this, each bar is measured in the interferometer at three temperatures over the range $15\text{ }^{\circ}\text{C}$ to $30\text{ }^{\circ}\text{C}$ and a value of α calculated.

5. Multiple-wavelength phase-stepping interferometry

Each bar is supported horizontally in the interferometer with a small reference flat attached to one end by 'wringing' (intra-surface attraction). The length of the bar is measured as a number of interference fringes at corrected wavelength λ'

$$L = (n + f)\lambda'$$

where n is an integer (fringe order) and f is a fraction of a fringe (fringe fraction). Values of f are measured by the interferometer for 3 wavelengths (633 nm, 543 nm and 612 nm) by using a five position phase-stepping algorithm (Hariharan *et al* 1987) to measure the surface form of the bar at each wavelength as a phase map. Each phase map contains measurements of the phase difference between light reflected from the surfaces of the bar and the reference flat. The method of exact fractions (Rolt 1929) is then used to combine the three phase measurements at the centre of the bar to calculate its length, based on an initial estimate to within $\pm 4.6\text{ }\mu\text{m}$ which is performed in another instrument.

6. Results and conclusions

Measurements have been made on bars ranging from 200 mm to 1000 mm in length. Repeatability of $\pm 1.9 \times 10^{-8}$ has been achieved over periods of up to 2 days. Comparisons with other interferometric measurements (Pugh and Jackson 1986) on bars up to 300 mm in length shows agreement to within $\pm 5 \times 10^{-8}$. This departure is thought to be due to different accuracies in determining the bar temperatures and refractive indices in the two instruments and also due to the difference in surface roughnesses of the two reference flats used.

7. References

- Hariharan P, Oreb B F and Eiju T 1987 *Digital phase-shifting interferometry: a simple error-compensating phase calculation algorithm* *Appl. Opt.* **26** 2504-2506
- Pugh D J and Jackson K 1986 *Automatic gauge block measurement using multiple wavelength interferometry* *Proc. SPIE* **656** 244-250
- Rolt F H 1929 *Gauges and Fine Measurement* (MacMillan & Co. : London) 49-50,212-213

END NOTE

This thesis was prepared on an Apple Macintosh SE/30 computer with a Radius Pivot display and printed on an Apple LaserWriter IIg. The text was prepared and set using Word (Microsoft). Graphs and illustrations were prepared in Cricket Graph (Cricket Associates), **Claris CAD** (Claris), **SuperPaint** (Silicon Beach) and *Mathematica* (Wolfram Research). Colour prints from the interferometer were printed on a Hewlett-Packard DeskJet 550C inkjet printer. Grey scale images were scanned at 16 grey levels on a Panasonic FXRS307 scanner, at 150 dpi resolution. Original video prints were obtained on a Mitsubishi P68B video printer. Photographs are courtesy of NPL Photographic Section and are © Crown Copyright. Colour and half-tone reproductions of diagrams and photographs were performed on a Canon colour copier. The text is set in 12 point Times-Roman with *10 point Times-Roman-Italic* (& **bold**) for figure and table captions. Chapter headings are set in 18 point **Helvetica bold**. Section headings are in **12 POINT BOLD CAPS** with sub-sections in **Bold**. Tables are set in 9 point Avant Garde. *Mathematica*-generated output contains Courier font and is imported as Encapsulated PostScript. Other diagrams were imported as either TIFF or PICT format. Computer code is represented in Monaco. The entire thesis, excluding colour diagrams, colour photographs and published papers occupies 13.5 Mb of disc space.

2002 Re-edition note: recreation of several plots in Microsoft Excel 2000 and PC Mathematica, and replacement of several Apple Macintosh fonts with nearest PC equivalent. Final correction, editing and preparation using Microsoft Word 2000 and Adobe Acrobat on a PC. No editing other than re-formatting for page breaks, re-creation of some images which did not import from Mac Word, image re-colouration, inclusion of references to significant changes, and some minor typo corrections.

Andrew Lewis, April 2002

“In the real world, absolute precision is an ideal that can never be reached”

R P Feynman

