

CRANFIELD UNIVERSITY

ANGUS BRIDGES

OPTICAL METROLOGY TO SUPPORT THE NEXT
GENERATION OF NANOPositionING

SCHOOL OF AEROSPACE, TRANSPORT AND
MANUFACTURING

Centre for Engineering Photonics

PhD

Academic Year: 2017–2021

Supervisors: Prof. Ralph P. Tatam, Dr. Thomas Kissinger,
Dr. Andrew Yacoot (NPL)

January 16, 2022

CRANFIELD UNIVERSITY

SCHOOL OF AEROSPACE, TRANSPORT AND
MANUFACTURING

Centre for Engineering Photonics

PhD

Academic Year: 2017–2021

ANGUS BRIDGES

Optical metrology to support the next generation of
nanopositioning

Supervisors: Prof. Ralph P. Tatam, Dr. Thomas Kissinger,
Dr. Andrew Yacoot (NPL)

January 16, 2022

© Cranfield University 2022. All rights reserved. No part
of this publication may be reproduced without the written
permission of the copyright owner.

Abstract

This thesis develops methods for the analysis and elimination of non-linearities, defined as errors in the measured displacement that are periodic with some harmonic of the illuminating wavelength, in homodyne displacement measuring optical interferometers. A particular focus is placed on phase-quadrature coating based interferometers.

A novel modelling technique is described, allowing the effects of cavity formation within networks of polarising optics to be analysed. This modelling technique is applied to investigate the sources higher order non-linearities introduced by unwanted cavity formation within the optics of the NPL plane plane mirror differential interferometer. Some non-linearities may be eliminated through improved optical design, however, this is not always possible, or practical. Three novel correction algorithms are therefore proposed, aiming to correct for non-linearities in situations where existing correction algorithms fail. The first, the two-wavelength non-linearity correction aims to correct for higher order non-linearities in cases where the displacement covers several optical fringes. Compared to a standard Heydemann ellipse fitting correction, the two-wavelength approach is shown to reduce estimated residual non-linearity amplitude from 84 pm to 11 pm. The second algorithm, the multiple intensity reference correction, aims to correct for non-linearities in cases where the displacement covers a small portion of the optical fringe. This algorithm is shown to reduce the amplitude of elliptical non-linearity harmonics to below the 10 pm level. Finally, the measurement arm normalisation correction also aims to correct for certain sources of higher order non-linearities, but reduced optical, electrical and data processing complexity as compared to the two-wavelength correction. Based on the findings of modelling work in combination with experimental non-linearity measurements, an improvement to an existing NPL interferometer design, utilising the measurement arm normalisation correction, is developed and tested. Peak-to-peak residual non-linearities of (23.0 ± 0.6) pm are achieved, reduced from (56.8 ± 0.4) pm by application of the measurement arm normalisation algorithm.

Acknowledgements

I would like to thank my partner, and parents, for their patience and support throughout this PhD.

Additionally I would like to thank my industrial supervisor at the National Physical Laboratory, Andrew Yacoot, and my Cranfield supervisors, Thomas Kissinger, and Ralph Tatam, for their guidance. I would like to specifically thank Andrew Yacoot for carrying out the x-ray interferometry measurements presented in this thesis. I would also like to thank the dimensional metrology group at the NPL as a whole, and specifically David Putland, for sharing his extensive experience with the operation, maintenance and repair of an occasionally cantankerous thin film coating plant. Furthermore, I would like to thank David Humphreys, also of the NPL, for suggesting the use of two-wavelength interferometry for non-linearity correction, leading to the development of the two-wavelength non-linearity correction methodology described in this thesis.

I gratefully acknowledge the funding of the PhD by the Engineering and Physical Sciences Research Council via grant EP/R511894/1 (Project No. 2199198) and the UK government Department for Business, Energy and Industrial Strategy via the National Measurement System Engineering Measurement Programme. I thank Queensgate Instruments for the loan of a NPS-X-15A nanopositioning stage, and for funding conference attendance. I also thank Santander UK plc for funding conference travel through the Santander Mobility Award programme.

Contents

Abstract	iii
Acknowledgements	iv
Contents	v
List of figures	ix
List of tables	xvi
List of acronyms and abbreviations	xviii
List of symbols	xx
1 Introduction	1
1.1 Displacement measuring optical interferometry	1
1.2 Aims and objectives	3
1.3 Thesis contents	3
2 Review of displacement measuring interferometry	5
2.1 Introduction	5
2.2 Basic principles of optical interferometry	6
2.2.1 Homodyne	6
2.2.2 Heterodyne	9
2.2.3 Other approaches	11
2.3 The state-of-the-art in displacement measuring interferometry	13
2.3.1 Homodyne	13
2.3.2 Heterodyne	17
2.3.3 Other approaches	22
2.4 Discussion and comparison of techniques	25
2.5 Summary	28
3 Uncertainty sources in homodyne interferometry	30
3.1 Introduction	30
3.2 Displacement measurement uncertainty budget	31
3.2.1 The measurement problem	31
3.2.2 Mathematical model	33
3.2.3 Contributory variances	33

3.2.4	Combined standard uncertainty	39
3.3	Phase measurement uncertainty budget	40
3.3.1	The measurement problem	40
3.3.2	Mathematical model	43
3.3.3	Contributory variances	44
3.3.4	Combined standard uncertainty	55
3.4	Shot noise limited homodyne interferometry	55
3.4.1	Detection noise	56
3.4.2	Intensity normalisation	57
3.4.3	Monte-Carlo simulation	61
3.5	Discussion and comparison of limiting factors	62
3.5.1	Displacement measurement uncertainty	62
3.5.2	Phase measurement uncertainty	64
3.6	Summary	70
4	Phase-quadrature coatings for homodyne interferometry	72
4.1	Introduction	72
4.2	Conservation of energy considerations	73
4.3	Thin film modelling	78
4.4	Phase-quadrature coating design	80
4.5	Coating fabrication	85
4.6	Coating optimisation and evaluation	87
4.6.1	Quadrature phase measurement	87
4.6.2	Reflectivity and transmission measurement	90
4.6.3	Optimisation	90
4.7	Discussion	94
4.8	Summary	95
5	Modelling displacement measuring interferometers	96
5.1	Introduction	96
5.2	Theory of the combined scattering and transfer matrix model	98
5.3	Matrices of common optical components	101
5.3.1	Free space propagation	102
5.3.2	Material interface	102
5.3.3	Linear polariser	102
5.3.4	Waveplate	103
5.3.5	Faraday rotator	104
5.3.6	Non-polarising beam splitter	104
5.3.7	Polarising beam splitter	105
5.3.8	Mirror	107
5.4	Modelling a polarisation splitting homodyne interferometer	108
5.4.1	Model construction	108
5.4.2	Model results	111
5.5	Discussion	114
5.6	Summary	115

6	Correction of non-linearities in homodyne interferometry	116
6.1	Introduction	116
6.2	Review of existing techniques	117
6.2.1	Ellipse-fitting algorithm comparison	117
6.2.2	Non-ellipse fitting algorithms	132
6.3	Two-wavelength non-linearity correction	134
6.3.1	Introduction	134
6.3.2	Theory and data processing methods	135
6.3.3	Experimental methods	140
6.3.4	Results	144
6.3.5	Discussion	145
6.3.6	Conclusions	150
6.4	Multiple intensity reference non-linearity correction	150
6.4.1	Introduction	150
6.4.2	Theory	151
6.4.3	Experimental methods	158
6.4.4	Results	160
6.4.5	Discussion	165
6.4.6	Conclusion	167
6.5	Measurement arm intensity reference normalisation	167
6.6	Summary	173
7	Displacement measuring phase-quadrature coating based interferometry	176
7.1	Introduction	176
7.2	Modelling the NPL plane mirror differential interferometer	177
7.3	Survey of existing interferometers	181
7.4	Improved optical design	185
7.5	Non-linearity evaluation	188
7.5.1	Methods	189
7.5.2	Results	192
7.6	Discussion	198
7.7	Summary	201
8	Conclusions and future work	203
8.1	Conclusions	203
8.2	Future work	208
8.2.1	Phase-quadrature coatings	208
8.2.2	Uncertainties	208
8.2.3	Two-wavelength non-linearity correction	208
8.2.4	Measurement arm normalisation non-linearity correction	210
8.2.5	Improved NPL plane mirror differential interferometer	210
A	Derivation of the covariance of Heydemann corrected quadrature signals	214
B	Phase uncertainty propagation calculations	216

C	Derivation of the covariance of the normalised quadrature signals	220
D	Derivation of the outputs of a phase-quadrature coating based homodyne interferometer	223
E	Outline of the thin film transfer matrix model	226
	References	232

List of figures

1.1	Illustration of the basic operating principle of a homodyne displacement measuring interferometer.	2
2.1	Schematic of plane mirror Michelson interferometer.	7
2.2	Phase measurement from quadrature signals.	8
2.3	Schematic of Fabry-Perot interferometer operated in reflection.	12
2.4	Schematic illustrating the operating principle of the multiple reflection path length multiplication interferometer.	13
2.5	Schematic representation of the two-wavelength common path homodyne interferometer.	14
2.6	Schematic representation of the AIST plane mirror differential homodyne interferometer.	15
2.7	Schematic representation of the ISI plane mirror differential homodyne interferometer.	16
2.8	Schematic representation of the NPL plane mirror differential homodyne interferometer.	17
2.9	Schematic representation of the PTB plane mirror differential heterodyne interferometer.	19
2.10	Schematic representation of the NIST plane mirror differential heterodyne interferometer.	20
2.11	Schematic representation of the Delft University of Technology plane mirror heterodyne interferometer.	21
2.12	Schematic representation of the McRae plane mirror heterodyne interferometer.	22
2.13	Schematic representation of the MIKES plane mirror heterodyne interferometer.	22
2.14	Schematic representation of the AIST Fabry-Perot interferometer.	24
2.15	Schematic representation of the dual Fabry-Perot interferometer.	25
3.1	Geometric representation of cosine (A) and Abbé (B) errors. In both cases the mirror is displaced a distance x_t along the displacement axis and a displacement x_m is measured along the measurement axis. Angles θ and α are equal for comparison. The effects of reflection from the mirror at non-normal incidence are neglected here for ease of comparison.	39
3.2	Demonstration of phase non-linearities introduced by higher order non-linearities without radial distortion.	52

3.3	Demonstration of the validity of variance and covariance estimates in intensity normalised quadrature signals.	62
3.4	Graphical illustration of uncertainty limits in homodyne interferometry.	65
3.5	Displacement uncertainties in an idealised shot noise limited phase-quadrature interferometer.	66
3.6	Variation of the maximum displacement uncertainty with measurement bandwidth and illuminating optical power.	67
3.7	Variation of the maximum displacement uncertainty with the intensity reference pick off reflectivity β_r	68
3.8	Variation of the maximum displacement uncertainty with quadrature phase shift δ	69
3.9	Demonstration of the phase distribution scaling effects of ellipse fitting corrections.	69
3.10	Variation of the maximum displacement uncertainty with reflectivity to transmission ratio.	70
4.1	Principle of operation of phase-quadrature coatings	73
4.2	Schematic of Michelson interferometer morphologies.	74
4.3	Geometry of thin film transfer matrix modelling methodology.	79
4.4	Variation of phase-quadrature coating properties with respect to changing layer thicknesses.	82
4.5	Angular sensitivity of phase-quadrature coatings.	83
4.6	Variation of phase-quadrature coating properties with respect to changing refractive index.	83
4.7	Sensitivity of phase-quadrature coating parameters to changes in refractive index and layer thicknesses.	84
4.8	Literature values of the complex refractive index of thin gold films.	85
4.9	E-beam gun schematic.	85
4.10	Schematic representation of quadrature phase measurement interferometer.	87
4.11	Example of poor quadrature measurement data resulting from beam precession.	88
4.12	Schematic representation of bench top reflectometer.	90
4.13	Optimisation plane fits to quadrature and reflectivity to transmission ratio measurements.	91
4.14	Optimisation progress of phase-quadrature coatings.	92
4.15	Photograph showing a number of the phase-quadrature beam splitter cubes fabricate in the course of this PhD, for applications across the NPL.	93
5.1	Schematic of the electric field components at a four port beam splitter.	100
5.2	Schematic of the electric field components across two plane parallel interfaces.	101
5.3	Schematic of the electric field components throughout a homodyne Michelson interferometer with polarising quadrature detection.	109
5.4	Ideal outputs of polarising quadrature interferometer.	111

5.5	Distorted outputs of polarising quadrature interferometer with quadrature phase and gain ratio errors.	112
5.6	Distorted outputs of polarising quadrature interferometer with mild back reflections.	113
5.7	Distorted outputs of polarising quadrature interferometer with extreme back reflections.	114
6.1	Estimated displacement uncertainty for varying numbers of fit points after applying the algorithms given in table 6.1 evaluated at the maximum error point within a fringe for an illuminating power of 1 mW. .	122
6.2	Estimated displacement uncertainty for varying numbers of fit points after applying the algorithms given in table 6.1 evaluated at the maximum error point within a fringe for an illuminating power of 0.1 mW.	123
6.3	Estimated displacement uncertainty for varying quadrature phase shifts after applying the algorithms given in table 6.1 evaluated at the maximum error point within a fringe for an illuminating power of 1 mW.	123
6.4	Estimated displacement uncertainty for varying phase quadrature coating reflectivity to transmission ratio after applying the algorithms given in table 6.1 evaluated at the maximum error point within a fringe for an illuminating power of 1 mW.	124
6.5	Estimated displacement uncertainty after applying the algorithms given in table 6.1 evaluated at the maximum error point within a fringe for an illuminating power of 1 mW for sub-fringe displacements.	125
6.6	Estimated displacement uncertainty after applying the algorithms given in table 6.1 evaluated at the maximum error point within a fringe for an illuminating power of 0.1 mW for sub-fringe displacements.	125
6.7	Repeat of the simulation shown in figure 6.5, restricted to larger fit ranges.	126
6.8	Average displacement errors for varying quadrature phase shifts when fitting one half of an optical fringe after applying the algorithms given in table 6.1 evaluated at the maximum error point within a fringe for an illuminating power of 1 mW.	126
6.9	Average displacement errors for varying phase quadrature coating reflectivity to transmission ratio when fitting one half of an optical fringe after applying the algorithms given in table 6.1 evaluated at the maximum error point within a fringe for an illuminating power of 1 mW.	127
6.10	Displacement uncertainties calculated with the described Monte-Carlo methodology for the LS, ODR and BLUE algorithms for full fringe fits including 100 data points.	129
6.11	Displacement uncertainties calculated with the described Monte-Carlo methodology for the LS, ODR and BLUE algorithms for full fringe fits including 20 data points.	129

6.12	Displacement uncertainties calculated with the described Monte-Carlo methodology for the LS, ODR and BLUE algorithms for half fringe fits including 100 data points.	130
6.13	Displacement uncertainties calculated with the described Monte-Carlo methodology for the LS, ODR and BLUE algorithms for quarter fringe fits including 100 data points.	130
6.14	Basic principles of the two-wavelength non-linearity correction methodology, illustrated here for the simple case of a linear movement, as applied to modelled data for $\lambda_a = 633\text{ nm}$ and $\lambda_b = 785\text{ nm}$. A - displacement is measured with interferometer 'a', operating at λ_a . B - displacement is measured simultaneously with interferometer 'b', operating at λ_b . C - the difference between the 'a' and 'b' displacements is calculated. D - a Fourier transform is applied to the displacement difference and windowed around spatial harmonics of λ_a . E - estimated non-linearity errors in the λ_a displacement are recovered by inverting the windowed Fourier transform of the displacement difference. F - estimated non-linearity errors are subtracted from displacement a, resulting in a displacement measurement with reduced non-linearity. Edge effects are visible, resulting from a non-integer number of periods being included in the Fourier transform.	136
6.15	Demonstration of the effect of multiple displacement values at each displacement point on the interpolation process for a measurement including a linear 50 pm drift over the full measurement period between the λ_a and λ_b displacement measurements.	139
6.16	Schematic of the two-wavelength interferometer. RR, retro-reflector; PQBS, phase-quadrature beam splitter; 90:10 BS, reference pick off beam splitter; Ref, intensity reference photodiode; Sin, sine quadrature photodiode; Cos, cosine quadrature photodiode.	141
6.17	Stages of the two-wavelength non-linearity correction process for three displacement profiles (A , B and C) as applied to correction of the λ_a displacement measurement. Offsets have been applied to the displacement profiles in A , B and C for clarity. D , E and F show the difference between the λ_a and λ_b interferometer displacement measurements. G , H and I show the DFT of the displacement difference following the interpolation, clipping and tiling process described in the text, with interpolation performed relative to the λ_a displacements. J , K and L show the errors in the λ_a displacement measurements, relative to the stage position, for the three displacement profiles, and for both Heydemann corrected and Heydemann and two-wavelength corrected data. M , N and O show DFTs of the data presented in J , K and L , the error harmonics relative to the stage position.	143
6.18	Variation of the non-linearity amplitude, estimated from the sum of the displacement error DFT peaks for the first three even harmonics, with displacement range.	145

6.19	Schematic of the phase quadrature coating based homodyne interferometer. FC, fibre collimator; PQBS, phase quadrature beam splitter; Pol, polariser; BS, 50:50 beam splitter; RR_{mes} and RR_{ref} , measurement and reference retroreflectors; I_{mes} , measurement beam reference intensity; I_{ref} , reference beam reference intensity; I_{sin} , sine quadrature output; I_{cos} , cosine quadrature output.	151
6.20	Flowchart demonstrating the MIR and constrained Heydemann (CH) non-linearity correction methodologies.	153
6.21	Comparison of the full-fringe performance of the Heydemann, beam shuttering, multiple intensity reference (MIR) and constrained Heydemann (CH) non-linearity correction algorithms when applied to a displacement covering $15\text{ }\mu\text{m}$ with a sample spacing of 15 nm . Inset shows region of interest. The Heydemann, MIR, and CH techniques overlap almost exactly away from harmonics of 632.8 nm	161
6.22	Comparison of the full-fringe performance of the Heydemann, beam shuttering, multiple intensity reference (MIR) and constrained Heydemann (CH) non-linearity correction algorithms in the presence of wavefront distortion, when applied to a displacement covering $15\text{ }\mu\text{m}$ with a sample spacing of 15 nm . All techniques plotted overlap almost exactly away from harmonics of 632.8 nm	162
6.23	Comparison of the full-fringe performance of the Heydemann, multiple intensity reference (MIR) and constrained Heydemann (CH) non-linearity correction algorithms in the presence of wavefront distortion, when applied to a displacement covering $15\text{ }\mu\text{m}$ with a sample spacing of 15 nm . Inset shows region of interest. All techniques plotted overlap almost exactly away from harmonics of 632.8 nm	163
6.24	Time stability of MIR and CH calibration process, compared to the Heydemann correction in the full-fringe case.	163
6.25	Sub-fringe performance of the Heydemann, MIR and CH algorithms.	164
6.26	Schematic of NPL plane mirror differential interferometer design with added intensity references. PQBS, phase-quadrature beam splitter; PBS, polarising beam splitter; QWP, quarter waveplate; M_{ref} , reference mirror; M_{mes} , measurement mirror.	168
6.27	Non-linearity resulting from introducing a $\lambda/300$ retardance error in the QWP, combined with a 1000:1 extinction ratio in the PBS, of the interferometer depicted in figure 6.26, for an illuminating wavelength of 632.8 nm , after applying a Heydemann correction.	170
6.28	Modelled reference and measurement arm normalised non-linearities in displacement (A) and spatial frequency (B) space, calculated relative to true displacement x , including the effects of unwanted back reflections as described in the text.	171
6.29	Modelled reference and measurement arm normalised non-linearities in displacement (A) and spatial frequency (B) space, calculated relative to true displacement x , including the effects of 0.01% back reflections from all four detectors, with other conditions as described for figure 6.27.	172

6.30	Modelled reference and measurement arm normalised non-linearities, calculated relative to true displacement x , including the effects of a 0.01 % back reflection from the interferometer input, with other conditions as described for figure 6.27. B shows the same data as A , with the ordinate axis magnified around the region of interest.	172
7.1	Schematic of NPL plane mirror differential interferometer with nodes labelled.	177
7.2	Effects of a 1° QWP retardance error on NPL plane mirror differential interferometer non-linearities.	178
7.3	Effects of QWP rotation angle and retardance on NPL plane mirror differential interferometer non-linearities.	179
7.4	Effects of varying PBS reflection extinction ratio on NPL plane mirror differential interferometer non-linearities.	180
7.5	Effects of a 0.1 % reflection from the front surface of the PBS on NPL plane mirror differential interferometer non-linearities.	180
7.6	Effects of a 0.1 % reflection from the front surface of the PBS combined with a 1° QWP retardance error on NPL plane mirror differential interferometer non-linearities.	181
7.7	Results of a survey of the peak-to-peak non-linearity of ten NPL plane mirror differential interferometers (frequency space).	183
7.8	Results of a survey of the peak-to-peak non-linearity of ten NPL plane mirror differential interferometers.	184
7.9	Schematic of the improved NPL plane mirror differential interferometer.	186
7.10	Photograph of improved NPL plane mirror differential interferometer prototype with key components labelled, as in figure 7.9. 3D printed beam baffles as discussed in the text (yellow in colour) are also visible. The Glan-Thompson polariser mounting coaxially with the fibre optic collimator in this photograph was later removed, and was not used when gathering the data presented in this work.	187
7.11	Principle of operation of the NPL XRI facility.	190
7.12	Photograph of the prototype improved plane mirror differential interferometer in operation against the NPL XRI facility.	191
7.13	Experimental measurement of reference and measurement arm normalised non-linearities, in the presence of back reflections.	193
7.14	Experimental measurement of reference and measurement arm normalised non-linearities.	194
7.15	Fourier transforms of the four output voltage signals of the prototype interferometer, with back reflections present, calculated as described in the text.	195
7.16	Fourier transforms of the four output voltage signals of the prototype interferometer, with back reflections minimised, calculated as described in the text.	196
7.17	Non-linearity errors in the prototype improved plane mirror differential interferometer, measured against the NPL XRI facility.	197

- 8.1 Proposed spatially separated two-wavelength interferometer, following the improved NPL plane mirror differential interferometer design. 209

List of tables

2.1	Comparison of state-of-the-art displacement measuring interferometer designs. Non-linearities are quoted as peak-to-peak values, converted from the published figures assuming a sinusoidal error distribution when required. Similar interferometer designs developed within the same research group have been grouped, and the year at which the quoted (minimum) non-linearity was reached is given. Entries are sorted by non-linearity first, and year second.	26
3.1	Standard values of atmospheric parameters and corresponding refractive index sensitivities evaluated at 633 nm.	35
3.2	Properties of a Hamamatsu S2386-18K/L photodiode, taken as an example of a typical photodiode, required to calculate the standard deviation of the photocurrent.	60
3.3	Typical values for parameters affecting interferometer uncertainties assuming a thermally stabilised single frequency HeNe operating at 632.8 nm with a beam waist radius of 1 mm. The values given are approximate, and are intended only to be used to give rough estimates of the measurement uncertainty.	63
3.4	Approximate uncertainty contributions to the measured displacement assuming an initial refractive index of unity, a change in refractive index during the measurement of 1×10^{-6} , a wavefront correction factor of zero, and other parameters as defined in table 3.3.	64
4.1	Displacement uncertainty due to shot noise contributions for the interferometer designs described in section 4.2 assuming an illuminating power of 1 mW at 633 nm with idealised unity quantum efficiency photodiodes and a 1 MHz measurement bandwidth.	78
4.2	Literature values of refractive index of materials used in phase-quadrature coatings at 632.8 nm.	81
4.3	Ideal thicknesses of chromium and gold layers and resulting absorption coefficients in a phase-quadrature coating operating at 632.8 nm, using literature values of refractive index, as calculated with the previously described thin film modelling technique.	82

4.4	Average values and the mean-centred range of values in which 95 % of coatings are expected to fall for the separately and batch coated phase-quadrature cube beam splitter measurements presented in figure 4.14. The range in which 95 % of coatings are expected to fall has been calculated assuming a student t distribution, as described in the text. All values are given for S polarised light.	93
6.1	Ellipse fitting algorithms included in the comparison.	118
6.2	Summary of the non-linearity correction techniques discussed in this chapter. Displacement ranges are approximate, given to the nearest order of magnitude, and assume a single pass interferometer operating at 632.8 nm. It should be noted that the minimum displacement range is not a sharp cut off, and that the short range performance of all algorithms considered will be dependent on optical wavelength, interferometer pass number, and noise levels.	173
7.1	Key properties of Queensgate Instruments X15-A nanopositioning stage [191]. Parameters are quoted across the full closed loop range of the stage, in the 'slow' control loop pre-set employed in this work. .	189

List of acronyms and abbreviations

NPL	National Physical Laboratory
XRI	X-ray Interferometry
HeNe	Helium-Neon
CMI	Czech Metrology Institute
AIST	National Institute of Advanced Industrial Science and Technology
ISI	Institute of Scientific Instruments of the Czech Academy of Sciences
AC	Alternating Current
DC	Direct Current
PTB	Physikalisch Technische Bundesanstalt
NIST	National Institute of Standards and Technology
QWP	Quarter Waveplate
AFM	Atomic Force Microscopy
JCGM	Joint Committee for Guides in Metrology
GUM	Guide to the expression of Uncertainty in Measurement
BIPM	the International Bureau of Weights and Measures
RMS	Root Mean Squared

PQBS	Phase Quadrature Beam Splitter
PBS	Polarising Beam Splitter
E-beam	Electron beam
QCM	Quartz Crystal Monitor
UV	Ultraviolet
FFT	Fast Fourier Transform
LS	Least Squares
DEF	Direct Ellipse Fit
BLUE	Best Linear Unbiased Estimator
ODR	Orthogonal Distance Regression
AML	Approximate Maximum Likelihood
FPGA	Field Programmable Gate Array
PF	Peak Finding
LEP	Low Error Point
KF	Kalman Filter
MIR	Multiple Intensity Reference
CH	Constrained Heydemann
BS	Beam Splitter
PLA	Polylactic Acid
PZT	Piezoelectric Transducer

List of symbols

δ	Quadrature phase (rad)	x, 57, 58, 61, 66, 67, 69, 74–77, 225
λ	Wavelength (m)	xii, xiii, 7, 31, 32, 36, 37, 56, 136–144, 146, 147, 149, 160–162, 170, 184, 192, 197, 199, 200, 227, 228
I	Optical intensity (W m^{-2})	xiii, 6–8, 10, 73–77, 151, 152, 155, 223–225
φ	Interferometric phase (rad)	1, 8, 9, 31–34, 43–46, 51, 54, 55, 61, 62, 64, 65, 72, 74–77, 98, 120, 121, 132, 142, 152, 154, 156, 157, 159, 160, 169, 170, 192, 216–219, 225
E	Scalar electric field (V)	6, 7
k	Wavenumber (m^{-1})	6–8, 10, 80, 98, 102, 112, 223–230
ω	Optical angular frequency (rad s^{-1})	6, 7, 10, 98, 223–227
t	Time (s)	6, 7, 10, 98, 223–227
f	Frequency (Hz)	11, 17, 32–34, 37, 62–64
x	Displacement (m)	31–34, 37, 38, 62, 64, 137–140, 142, 160, 170, 192, 197

c	Speed of light in the vacuum (m s^{-1})	32–34, 37, 56, 62, 224
n	Refractive index	32–34, 37, 62–64, 227
σ	Standard deviation	33, 34, 44–46, 48, 49, 54, 57–60, 62, 64, 65, 121, 214– 222
P	Optical power (W)	40, 56–59, 61, 169, 171
i	Electrical current (A)	40, 56–58
V	Voltage (V)	40, 44, 45, 47, 48, 58, 152, 155, 159, 160, 169, 195, 199, 220–222
S	Measured signal (A.U.)	40–42, 44–46, 48, 51, 55, 56, 58, 60, 61, 66, 154–158, 169, 220–222
C	Corrected quadrature signal (A.U.)	43, 51, 157–160, 169
N	Integer count	43, 56, 57, 65
e	Error	44, 46, 54, 197, 216
Δf	Bandwidth (Hz)	56–58
χ	Photodiode responsivity (A W^{-1})	56–58
q	Charge on an electron (C)	56–58
h	Planck constant ($\text{m}^2\text{kg/s}$)	56
QE	Quantum efficiency	56
μ	Mean	57, 59, 60, 214, 215, 220– 222
k_B	Boltzmann constant ($\text{m}^2\text{kgs}^{-2}\text{K}^{-1}$)	57, 58
T	Temperature (K)	57, 58
R	Resistance (Ω)	57, 58
G	Gain coefficient (A.U.)	58–60

$\mathfrak{R}$	Intensity reflectivity coefficient	61, 66, 67, 74–78, 80, 120, 152, 154, 155, 224, 225
$\mathfrak{T}$	Intensity transmissivity coefficient	61, 66, 67, 74–77, 80, 120, 152, 154, 155, 224, 225
φ_t	Phase change upon transmission (rad or °)	72, 73, 80
φ_r	Phase change upon reflection (rad or °)	72, 73, 80
$\mathfrak{r}$	Amplitude reflectivity coefficient	73, 80, 102, 105, 106, 112, 228–230
$\mathfrak{t}$	Amplitude transmissivity coefficient	73, 80, 102, 105–107, 112, 228–230
Δ	Phase difference (rad)	76, 77
d	Layer thickness (m)	79, 102, 112, 230
$\tilde{n}$	Complex refractive index	79, 227–229
$\mathbf{T}$	Transfer matrix	78–80, 100, 101, 110, 111, 230, 231
$\mathbf{P}$	Propagation matrix	79, 101, 230, 231
$\tilde{E}$	Complex scalar electric field (V)	79, 98, 99, 227–231
$\tilde{\mathbf{E}}$	Complex electric field vector (V)	98, 227
$\tilde{\mathbf{A}}$	Complex electric field amplitude vector (V)	98, 227
A	Amplitude	98, 223, 226
$\tilde{A}$	Complex scalar electric field amplitude (V)	98, 112, 227
a	‘Forward’ propagating complex electric field vector component (V)	99–106, 108–111
b	Counter propagating complex electric field vector component (V)	99–106, 108–112
$\mathbf{Q}$	Scattering matrix	99, 100, 103, 104, 106–109
θ	Rotation angle (rad or degree)	102–104, 107
ρ_{ex}	Polarisation extinction ratio	102, 103, 105–107

α	Loss coefficient	102–107
$\mathbf{R}$	Rotation matrix	103, 104, 107
η	Waveplate retardance (rad or °)	103, 104
ψ	Polarisation rotation angle (rad or degree)	104
ζ	Intensity reflectivity to transmission ratio	104, 105
$\mathbf{U}$	Rotation sub-matrix	107
δ'	Quadrature phase error (rad)	152, 154–158
$\mathbf{M}$	General matrix	156, 158
ϵ_0	Vacuum permittivity (F m^{-1})	223, 224
$\mathbf{E}$	Electric field vector (V)	223, 224, 226, 227
$\mathbf{k}$	Wavevector (m^{-1})	223
$\mathbf{r}$	Position vector (m)	223
$\mathbf{S}$	Poynting vector (W m^{-2})	224, 225
μ_0	Vacuum permeability (H m^{-1})	224
$\mathbf{B}$	Magnetic flux density vector (T)	224
κ	Imaginary component of refractive index.	227

Chapter 1

Introduction

1.1 Displacement measuring optical interferometry

Traceable displacement measurement is of vital importance to many manufacturing processes and research tasks, with traceability back to the SI definition of the metre [1] required for any meaningful comparison of measurements to be made. This is no less true of nanoscale measurements than of those made at the more familiar macroscopic scales, and an increasing demand for dimensional metrology at the nanoscale has been predicted [2–5]. This increasing demand for nanoscale metrology is driven both by the requirement to measure the form of nanoscale structures [5], and by the need to characterise and calibrate nanopositioning devices [6].

A range of technologies are capable of traceable measurements with sub-nanometre uncertainties, however, all are ultimately traceable back to the wavelength of light and therefore, with the exception of large area grating period measurement, to displacement measuring optical interferometry [7]. A secondary realisation of the metre standard has been added to the *mise en pratique* [7], through the well characterised lattice parameter of silicon [8]. This secondary realisation is particularly well suited to sub-nanometre metrology, however, the lattice parameter of silicon is again ultimately traceable to measurements made with optical interferometry [9, 10]. Improved displacement measuring optical interferometers have therefore been identified as a key enabling technology for the advancement of the field of nanotechnology [4, 11].

At the nanoscale, as will be demonstrated in chapter 3, the measurement uncertainty of displacement measuring optical interferometers may be dominated by non-linearities, defined throughout this thesis as errors in the measured phase (and therefore displacement) that are periodic with harmonics of the wavelength of the illuminating light. The basic operating principle of a homodyne interferometer is illustrated in figure 1.1, along with the origins of non-linearities, which in homodyne interferometry are introduced by distortions of the ideally sinusoidal outputs of the interferometer. As the measurement retroreflector (RR_{mes}) is displaced relative to the fixed reference retroreflector (RR_{ref}) (**A**), sinusoidal intensity changes are observed at outputs x and y . Plotting these outputs against each other forms a Lissajous figure, with the instantaneous interferometric phase φ being measured by observing the angle of an individual data point in polar coordinates (figure 1.1

B). In the ideal case, x and y take the form of sine and cosine signals, resulting in a circular Lissajous. When the outputs deviate from this ideal form, the Lissajous is distorted, and a non-linearity error is present in the measured displacement (figure 1.1 **C**). The operating principle of homodyne interferometry is explained in more detail in section 2.2.1.

Commercial interferometers currently achieve non-linearities of $1 - 0.1$ nm, with some state-of-the-art interferometer designs reaching 10 pm peak-to-peak non-linearities under ideal circumstances [11]. As will be discussed in chapter 2, all interferometer designs currently capable of achieving these low non-linearities have limitations, in measurement range, cost, or optical complexity, and whether such low non-linearities can be achieved without careful optimisation referenced against a highly linear X-ray interferometer [12] remains unclear. Existing interferometer designs that are capable of picometre non-linearities are therefore poorly suited to many practical displacement measurement tasks, in particular the characterisation and calibration of the next generation of nanopositioning devices.

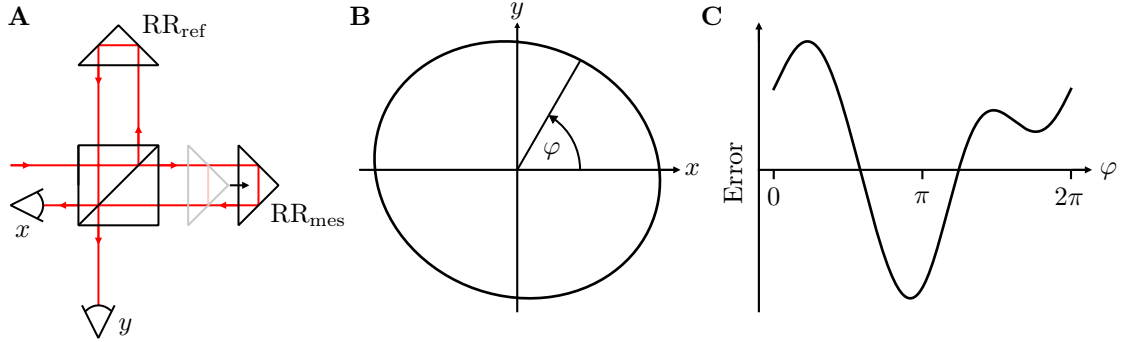


Figure 1.1: Illustration of the basic operating principle of a homodyne displacement measuring interferometer.

A research gap therefore exists for the investigation of improved displacement measuring optical interferometer designs capable of achieving picometre level non-linearities over the displacement ranges typically encountered when calibrating or characterising nanopositioning stages. Phase-quadrature coating-based homodyne (single optical frequency) interferometers, which introduce the phase change between quadrature signals required for bi-directional fringe counting by using an absorbing thin film coating, have been developed previously, however, such designs currently suffer from non-linearities introduced by parasitic reflections and unwanted formation of weak optical cavities. Phase-quadrature coating interferometer designs have well understood non-linearity sources in comparison to heterodyne systems and can be significantly simpler in design than comparable polarisation-based quadrature interferometers, reducing the potential for non-linearities to be introduced. The development of such designs is the focus of this thesis. A tool kit of novel techniques for the analysis, elimination, and correction, of non-linearities in phase-quadrature coating-based interferometers are developed, and the application to the development of improved phase-quadrature coating-based interferometry is demonstrated. Many of these techniques are equally applicable to other interferometer designs.

1.2 Aims and objectives

The overall aim of this PhD is to produce a tool-kit of techniques that allow homodyne displacement measuring interferometers to be constructed with repeatable non-linearities in the range of tens of picometres or less.

This will be achieved through the follow research objectives:

- Investigate the design and fabrication of phase-quadrature coatings.
- Develop a modelling technique capable of analysing higher order non-linearities in homodyne interferometers.
- Research novel approaches to the correction of higher order non-linearities in homodyne interferometers.
- Develop and test a phase-quadrature coating-based interferometer design that is capable of repeatably matching or exceeding the performance of existing state-of-the-art precision displacement measuring interferometers.

1.3 Thesis contents

This work begins with a brief overview of the fundamental principles of optical interferometry as applied to the problem of displacement measurement, focussing on the most commonly employed optical techniques within this field: homodyne, heterodyne and Fabry-Perot interferometry. Following this overview is a review of the current state-of-the-art as of the publication date of this thesis, again broken down into the four most commonly employed optical techniques.

Sources of uncertainty in homodyne interferometry, the technique employed throughout this work, are then discussed, primarily with reference to existing literature where possible, along with new derivations where required. Uncertainty sources are discussed from the viewpoint of precession interferometry with target uncertainties in the sub-nanometre range. The goal of this discussion is the creation of a summary of uncertainty sources that can be used to guide the creation of uncertainty budgets for individual interferometer systems, rather than to create an uncertainty budget for a specific interferometer. To conclude this chapter the dominant uncertainty sources under various conditions are discussed, with the results of this discussion guiding the remainder of the thesis.

Throughout this thesis, the focus is upon phase-quadrature coating-based interferometry, and the fourth chapter of this thesis discusses the properties, design and fabrication of such coatings. The fundamental limits of phase-quadrature coating-based interferometers are described. The thin film modelling techniques required to design such a coating are then detailed, followed by a discussion of the coating design process. The experimental fabrication of this coating design is then covered, along with experimental measurements of the coating properties, and the empirical optimisation process required to account for deviations from the ideal model.

The fifth chapter of this thesis deals with a novel modelling technique developed in the course of this work, with the primary goal of describing the effects of unwanted multiple reflections, sometimes termed ghost reflections, within interferometer optics. The theory behind this modelling technique is described and examples of the matrices employed by the technique to model various optical components commonly found in displacement measuring interferometers are given. The application of this modelling technique to a polarisation-based homodyne interferometer is demonstrated, and compared to an exact analytic solution.

Whilst in some cases non-linearities in homodyne interferometers may be eliminated through improved optical design, in other cases such improvements are either not possible, or not practical. As such it is desirable to correct for non-linearities in interferometric displacement measurements where they do occur. This is the topic of the sixth chapter of this thesis. Firstly, a review of non-linearity correction techniques proposed in the literature is carried out. This review of existing techniques is followed by a discussion of three novel non-linearity correction techniques developed in the course of this work. Finally, a summary is presented of both the techniques proposed in the literature, and those developed in this work, and the circumstances under which each technique is best applied are highlighted.

Combining the foundations laid in the previous chapters, the seventh chapter of this thesis describes the development of a plane mirror differential interferometer design, based upon an existing National Physical Laboratory (NPL) design. The modelling technique described in chapter 5 is combined with a novel non-linearity correction approach described in chapter 6 to identify non-linearity sources in the existing NPL design, and develop an improved design. The non-linearities of the improved design are evaluated against the NPL X-ray Interferometry (XRI) facility.

The conclusions of this work form the final chapter, along with suggestions for future research avenues within the field.

Chapter 2

Review of displacement measuring interferometry

2.1 Introduction

Optical interferometry finds a wide range of applications, a review of which may be found in, for example, [13]. This thesis is concerned only with displacement measuring interferometer designs, specifically those that are capable of measuring displacements over ranges of millimetres or greater, with sub-nanometre uncertainties, although measurements at the nanoscale are also of interest. Displacement measuring interferometer designs fulfilling the above criteria fall into two main categories: homodyne designs, employing a single optical frequency, and heterodyne designs, where beams having different optical frequencies are interfered. This chapter will first outline the basic operating principles of homodyne and heterodyne displacement measuring interferometers. Other displacement measuring interferometer approaches that are less commonly employed for measurements in this regime will then be briefly discussed.

Following this overview of the working principles of displacement measuring optical interferometry, the current state-of-the-art will be reviewed in section 2.3. Section 2.3 will focus on describing the performance and key features of state-of-the-art optical interferometer designs, with detailed comparison reserved for section 2.4. As will be elaborated upon in chapter 3, many of the uncertainty sources affecting interferometric measurements of displacement will affect all interferometer designs approximately equally if operated under similar conditions, for example, optical wavelength uncertainty. One important uncertainty contribution, particularly when sub-nanometre uncertainties are required, that is highly dependent on the exact design of an interferometer system is the non-linearity in the measurement of the change in the interferometric phase. As such, the focus will be upon comparing the non-linearity of state-of-the-art displacement measuring interferometers, with the review primarily restricted to interferometer designs that have been shown to achieve residual non-linearities, after the application of any non-linearity correction algorithm, of less than 100 pm. Wider ranging discussion of uncertainty sources in homodyne interferometry is reserved for chapter 3.

Only interferometer designs with publicly available non-linearity specifications

will be discussed, with the result that this review is restricted to interferometers developed in universities and national measurement institutes. Commercial displacement measuring interferometers are also available, however, in many cases these interferometers are intended for calibration tasks, for example of machining equipment, where sub-nanometre uncertainty is not required. Many commercial interferometers therefore exhibit non-linearities in the range of hundreds of picometres to nanometres [14–17], and as such are outside the scope of this review. Commercial interferometers targeting higher accuracy applications are also available, for example the SIOS Meßtechnik SP 5000 DI, with a manufacturer quoted displacement resolution of 5 pm [18], however, due to the lack of published non-linearity specifications, it is not possible to meaningfully review the performance of these interferometers here.

Finally, the strengths and weaknesses of the reviewed interferometer designs will be discussed, along with the strengths and weaknesses of the various approaches to displacement measurement with optical interferometry. Key properties of leading interferometer designs will be identified, so that these properties may be incorporated into future designs. Research areas where the literature suggests significant performance improvements are possible will also be identified.

2.2 Basic principles of optical interferometry

2.2.1 Homodyne

Homodyne interferometry is the simplest of the techniques considered here. Beams of nominally identical optical frequency are interfered, and the intensity of the resulting interference pattern measured directly. A Michelson interferometer configuration [19] is depicted in figure 2.1, in which the source beam from the laser is split into measurement and reference arms by a beam splitter. Each arm is terminated by a plane mirror, with the reflected beams being recombined at the beam splitter, producing a single output with intensity I , and a back reflection into the laser which will not be considered here. Similarly, the arms of the interferometer may be terminated by cube corner retroreflectors.

Taking the simplest possible case of a beam splitter with a 50:50 power splitting ratio, assuming unity refractive index throughout the interferometer, quasi-monochromatic light [20], and neglecting the vector nature of the electric field, with the result that all polarisation effects are also neglected, the electric field magnitude of the measurement and reference beams at the output is given by [13]

$$E_{\text{mes}} = \frac{1}{\sqrt{2}} \cos(2kx_{\text{mes}} - \omega_{\text{mes}}t) E_0, \quad (2.1)$$

$$E_{\text{ref}} = \frac{1}{\sqrt{2}} \cos(2kx_{\text{ref}} - \omega_{\text{mes}}t) E_0, \quad (2.2)$$

where x_{mes} and x_{ref} are the distances from the beam splitter to the measurement and reference mirrors respectively (as indicated in figure 2.1), k is the wavenumber, ω_{mes} and ω_{ref} are the optical frequencies in the reference and measurement arms

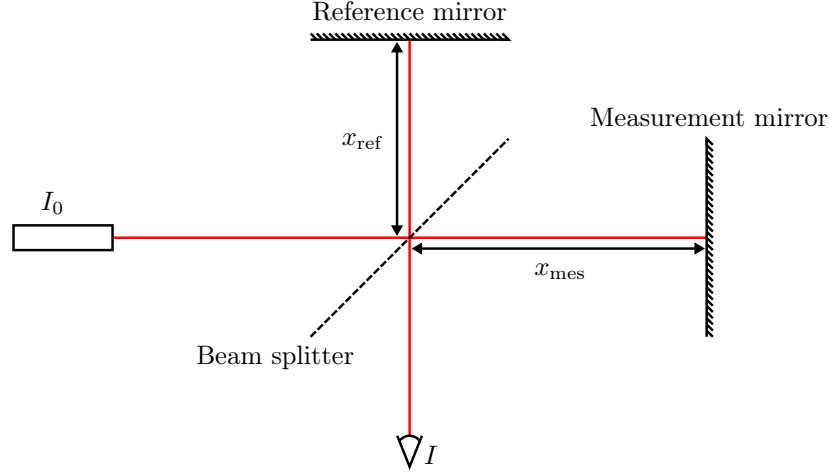


Figure 2.1: Schematic of plane mirror Michelson interferometer.

of the interferometer respectively, t is time, and E_0 is the electric field amplitude at the input to the interferometer. In the homodyne case, $\omega_{\text{mes}} = \omega_{\text{ref}} = \omega$. The wavenumber is defined in this work as [20]

$$k = \frac{2\pi}{\lambda}, \quad (2.3)$$

where λ is the optical wavelength. The total electric field magnitude at the output is given through the principle of superposition [20] by the sum of these two magnitudes, and the resulting intensity is proportional to the square of this sum as

$$I \propto (E_{\text{mes}} + E_{\text{ref}})^2, \quad (2.4)$$

$$\propto (\cos(2kx_{\text{mes}} - \omega t) + \cos(2kx_{\text{ref}} - \omega t))^2. \quad (2.5)$$

Expanding and applying the trigonometric identity

$$\cos(x) \cos(y) = \frac{1}{2} (\cos(x - y) + \cos(x + y)) \quad (2.6)$$

results in

$$I \propto \cos^2(2kx_{\text{mes}} - \omega t) + \cos^2(2kx_{\text{ref}} - \omega t) + \cos(2k(x_{\text{mes}} - x_{\text{ref}})) + \cos(2k(x_{\text{mes}} + x_{\text{ref}}) - 2\omega t). \quad (2.7)$$

Three of these terms are at the optical frequency or higher, and will therefore be averaged out when detected electronically, resulting in the observed intensity

$$I_{\text{obs}} \propto \frac{1}{2\pi} \int_0^{2\pi} I d(\omega t), \quad (2.8)$$

$$\propto 1 + \cos(2k(x_{\text{mes}} - x_{\text{ref}})), \quad (2.9)$$

where the integral serves to average over one period of the optical frequency.

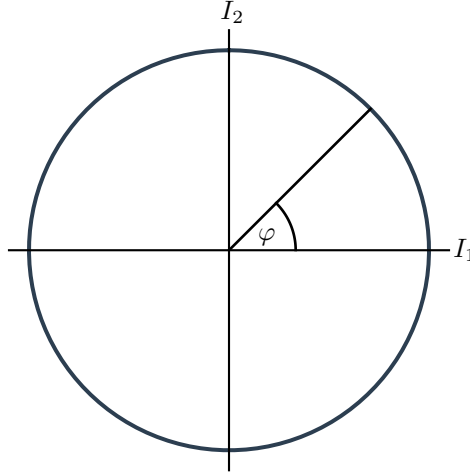


Figure 2.2: Illustration of phase measurement from quadrature signals, as described in equation 2.13

From this relationship it can be seen that the output intensity of the homodyne Michelson interferometer will vary sinusoidally as one of the mirrors is displaced, however, due to the symmetry of the cosine function about the extrema, the direction of the displacement cannot be determined. In order to determine the direction of displacement, a second signal must be introduced that is phase shifted relative to the first. Ignoring for the moment the specifics of how such a second phase shifted signal may be produced or separated from the first signal, in the ideal case the second signal is generated with a 90° phase shift relative to the first, such that the two signals are in quadrature, resulting in two observed intensity outputs

$$I_1 \propto 1 + \cos(2k(x_{\text{mes}} - x_{\text{ref}})), \quad (2.10)$$

$$I_2 \propto 1 + \sin(2k(x_{\text{mes}} - x_{\text{ref}})). \quad (2.11)$$

The instantaneous interferometric phase, φ , defined as

$$\varphi = 2k(x_{\text{mes}} - x_{\text{ref}}) \quad \text{modulo} \quad 2\pi, \quad (2.12)$$

may then be recovered from the measured intensities by first centring the measured signals on zero, and then applying the four-quadrant arctangent function. In the ideal case, neglecting the constant of proportionality, this may be accomplished in this idealised case with [21]

$$\varphi = \text{atan2}\left(\frac{I_2 - 1}{I_1 - 1}\right), \quad (2.13)$$

as illustrated in figure 2.2, where φ is in the interval $(-\pi/2, \pi/2]$. The instantaneous phase (described as such in this work on time-scales significantly larger than the optical period) as recovered by the four-quadrant arctangent function is only known

to a multiple of 2π , and therefore the absolute path length difference $x_{\text{mes}} - x_{\text{ref}}$ is not known. Consequently, homodyne interferometers measure a change in the interferometric phase, corresponding to a displacement of the measurement target, as opposed to an absolute position. To measure changes in the interferometric phase over a range greater than 2π the number and direction of crossings of the discontinuity in the four-quadrant arctangent function, from $-\pi/2$ to $\pi/2$ and vice versa, must be tracked. This is commonly referred to as fringe counting, or phase unwrapping [22], with the directional number of integer multiples of 2π covered by the interferometer counted. The phase change between two measurement points, $\Delta\varphi$, is then calculated from the sum of the difference between the instantaneous phase measurements at each point, and the difference between the integer fringe counts at each point. The displacement, assuming the wavenumber remains constant throughout the measurement, may then be recovered by dividing the change in interferometric phase by twice the wavenumber, again assuming a Michelson configuration as depicted in figure 2.1. Where the optical wavelength, and therefore wavenumber, is not constant, other effects must be accounted for. Discussion of these effects is reserved for chapter 3.

In real world interferometers, the offsets of the quadrature signals from zero are not known, the amplitudes of the two quadrature signals may not be equal, and the quadrature phase shift between the two signals may not be exactly 90° . Additionally, unwanted harmonics may also be present in the quadrature signals. All of these effects will, if uncorrected, introduce a periodic error into the measured instantaneous interferometric phase, φ , with the error being periodic with some harmonic of the illuminating wavelength. In this work, these periodic errors are defined as non-linearities, and the reduction of these error sources is the focus of this thesis. A number of other factors also affect the displacement measurement, primarily as a result of the wavenumber not remaining constant during measurement. Error sources in homodyne interferometry are discussed in detail in chapter 3.

Several optical techniques exist to introduce the required quadrature phase shift in homodyne interferometry. Most common in the field of displacement measuring optical interferometry, is a polarisation-based approach, in which orthogonal polarisations propagate in each arm of the interferometer, and the phase retardance properties of optical waveplates or Fresnel rhomb like constructions [23] are used to introduce a phase shift between the arms of the interferometer for one of the two outputs. An alternative approach employs absorbing thin film coatings to produce the required phase shift [24, 25]. A final method, rarely employed in the field of precision optical interferometry, involves introducing a wedge angle between the interfering wavefronts, and spatially separating the detectors. A more detailed analysis of the polarisation and thin film coating based methods of achieving phase quadrature is presented in section 4.2.

2.2.2 Heterodyne

Heterodyne interferometers, as the name suggests, utilise interference between beams of light having differing optical frequencies. The resulting beat frequency is then detected, and the interferometric phase deduced. Returning to equation 2.7, and

including the effects of different optical frequencies propagating in each arm of the interferometer, results in

$$I \propto \cos^2(2k_{\text{mes}}x_{\text{mes}} - \omega_{\text{mes}}t) + \cos^2(2k_{\text{ref}}x_{\text{ref}} - \omega_{\text{ref}}t) \\ + \cos(2(k_{\text{mes}}x_{\text{mes}} - k_{\text{ref}}x_{\text{ref}}) - (\omega_{\text{mes}} - \omega_{\text{ref}})t) \\ + \cos(2(k_{\text{mes}}x_{\text{mes}} - k_{\text{ref}}x_{\text{ref}}) - (\omega_{\text{mes}} + \omega_{\text{ref}})t), \quad (2.14)$$

where the optical frequency of the light propagating in the reference arm of the interferometer is ω_{ref} and the optical frequency of the light incident to the measurement arm of the interferometer is ω_{mes} , and the respective wavenumbers are k_{ref} and k_{mes} . Making the assumption that the optical frequencies are closely spaced, with heterodyne frequencies of MHz to tens of MHz typical [26–28], such that

$$\omega_{\text{mes}} \gg |\omega_{\text{mes}} - \omega_{\text{ref}}|, \quad (2.15)$$

$$\omega_{\text{ref}} \gg |\omega_{\text{mes}} - \omega_{\text{ref}}|, \quad (2.16)$$

the first, second, and fourth terms of equation 2.14 will be at optical frequencies, and will be averaged out by electronic detection. Averaging out each term over the appropriate period, as was done in the homodyne case, results in an observed intensity given by

$$I_{\text{obs}} \propto 1 + \cos(2(k_{\text{mes}}x_{\text{mes}} - k_{\text{ref}}x_{\text{ref}}) - (\omega_{\text{mes}} - \omega_{\text{ref}})t). \quad (2.17)$$

In the case that the measurement target is stationary, equation 2.17 describes a time varying sinusoidal signal with a frequency given by the heterodyne beat frequency, the difference between the measurement and reference arm frequencies. When the measurement target is in motion, the frequency of the output is shifted [29], with the frequency reaching zero when

$$\frac{d}{dt}(2(k_{\text{mes}}x_{\text{mes}} - k_{\text{ref}}x_{\text{ref}})) = \frac{d}{dt}((\omega_{\text{mes}} - \omega_{\text{ref}})t). \quad (2.18)$$

Assuming the reference arm length remains fixed, and only the measurement target is in motion, this results in a zero frequency output when the velocity of the measurement target is given by

$$\frac{dx_{\text{mes}}}{dt} = \frac{\omega_{\text{mes}} - \omega_{\text{ref}}}{2k_{\text{mes}}}. \quad (2.19)$$

The maximum target velocity that can be measured is therefore limited not only by the bandwidth of the measurement electronics, as in the homodyne case, but also by the heterodyne frequency. In practice, the electronics used in heterodyne detection have a finite bandwidth centred on the heterodyne frequency, that does not reach zero, and as such the actual velocity limit is more stringent than suggested by equation 2.19.

For velocities resulting in a frequency within the pass band of the detection electronics, the electrical output of a heterodyne interferometer takes the form of a continuous sinusoidal signal. The interferometric phase may therefore be measured

with phase-sensitive detection, allowing the effects of $1/f$ noise in the electronics to be minimised [30]. Typically, a phase and frequency reference output is provided by heterodyne interferometers, with a portion of the measurement and reference arm beams being interfered prior to passing through the interferometer arms. This reference output may then be utilised as the reference input to a lock-in amplifier, which is used to measure the interferometric phase [26]. Having obtained a measurement of a change in the interferometric phase between two measurement points, the displacement may be calculated as in the homodyne case, with the caveat that the small difference in wavelength between the two heterodyne frequencies must be accounted for.

The heterodyne frequency shift may be introduced by a number of methods. Historically, the most common approach [26] has been to utilise Zeeman split Helium-Neon (HeNe) lasers, which produce two orthogonally polarised components with optical frequencies separated by a heterodyne frequency shift of the order of megahertz [27]. Alternative approaches include frequency shifting a portion of the laser beam using acousto-optic modulation [28], and more advanced alternatives to Zeeman splitting that introduce birefringent materials to the laser cavity [31], resulting in a greater heterodyne frequency shift. Discussion of the benefits of the various methods of introducing the heterodyne frequency shift is reserved for section 2.4.

2.2.3 Other approaches

Although the previously described homodyne and heterodyne two-beam interferometer designs are the most commonly employed where displacements must be measured over ranges of millimetres or greater with sub-nanometre uncertainty, other approaches are possible.

Frequency modulation

In addition to the heterodyne approach previously discussed, approaches relying upon the modulation of the optical frequency have also been described in the literature, where the interferometer is illuminated with a optical frequency modulated laser source, as opposed to two discrete frequencies [32]. Both linear [32] and sinusoidal frequency modulations [33, 34] have been explored. These approaches rely upon the fact that a modulated illuminating optical frequency will introduce a modulation in the phase at the output of the interferometer that is dependent on both the frequency modulation profile and the path length difference between the interfering beams. A range of methods of recovering the interferometric phase exist [32, 33, 35, 36], and the exact signal processing methodology depends on the optical frequency modulation profile employed. A significant benefit of this approach is that the absolute length difference between the arms of the interferometer can be calculated, although this measurement has a significantly higher uncertainty than the displacement measurement. Multiple interferometers can also be interrogated by some approaches [33, 37], as long as the optical path lengths of the interferometers are sufficiently different, with a minimum difference in optical path length of 5 mm quoted for one implementation [37]. Whilst such approaches have been employed

for precision displacement measurement [37], further work is required to verify the uncertainties in applications where sub-nanometre uncertainties are required.

Fabry-Perot

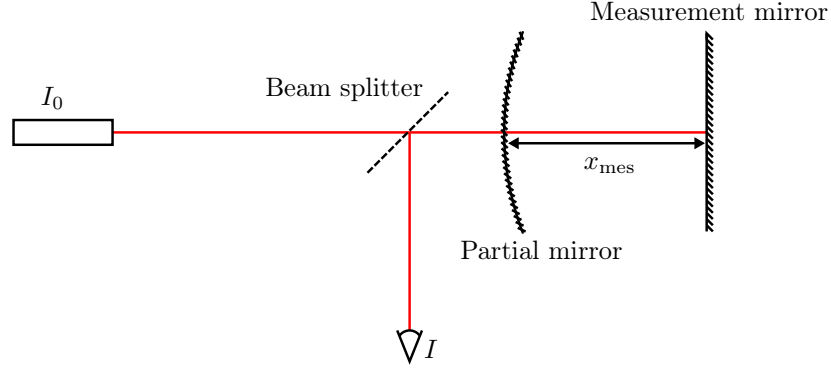


Figure 2.3: Schematic of Fabry-Perot interferometer operated in reflection.

Moving beyond two-beam interference, displacements may also be measured with interferometers operating in a Fabry-Perot configuration. Some of the interferometer designs previously considered, that make use of measurement and reference mirrors that are both parallel and coaxial [33], could be thought of as operating in a Fabry-Perot configuration, however, the operating principle of the multi-beam Fabry-Perot interferometers considered here is fundamentally different to that of the interferometer systems considered up to this point. Multi-beam Fabry-Perot interferometers consist of an optical cavity, which in displacement measuring applications typically consists of one fully and one partially reflecting mirror, with the cavity being interrogated in reflection [38–43] as shown in figure 2.3. Fabry-Perot cavities constructed in this manner have a reflectivity that is dependent on both the length of the cavity, and the illuminating wavelength, with the reflectivity of the cavity following an Airy function with respect to wavelength and length [20]. Rather than illuminating the interferometer with a single, or two closely spaced optical frequencies, displacement measuring Fabry-Perot interferometers are instead illuminated with a tunable laser source, with the frequency of the laser source being tuned so as to maximise the reflectivity of the cavity [38], for example using the Pound-Drever-Hall technique [44]. By interfering the output of the tuned laser with a second laser of known optical frequency, the optical frequency of the tuned laser may then be measured by recording the beat frequency that arises as a result of the heterodyne interference [38]. Alternatively, two tunable lasers may be locked to two separate cavities, and the optical frequency difference between the two tuned lasers measured with the same technique [42]. This approach has the advantage that

differential measurements may be made between two separate cavities, rather than only between ends of a single cavity.

Multiple reflection

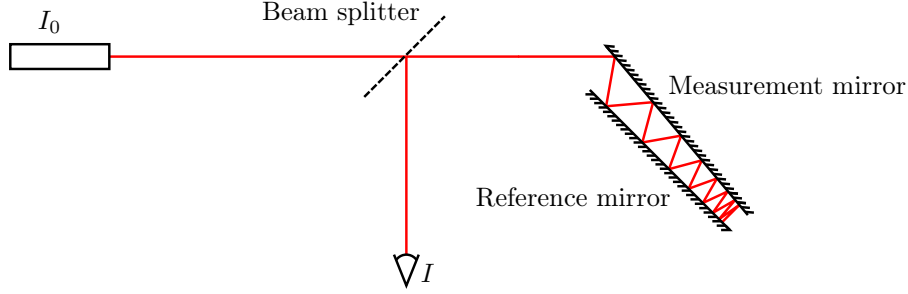


Figure 2.4: Schematic illustrating the operating principle of the multiple reflection path length multiplication interferometer.

A final approach that has been proposed in the literature with the goal of obtaining sub-nanometre interferometric displacement uncertainties is to amplify the displacement sensitivity of a conventional homodyne or heterodyne interferometer by increasing the number of times the measurement arm beam is reflected from the measurement mirror [45, 46], as illustrated in figure 2.4. This approach results in a very sensitive instrument, however, the measurement is also very sensitive to angular deviations of the measurement mirror, and has a limited displacement range, with a maximum range of $100\text{ }\mu\text{m}$ reported [47].

2.3 The state-of-the-art in displacement measuring interferometry

2.3.1 Homodyne

The first optical interferometers worked on the homodyne principle [19], and as such homodyne interferometry has been extensively studied for over a century. A number of homodyne interferometer designs have been described in the literature that are capable of achieving non-linearities in the range of tens of picometres or less, and these will be reviewed here.

The smallest reported traceably measured non-linearity for a homodyne interferometer to date has been achieved by a single pass (fringe period $\lambda/2$) differential two-wavelength system developed by the Czech Metrology Institute (CMI) [48], a schematic representation of which is shown in figure 2.5. Although this interferometer makes use of two wavelengths (633 nm and 532 nm), each wavelength is operated as a separate interferometer, with polarisation-based homodyne detection

employed. The two displacement measurements, made at different wavelengths, are then used to provide drift compensation [48]. This interferometer has been found to have 10 pm of non-linearity, measured as a peak-to-peak value [11]. The use of very high extinction ratio (100000:1) calcite polarising beam displacers to separate and recombine the polarisation states in the measurement and reference arms of the interferometer is identified as a key factor in achieving this low non-linearity, with the high extinction ratio minimising polarisation mixing between the interferometer arms [11]. By employing a differential design, with the measurement and reference beam paths closely spaced (6 mm separation, increased to 14 mm for non-linearity measurement [11]), the effects of air turbulence are reduced compared to a 90° Michelson configuration, where the reference and measurement arms pass through entirely different regions of the optical medium. This configuration also permits the dead path [49] to be minimised, the importance of which will be demonstrated in chapter 3.

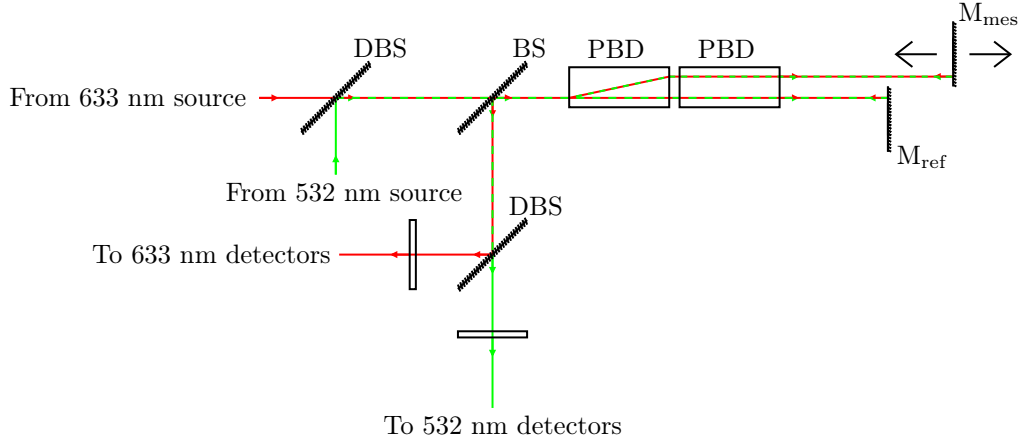


Figure 2.5: Schematic representation of the two-wavelength common path homodyne interferometer described in [48]. The homodyne detection system is not described in [48], and therefore is not shown. DBS, dichroic beam splitter; BS, beam splitter; PBD, polarising beam displacer.

Another differential polarisation-based homodyne interferometer design, developed by the National Institute of Advanced Industrial Science and Technology (AIST), and shown in figure 2.6, has reported non-linearities of 10 pm in amplitude [50], corresponding to a 20 pm peak-to-peak value. This interferometer again employs a stabilised HeNe laser operating at approximately 633 nm for illumination. Non-linearities were estimated based upon the amplitude deviations of the quadrature signal Lissajous figure from the ideally circular state. This approach to non-linearity measurement is shown in section 3.3.3 to underestimate non-linearity. The interferometer design makes use of four separate outputs, separated by 90° quadrature phase differences, minimising lost light, and allowing for offsets in the quadrature signals to be partially compensated for [50]. The measurement and reference beams are also closely spaced, described as a quasi-common path arrangement,

reducing the effects of air turbulence [50]. In contrast to the CMI system [48], this interferometer operates in a double pass configuration, resulting in a fringe period of $\lambda/4$.

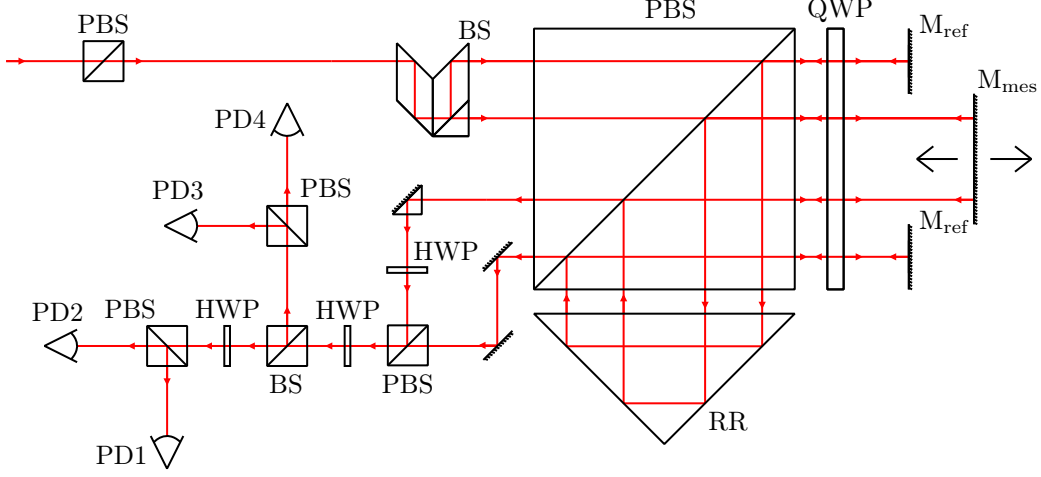


Figure 2.6: Schematic representation of the AIST plane mirror differential homodyne interferometer described in [50]. The measures taken to ensure beam paths in air and glass are equal are shown. PBS, polarising beam splitter; BS, beam splitter; QWP, quarter waveplate; RR, retroreflector; HWP, half waveplate; PD, photodiode.

A final differential polarisation-based homodyne interferometer design, shown in figure 2.7, and developed by the Institute of Scientific Instruments of the Czech Academy of Sciences (ISI) has been found to achieve non-linearities of 55 pm, reported in this case as a standard deviation [51]. Assuming the non-linearities follow a sinusoidal error distribution, this corresponds to a peak-to-peak non-linearity of 156 pm [52], as calculated with equation 3.62. The authors of [51] note that this is a preliminary result, and further improvements to the optical configuration and data processing are required to reduce the non-linearity. Illumination is in this case provided by an iodine-stabilised 633 nm diode laser [53]. This system, in common with that described in [50], makes use of a quasi-common path double pass (fringe period $\lambda/4$) differential configuration, and four quadrature outputs with 90° phase differences. In addition to displacement measurement, this interferometer incorporates a four-quadrant position-sensitive detector, with the optical configuration of the interferometer arranged such that changes in the beam position on the position-sensitive detector are correlated to changes in the tip and tilt angle of the measurement mirror [51]. Due to the optical configuration of the interferometer, which employs corner cube retroreflectors for beam steering, angular deviations of the target mirrors result only in translations of the reference and measurement beams in the detector plane, with the angle of incidence of the beams on the detector remaining unchanged [51]. This property, although not described by the authors, is shared by the previously discussed AIST interferometer [50]. In addition to simplifying the implementation of tilt sensing, this property of the optical configuration also avoids fringes appearing across the beams on the detectors due to wavefront misalignment, which depending

on the beam diameters, and detector sizes, may reduce the sensitivity of the fringe contrast to angular deviations.

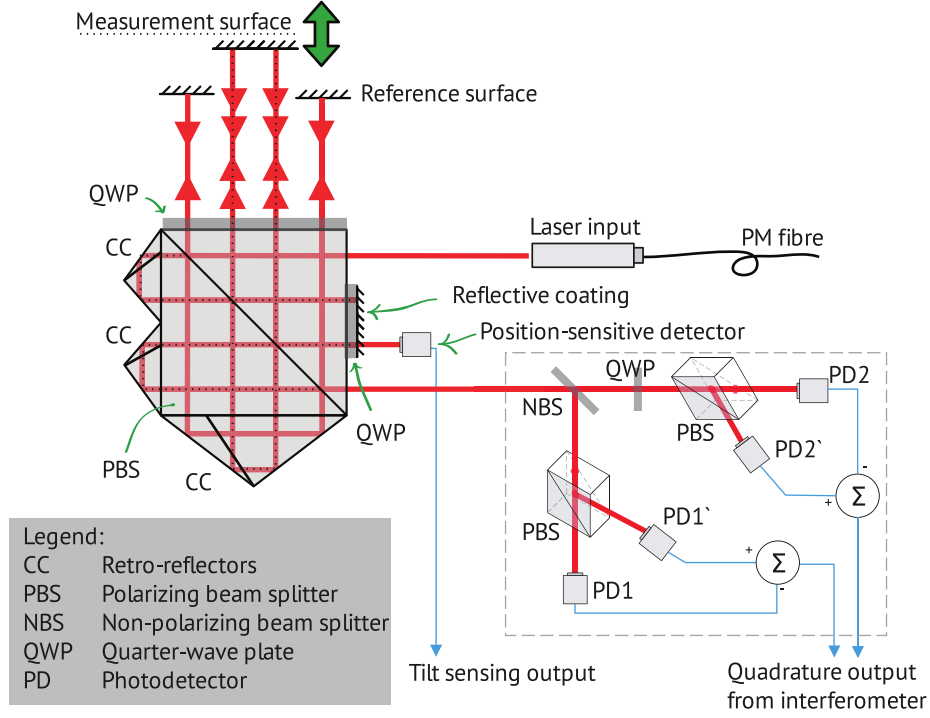


Figure 2.7: Schematic representation of the ISI plane mirror differential homodyne interferometer described in [51]. Reproduced from [51] (<https://doi.org/10.1016/j.optlaseng.2021.106568>) under the terms of the Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International (CC BY-NC-ND 4.0) licence.

In addition to polarisation-based designs, a phase-quadrature coating [25] based homodyne interferometer developed by NPL has also been shown to be capable of achieving non-linearities in the range of tens of picometres. A schematic representation of this interferometer is shown in figure 2.8. Reported non-linearities vary with illuminating laser source, with a minimum peak-to-peak non-linearity of 40 pm reported when employing a stabilised HeNe laser source operating at approximately 633 nm [54]. In common with the previously described interferometer designs, this interferometer operates with closely spaced parallel measurement and reference beams. A double pass (fringe period $\lambda/4$) configuration is again employed. This interferometer design has been utilised in a range of metrological instruments [6, 55, 56], and is further discussed in chapter 7.

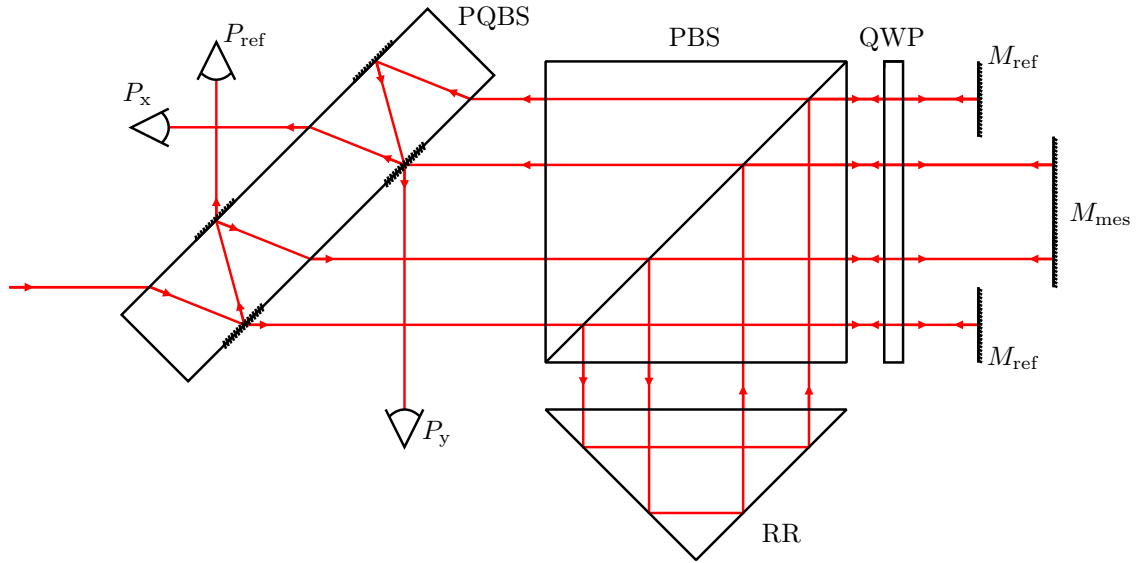


Figure 2.8: Schematic representation of the NPL plane mirror differential homodyne interferometer described in [54]. PQBS, phase-quadrature beam splitter; PBS, polarising beam splitter; QWP, quarter waveplate; RR, retroreflector; QWP, quarter waveplate; P_x and P_y , quadrature outputs; P_{ref} , intensity reference output.

2.3.2 Heterodyne

Although a relatively modern development compared to homodyne interferometry, heterodyne interferometry has been employed for precision displacement measurement for over 50 years [27], at the time of writing. The primary factor driving the development of heterodyne detection has historically been the comparative ease of measuring phase changes of Alternating Current (AC) signals, as opposed to the Direct Current (DC) alternative. By measuring an AC signal, it is also possible to avoid the effects of $1/f$ noise, which is often a factor limiting the precision of low frequency electronics [30]. As such, even very early heterodyne interferometers were able to resolve displacements of 60 pm [27]. Resolution should not, however, be confused with accuracy, and whilst heterodyne interferometers may have a high precision, heterodyne designs are also affected by non-linearities in the phase measurement [57], with non-linearities in the range of 1 nm to 4 nm typical [58] for the classic heterodyne interferometer design employing a Zeeman split laser, as described in, for example [27]. Whilst non-linearities of the order of single digit nanometres may be sufficient for some applications, more recent developments, driven by an improved understanding of the sources of non-linearity in heterodyne interferometers [59–61], have lead to the development of heterodyne interferometers with non-linearities in the range of tens of picometres or less. Such designs will be reviewed here.

Beginning once again with the lowest reported traceably measured non-linearity, a heterodyne interferometer design developed by Physikalisch Technische Bundesanstalt (PTB) has been shown to achieve non-linearities with a peak-to-peak value of less than 10 pm [11, 28]. This interferometer, a schematic representation of which

is shown in figure 2.9, employs an iodine-stabilised frequency doubled Nd:YAG laser, operating at approximately 532 nm as the primary light source, with the heterodyne frequency separation introduced by a pair of acousto-optic modulators, one operating at 80 MHz and the other operating at 78.4375 MHz resulting in a heterodyne beat frequency of 1.5625 MHz [28]. By producing each frequency separately, and maintaining spatial separation of the two frequencies until interference occurs, as first proposed in [62], the potential for frequency mixing, a cause of non-linearity in heterodyne interferometry [59], is eliminated. For additional non-linearity mitigation, all optical surfaces in the interferometer are wedged, minimising the potential for back reflections, another source of non-linearity [61]. The spatial separation of the two heterodyne frequencies, which in this interferometer design are fibre fed to the interferometer through separate fibres, introduced the potential for phase changes to occur between the two frequencies, for example due to mechanical stress in the fibre. To compensate for this effect, the interferometer is effectively implemented as two separate interferometers measuring the same displacement, with the measurement and reference arm frequencies reversed for the second interferometer. The interferometric phase change is then measured as the phase change between the outputs of the two interferometers. This configuration acts to effectively half the fringe period of the interferometer, and cancels out the effects of any disturbances common to both interferometers, as would be introduced, for example, by varying fibre stress [28]. A further benefit of this configuration is that the interferometer is truly differential; neglecting the effects of refractive index variations across the beam paths, a displacement of the measurement or reference mirror by the same distance will result in an identical change in the measured interferometric phase, with the fringe period being determined by the average of the two heterodyne frequencies. In contrast, if each heterodyne frequency were to propagate in a separate arm of the interferometer, the fringe period resulting from a movement of the measurement arm mirror would differ slightly from that produced by a displacement of the reference arm mirror.

An earlier interferometer design utilising a similar separation of the heterodyne frequencies, shown in figure 2.10, developed by National Institute of Standards and Technology (NIST), has been estimated to have non-linearities of 7.4 pm RMS [15], corresponding to a peak-to-peak non-linearity of approximately 22 pm, assuming sinusoidal non-linearity. In this work, the authors separate non-linearities with optical and electronic origins, and including the observed non-linearity in the phase measurement electronics increases the RMS non-linearity to 12.2 pm, equivalent to a 34.5 pm peak-to-peak non-linearity. The two heterodyne frequencies are again produced by a pair of acousto-optic modulators, acting in this case on the measurement and reference arms separately. To avoid the previously described problem of different optical frequencies propagating in each arm of the interferometer preventing true differential measurement, polarisation sensitive acousto-optic modulators are used, and the polarisation is rotated by 90° after reflecting from the target mirrors by means of a double pass through a Quarter Waveplate (QWP). The acousto-optic modulators are aligned relative to the input polarisation to the interferometer such that no frequency shift is imparted to the outgoing beams, with the heterodyne

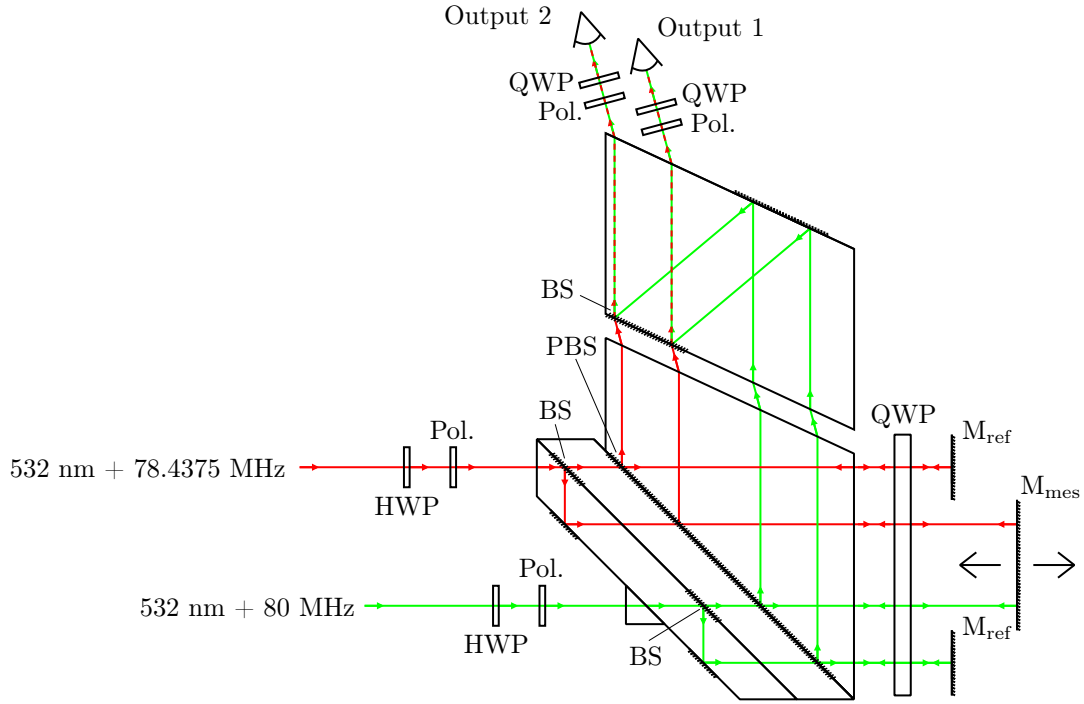


Figure 2.9: Schematic representation of the PTB plane mirror differential heterodyne interferometer described in [28]. BS, beam splitter; PBS, polarising beam splitter; QWP, quarter waveplate; HWP, half waveplate; Pol., polariser. Angles and optic sizes are not shown to scale. Beam conditioning and heterodyne frequency generation optics are not shown.

shift instead being introduced into the return beams from each arm of the interferometer. In this case, the acousto-optic modulators were driven at 39.98 MHz and 40 MHz, resulting in a heterodyne beat frequency of 40 kHz [15]. This interferometer is operated in a single pass configuration, with a fringe period of $\lambda/2$, and an illuminating wavelength of 632.8 nm provided by a polarisation stabilised HeNe laser. The sources of non-linearity in this interferometer design were not conclusively identified, however, the authors suggest parasitic reflections and optical feedback back into the laser as two possible causes [15]. This system is a further development of the interferometer described in [63], with the interferometer described in [63] having 40 pm peak-to-peak non-linearities, which were again attributed to parasitic reflections.

A third interferometer incorporating spatially separated heterodyne frequency generation has been developed by Delft University of Technology [64]. This interferometer design has been shown to have non-linearities below the measurement noise floor of 20 pm in amplitude, corresponding to peak-to-peak non-linearity below 40 pm. In common with the interferometer described in [28], both heterodyne frequencies propagate in each arm of the interferometer, however each frequency undergoes a double pass, resulting in a fringe period of $\lambda/4$. A corner cube based 90° Michelson configuration is also presented, with a fringe period of $\lambda/4$, with the

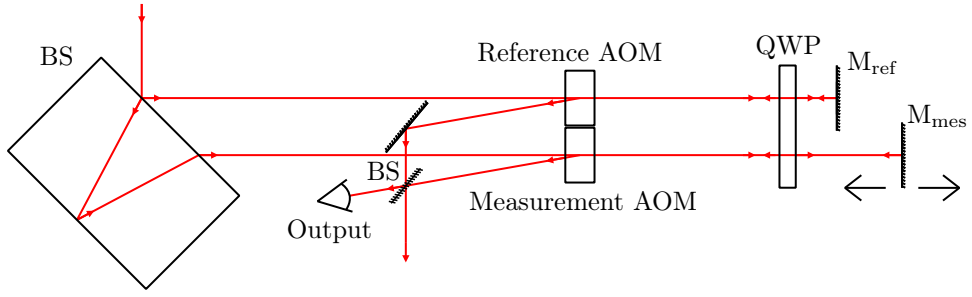


Figure 2.10: Schematic representation of the NIST plane mirror differential heterodyne interferometer described in [15]. BS, beam splitter; AOM, acousto-optic modulator; QWP, quarter waveplate. AOMs are polarisation dependent, diffracting only the return beams.

amplitude of any non-linearity again below the 20 pm noise floor. The experimentally realised interferometer design does not operate in a differential configuration, with the reference arm instead being internal to the interferometer, however a proposed design is presented for a differential plane mirror interferometer [64]. This interferometer is a further development to the design presented in [65], with the interferometer presented in [65] having peak-to-peak non-linearities of approximately 150 pm, attributed to frequency mixing in the Zeeman split HeNe laser used as the source of the heterodyne frequencies, demonstrating the importance of separating the heterodyne frequency generation. Further development of the work presented in [65] and [64] resulted in a fibre fed interferometer, allowing the laser and acousto-optic modulators to be separated from the metrology environment [66]. A fourth interferometer design operating on similar principles has been published [17], with the key difference that the interferometer described in [17] operates in a single pass differential configuration, with one frequency propagating in both the measurement and reference arms of the interferometer. The second heterodyne frequency propagates through a third, fixed, beam path, and is interfered with the measurement and reference arm return beams separately, resulting a measurement following the same concept as [15]. Non-linearities were again below the measurement noise floor, with the noise floor in this case at an amplitude of approximately 200 pm, corresponding to an upper limit on the peak-to-peak non-linearity of 400 pm.

Spatially separated heterodyne frequency generation was again employed for the single pass 90° Michelson interferometer design presented in [67], shown in figure 2.12. This interferometer operates with an illuminating wavelength of 632.8 nm, with a heterodyne beat frequency of 400 kHz. In this case the non-linearity measurement technique also described in [67] was used to measure the non-linearity of the interferometer, allowing the alignment of the polarisation components making up the interferometer to be optimised, and the non-linearity to be minimised. A first harmonic (spatial period $\lambda/2$) non-linearity amplitude of 44 pm was achieved, corresponding to a peak-to-peak non-linearity of 88 pm. An additional non-linearity term was present at a spatial period of $\lambda/4$, however, no numeric value for this component was given.

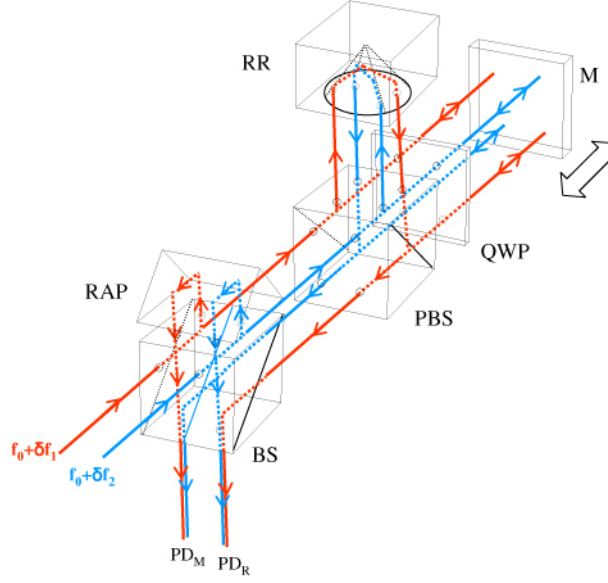


Figure 2.11: Schematic representation of the Delft University of Technology plane mirror heterodyne interferometer described in [64]. BS, beam splitter; RAP, acousto-optic modulator; PBS, polarising beam splitter; QWP, quarter waveplate; RR, retroreflector; PD, photodiode. Heterodyne frequency generation optics are not shown. Reproduced from [64] (<https://doi.org/10.1364/OE.18.001159>) under the terms of the Optica Publishing Group Open Access Publishing Agreement.

A final heterodyne design combining a heterodyne optical interferometer with a non-linearity of 300 pm with a capacitance sensor has been proposed [68], with linearisation based upon the capacitance sensor shown to reduce the RMS errors relative to displacement measurements made with an x-ray interferometer to 14 pm [11], although it is not clear whether these residual errors are a result of non-linearities, or other non-periodic error terms. A schematic representation of this interferometer design is shown in figure 2.13. The 300 pm residual non-linearities are corrected by fitting a truncated Fourier series to the difference between the displacement measured by the heterodyne interferometer and the capacitance sensor, with Fourier components corresponding to spatial harmonics of the illuminating wavelength (632.8 nm, provided by a commercial heterodyne HeNe laser source) fitted [68]. This approach is limited to the displacement range of the capacitance sensor, in this case limiting the working range to 20 μm 300 μm . A related technique developed for the correction of higher order non-linearity effects is described in section 6.3 and [69].

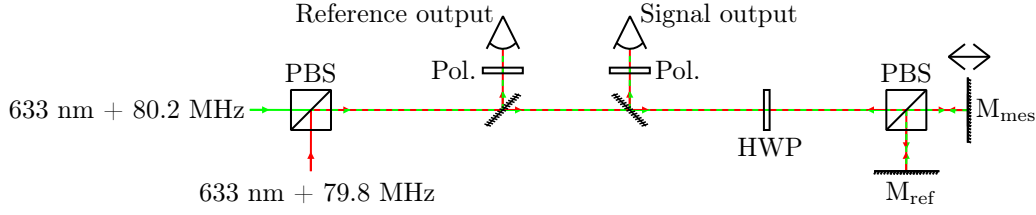


Figure 2.12: Schematic representation of the McRae plane mirror heterodyne interferometer described in [67]. PBS, polarising beam splitter; Pol., polariser; HWP, half waveplate.

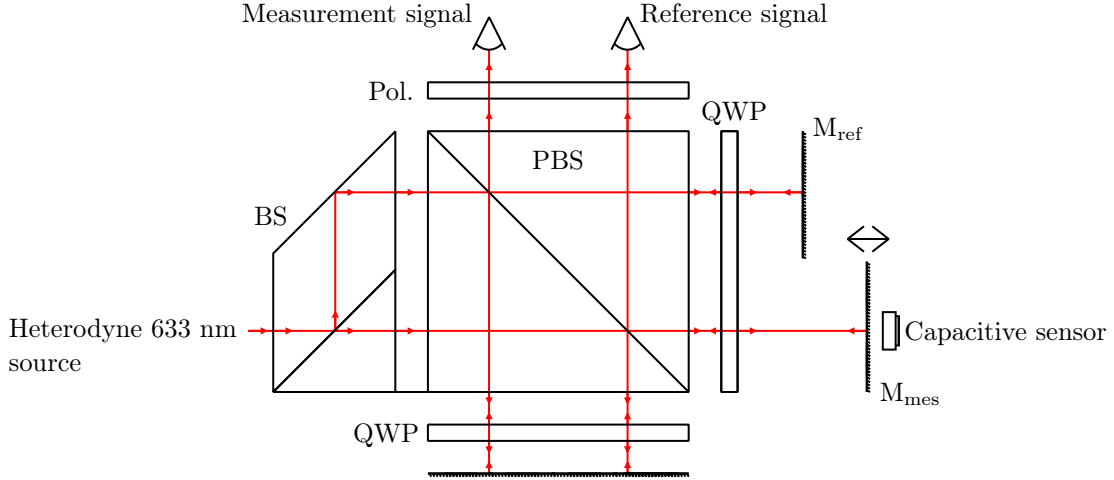


Figure 2.13: Schematic representation of the MIKES plane mirror differential heterodyne interferometer described in [68]. BS, beam splitter; PBS, polarising beam splitter; QWP, quarter waveplate; Pol., polariser; HWP, half waveplate.

2.3.3 Other approaches

In addition to the homodyne and heterodyne interferometer designs previously discussed, a number of other interferometer designs have also been developed that target displacement measurements that are linear to tens of picometres. Frequency modulated approaches, as previously described, whilst having a number of advantages, are generally aimed at lower accuracy applications. Phase modulation techniques have been developed for extremely precise measurement of displacements covering an integer number of optical fringes [9, 70, 71], however, as these approaches are restricted to integer fringe counting, or monitoring of very small (tens of picometres) perturbations around a locked fringe, they will not be discussed in detail here. Fabry-Perot interferometers that are capable of highly linear displacement measurements over ranges of millimetres or greater have, however, been described in the literature, and these will be reviewed here. Some Fabry-Perot systems that are restricted to ranges of the order of micrometres will also be discussed to provide further context for the state-of-the-art of Fabry-Perot interferometry. Although fibre Fabry-Perot

interferometers, with the cavity formed between a fibre tip and a reflecting surface, have been developed with picometre sensitivities [43, 72], and applied as an alternative to beam deflection in Atomic Force Microscopy (AFM) systems [73–75], such interferometers are limited to extremely short stand off distances and displacement ranges, and as such will also not be considered here. As discussed in section 2.2.3, the operating principle of Fabry-Perot interferometers is fundamentally different to that of two-beam interferometers. Therefore, rather than restricting discussion to periodic non-linearities, the displacement uncertainty of Fabry-Perot designs will be considered. When comparing the uncertainties discussed to the non-linearities discussed for the homodyne and heterodyne interferometers, it should be noted that the non-linearity is only one component of the uncertainty in a homodyne or heterodyne displacement measurement.

A Fabry-Perot interferometer developed by AIST, based upon a wavelength, rather than frequency, locked laser source, has been reported to achieve sub-picometre resolutions for measurements made in air [76], demonstrating the high displacement resolution potential of Fabry-Perot interferometry. A schematic of both the interferometer, and the wavelength stabilisation scheme, is shown in figure 2.14. Wavelength stabilisation is achieved by locking a tunable Nd:YAG ring laser to a fixed length Fabry-Perot cavity, consisting of two low thermal expansion glass (Ohara Inc. CLEARCERAM-Z) mirrors optically contacted to a low thermal expansion glass block [76]. A second Fabry-Perot interferometer is then employed for displacement measurement, with a tunable (50 GHz tuning range) continuous-wave fibre laser locked to the second cavity, one mirror of which may be displaced. To allow measurement of the wavelength difference between the two Fabry-Perot cavity locked lasers at >10 GHz beat frequencies, a femtosecond frequency comb is locked to the wavelength stabilised laser, and the wavelength of the fibre laser locked to the second Fabry-Perot cavity is measured relative to this comb, operating as a tunable single wavelength generator as described in [77], resulting in a wavelength accuracy of 10 am over a wavelength tuning range of 0.2 nm. Displacement measurement uncertainties have been calculated to be less than 1 pm over a measurement range of 10 μ m, with the uncertainty dominated by uncertainty in the determination of the free spectral range of the Fabry-Perot cavity, which determines the relationship between cavity length and resonant optical wavelength [76].

A series of Fabry-Perot interferometer designs have also been developed by NIST for the purposes of high accuracy displacement measurement [38, 40, 41, 78, 79]. Ultimately, the Gouy phase, the phase difference that accumulates between focused light in a Fabry-Perot cavity employing curved mirrors, and the plane wave case, has been identified as the limiting uncertainty source for displacement measurements made with Fabry-Perot interferometers where displacement measurements must be made over ranges of millimetres under vacuum [79]. Measurements of the Gouy phase have been made for a Fabry-Perot cavity with sufficient accuracy to allow 5 pm uncertainties over a displacement range of 50 mm [79]. As the Gouy phase is dependent on mirror curvature and astigmatism [79], this method must be repeated for each Fabry-Perot cavity. For measurements made in air, as is the case for two-beam interferometry (discussed in chapter 3), refractive index uncertainty dominates over the effects of wavefront curvature [80, 81]. A number of frequency reference sources have

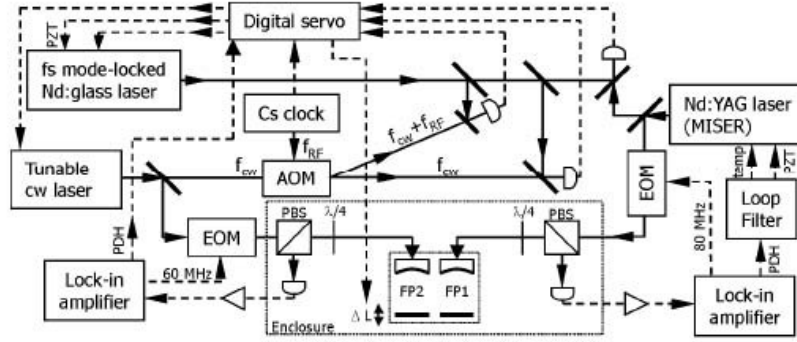


Figure 2.14: Schematic representation of the AIST Fabry-Perot interferometer described in [76]. Wavelength stabilisation and interferometer systems are shown. AOM, acousto-optic modulator; EOM, electro-optic modulator; PBS, polarising beam splitter; FP, Fabry-Perot cavity; PDH, Pound-Drever-Hall; $\lambda/4$, quarter waveplate; MISER, Monolithic Isolated Single-mode End-pumped Ring oscillator. Reproduced from [76] (<https://doi.org/10.1364/OE.14.005984>) under the terms of the Optica Publishing Group Open Access Publishing Agreement.

been suggested, including single frequency stabilised lasers [38] and frequency comb based approaches [41], with comb based approaches considered preferable for long range (tens of millimetres) measurements [41]. Displacement uncertainties of 10 pm have been achieved in vacuum over displacement ranges of up to 50 mm [78, 81].

The Fabry-Perot interferometers discussed up to this point measure the change in length of a Fabry-Perot cavity. Whilst the length change of a Fabry-Perot cavity could be considered as a differential measurement, differential measurement between two mirrors that are separate from the interferometer body is often desirable, and a Fabry-Perot interferometer capable of measuring the differential displacement between two planar mirrors has been proposed in the literature [42]. This system, shown in figure 2.15 is based upon two Fabry-Perot cavities, each consisting of one spherical mirror (75 mm radius of curvature) and one planar mirror, and two extended cavity diode lasers. Each laser is locked to one of the cavities using the PD1 and PD2 outputs in figure 2.15. Separate beams picked off from each laser are also interfered. The resulting beat frequency is measured at the frequency counter output, yielding the frequency difference between the two lasers. In order to give an absolute frequency reference, the lasers are switched between locking to the Fabry-Perot cavities, and to a rubidium vapour cell, locking to the Rubidium D_2 absorption line in the latter case, resulting in a reference wavelength of 780.24 nm [42]. Some non-linearity (with the term used here in the literal sense, not to refer exclusively to the periodic errors that occur in two-beam interferometry, as is the case elsewhere in this thesis) in the displacement measurements made with this interferometer was observed in comparison with x-ray interferometer measurements, with the non-linearity having a standard deviation of 2.84 pm over a 160 nm displacement [11]. The authors note that the non-linearity appears to have a periodic component, with a peak-to-peak value of approximately 4 pm. The source of this periodic non-linearity could not

be conclusively identified, however, mechanical effects are suggested as a cause [42]. This interferometer has a maximum measurement range of $1\text{ }\mu\text{m}$ [11].

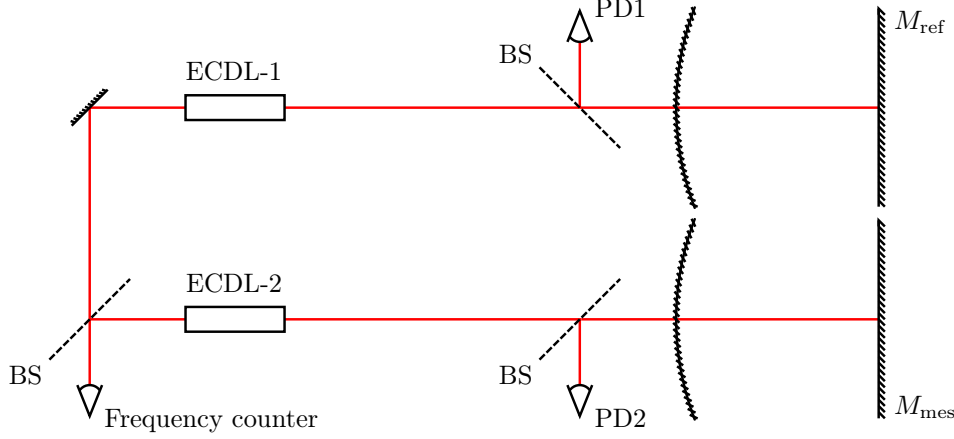


Figure 2.15: Schematic representation of the dual Fabry-Perot interferometer described in [42]. ECDL, external cavity diode laser; BS, beam splitter, PD, photodiode.

2.4 Discussion and comparison of techniques

The interferometers reviewed in this chapter are summarised in table 2.1. One immediately apparent conclusion is that Fabry-Perot based systems are able to achieve the lowest measurement uncertainties, as compared to two-beam interferometry, with sub-picometre displacement uncertainties reported for two such systems [76, 79].

Fabry-Perot based interferometers do, however, have some disadvantages that may limit the wide spread application of such devices in metrological applications. The most readily apparent limitation of Fabry-Perot interferometry is the limited displacement range. Two-beam interferometry may operate over displacement ranges that are limited only by the coherence length of the laser source, whereas Fabry-Perot interferometers, if continuous displacements must be monitored, are limited by both the tuning range of the illuminating laser, and the cavity geometry [38]. In some cases the displacement range of Fabry-Perot interferometers is restricted to the order of micrometres [42, 76], although Fabry-Perot interferometers that are capable of measuring displacements over ranges of tens of millimetres have also been demonstrated [78, 79]. A second limitation is the requirement to accurately track the optical frequency of a tunable laser source, with spectroscopic frequency references [40–42, 78] or frequency combs [39, 76] typically required in high accuracy applications. This requirement significantly increases both the cost and physical size of the associated optics and electronics. The required bandwidth of the measurement electronics is often also significantly increased relative to two-beam interferometry, with the required bandwidth increasing with displacement range for Fabry-Perot

Table 2.1: Comparison of state-of-the-art displacement measuring interferometer designs. Non-linearities are quoted as peak-to-peak values, converted from the published figures assuming a sinusoidal error distribution when required. Similar interferometer designs developed within the same research group have been grouped, and the year at which the quoted (minimum) non-linearity was reached is given. Entries are sorted by non-linearity first, and year second.

Reference	Year	Approach	Schematic	Non-linearity	Range
[76]	2006	Fabry-Perot	2.14	$<1 \text{ pm}^1$	$10 \text{ }\mu\text{m}$
[38, 40, 41, 78, 79]	2012	Fabry-Perot	Not shown	$<1 \text{ pm}^1$	50 mm
[11, 42]	2012	Fabry-Perot	2.15	4 pm^2	$1 \text{ }\mu\text{m}$
[11, 48]	2009	Homodyne	2.5	10 pm	$>\text{cm}$
[11, 28]	2012	Heterodyne	2.9	10 pm	$>\text{cm}$
[50]	2018	Homodyne	2.6	20 pm^3	$>\text{cm}$
[15, 63]	2000	Heterodyne	2.10	22 pm	$>\text{cm}$
[17, 64–66]	2010	Heterodyne	2.11	$<40 \text{ pm}$	$>\text{cm}$
[54]	2000	Homodyne	2.8	40 pm	$>\text{cm}$
[11, 68]	2011	Heterodyne	2.13	40 pm	$280 \text{ }\mu\text{m}$
[67]	2012	Heterodyne	2.12	$>88 \text{ pm}^4$	$>\text{cm}$
[51]	2021	Homodyne	2.7	156 pm	$>\text{cm}$

¹ Estimated displacement uncertainties are quoted, rather than non-linearity.

² Peak-to-peak value of observed periodic error.

³ Non-linearities are evaluated based on Lissajous figure radius errors, which may underestimate the non-linearity error.

⁴ Non-linearities are only measured at a single spatial period. Non-linearities are visible in the results at other spatial periods.

interferometry, as opposed to with maximum target velocity for two-beam interferometry. In some cases Fabry-Perot interferometers are also limited by the cavity being interrogated in transmission [78, 79], requiring optical access to the back side of both mirrors, although this may be overcome by interrogating the cavity in reflection [39, 40, 42, 76]. A related limitation is that many Fabry-Perot interferometer systems measure changes in the length of a single cavity [38, 40, 41, 76, 78, 79], which may limit the applicability of such systems where differential measurement is required. Differential measurement with Fabry-Perot interferometry has been realised by employing two cavities [39, 42], however, this again increases the size and cost of the associated electronics and optics. A final limitation of Fabry-Perot interferometry that is related, but not identical, to the limited displacement range, is that the stand off distance is limited to the cavity length. Quasi-common path two-beam interferometer designs allow the interferometer optics to be located some distance from the measurement and reference mirrors, whilst remaining relatively insensitive

to refractive index fluctuations in the space between the interferometer optics and the mirrors [50], and only small access ports are required for the beams to pass through. An equivalent Fabry-Perot implementation would require two cavities operating differentially, with the associated increase optical and electronic complexity, size and cost, and significantly larger optical access ports would be required. Additionally, in the Fabry-Perot case the cavity must be designed specifically for the required stand off distance. Precision interferometers are often integrated into larger instruments, for example scanning probe microscopes or stage calibration systems, and having the ability to separate the interferometer optics from the measurement target may simplify the instrument design process.

As a result of these limitations, it is desirable to develop two-beam interferometers that are able to match the accuracy of Fabry-Perot interferometers, but with increased measurement range, and reduced physical size, optical and electronic complexity, and cost.

Considering the possible implementations of two-beam interferometry, it can be seen from table 2.1 that both homodyne and heterodyne interferometers have achieved traceably measured peak-to-peak non-linearities of 10 pm [28, 48], with a number of other homodyne and heterodyne interferometers achieving non-linearities in the sub-100 pm range. There is therefore no clear performance advantage to either state-of-the-art heterodyne or homodyne interferometry with regards to non-linearity.

In the heterodyne case, unwanted mixing of the two heterodyne frequencies between the arms of heterodyne interferometers has been identified as a key cause of non-linearity [59, 82]. This frequency mixing effect has been minimised in low non-linearity heterodyne interferometer designs, as first proposed in [62], by spatially separating the generation of the two heterodyne frequencies, rather than relying on orthogonal polarisation states to separate the frequencies [15, 17, 28, 63, 64, 66], although this approach may add significant complexity to the design. In cases where polarisation-based separation of the heterodyne frequencies was employed, the use of high extinction ratio polarising optics, along with careful alignment, was required to achieve sub-100 pm non-linearities [65, 67]. Residual non-linearities after minimising frequency mixing are attributed to unwanted back reflections in one case [63]. In other cases the importance of minimising unwanted back reflections are discussed, and mitigation for back reflections is built into the interferometer design [15, 28]. Faraday isolators have been used to minimise the effects of free space back reflections in the laser cavity [15]. Tilted optical surfaces have been used to reduce the effects of back reflections within the interferometer optics, and quarter wave plates and polarisers have been used to reduce the effects of back reflections of photodiode detectors [28].

In the homodyne case polarisation mixing between the arms of the interferometer has also been identified as a cause of non-linearity [83], and high extinction ratio optics have again been employed to minimise this effect [48]. Beyond the use of high extinction ratio polarising optics, few design choices explicitly aimed at reducing non-linearities in homodyne interferometry are apparent from the literature. One general trend that is apparent is that non-linearities appear to increase with increasing optical complexity, with the single pass interferometer design of [48]

achieving the lowest reported non-linearities, and more complex double pass designs having larger non-linearities [50, 51, 54]. Given the relatively small number of interferometer designs that are capable of achieving sub-100 pm non-linearities, however, this trend cannot be stated with certainty. The potential for unwanted back reflections to introduce non-linearities in homodyne interferometry has been established in the literature [84, 85], and further reinforces the conclusion that, without specific measures being taken to misalign back reflections, increasing the number of optical surfaces that may introduce reflections is likely to increase non-linearity.

A common factor between the majority of two-beam interferometer designs reviewed here is that they are operated in a differential configuration, with the reference and measurement arm beam paths parallel and closely spaced [15, 17, 28, 48, 50, 51, 54, 63, 68]. This design choice is not directly related to the non-linearity of the interferometer, however, where measurements are to be made in air, a differential quasi-common path configuration does reduce the sensitivity of the interferometer to air turbulence effects [50], an important consideration in precision interferometry [86]. An additional factor common to all interferometer designs reviewed here is the use of stabilised laser sources, again critical if displacements are to be measured with sub-nanometre uncertainties, as will be established in chapter 3.

Considering research areas where the sub-nanometre performance of displacement measuring two-beam interferometers may be improved, sources of residual non-linearity in the leading heterodyne interferometer design could not be identified [28]. In contrast, sources of residual non-linearity in homodyne interferometry are relatively well understood, with unwanted back reflections [84, 85], or weak optical cavities forming through networks of polarising optics [85], likely to be the cause of non-linearities that cannot be eliminated by the commonly employed Heydemann correction [87]. Further discussion of the Heydemann correction may be found in sections 3.3 and 6.2.1. Homodyne interferometry may introduce the quadrature phase shift that is required for directional fringe counting through either polarisation, as in [48, 50, 51], or absorbing thin film coating, as in [54], based approaches. The two approaches may be most readily compared when considering the interferometer designs described in [50] and [54], both of which are double pass differential homodyne interferometers. Taking the number of discrete optical components (beam splitters, waveplates, steering mirrors and prisms) included in the design as an indicator of the optical complexity, and thus susceptibility to parasitic reflections, the polarisation-based interferometer consists of 15 optical components [50], whereas the absorbing thin film based interferometer consists of 4 optical components [54]. Phase-quadrature coating based homodyne interferometers therefore appear to be a promising candidate for further research, with scope for the 40 pm peak-to-peak best case non-linearity of the existing design [54] to be improved upon.

2.5 Summary

The two primary methods of measuring displacement with two-beam optical interferometry, the homodyne and heterodyne approaches, have been described. Other methods of making displacement measurements with sub-nanometre uncertainties using optical interferometry have also been discussed. Interferometers described

in the literature that are capable of measuring displacement with sub-100 pm non-linearities have been reviewed, and are summarised in table 2.1.

Fabry-Perot interferometers currently offer the lowest displacement measurement uncertainties [42, 76, 79], however, the Fabry-Perot approach is limited by displacement range, stand off distance, and optical and electronic complexity, size, and cost. As a result of these limitations, improved two-beam interferometry is desirable.

Comparing the homodyne and heterodyne approaches, both approaches have been shown in the literature to be capable of reaching 10 pm peak-to-peak non-linearities [28, 48]. Sources of residual non-linearities in homodyne interferometry are well understood [83–85], whereas sources of residual non-linearity in leading heterodyne interferometer designs could not be identified [28]. The homodyne approach is therefore a good candidate for further research into highly linear two-beam interferometry.

Within the field of homodyne interferometry, two techniques for introducing the quadrature phase shift required for directional fringe counting are represented in the reviewed interferometer systems: polarisation-based [48, 50, 51], and absorbing thin film based [54] approaches. The absorbing thin film based approach requires significantly fewer optical components compared to the polarisation-based approach. Furthermore, the absorbing thin film based approach avoids any potential for leakage between the interferometer arms; the beam is either reflected or transmitted, with no difference in polarisation between the light propagating in either arm of the interferometer. Parasitic reflections have been identified as a key source of non-linearity in homodyne interferometers [84, 85], and therefore the reduced optical complexity of absorbing thin film based homodyne interferometers suggests further research into this interferometer design may yield reduced non-linearity. The relative benefits of the polarisation and absorbing thin film based approaches are further compared in chapter 4.

Chapter 3

Uncertainty sources in homodyne interferometry

3.1 Introduction

A wide range of uncertainty sources contribute to the overall uncertainty in displacement measurements made with homodyne optical interferometers, many of which have been well studied (see for example [26] for a wide ranging overview) in the literature. This chapter aims primarily to review known uncertainty contributions, with the ultimate goal of identifying factors limiting the performance of homodyne displacement measuring interferometry.

Contributions to the uncertainty in displacement measurements made with homodyne interferometry may be broken down into two broad groups: firstly, uncertainties in the measurement of a change in the optical phase, and secondly, uncertainties in the conversion of this measured phase change to a displacement. This chapter first reviews uncertainty contributions that affect the displacement measurement directly, and then moves on to review contributions to the phase change measurement uncertainty. This distinction is made as in some circumstances it may be desirable to consider uncertainties in the phase and displacement separately. In particular, in cases where displacement measurements are made over large ranges, or in unstable environments, the uncertainty in the phase measurement may be a minor contribution to the displacement measurement uncertainty, and a detailed evaluation of the phase uncertainty may not be warranted. Alternatively, the phase measurement may be converted to a displacement through non-standard methods, for example multi-wavelength approaches [88], and the displacement measurement uncertainty calculations presented in this work may not be applicable, whilst the phase measurement uncertainty calculations remain valid. In order to provide a clear overview of uncertainty contributions these reviews will be laid out in the form of two uncertainty budgets: one for the displacement measurement process, and one for the phase measurement process. The uncertainty budgets will follow the guidelines described in Joint Committee for Guides in Metrology (JCGM) 100:2008 [89] (commonly referred to as the Guide to the expression of Uncertainty in Measurement (GUM)) with an equation modelling the measurement process being formed, and analysed. With further development to account for specific interferometer systems,

this work may form the basis of an experimental uncertainty budget.

Following the two uncertainty budgets, which focus on practical methods for the estimation of measurement uncertainties, the effects of shot noise, which establish the ultimate performance limit for homodyne interferometry operating with a given bandwidth and optical power, will be considered. Whilst the theoretical findings have limited practical application when producing an uncertainty budget for an experimental measurement, where direct measurement of the experimental noise levels is preferred, an analysis of the theoretical noise limit is beneficial to identify performance limits, and to provide a baseline noise level and distribution for further discussion throughout this thesis.

The dominant uncertainty sources in both the phase and displacement measurements will then be discussed, aiming to identify limiting factors, with a particular emphasis on factors that may limit the performance of homodyne interferometry at the nanoscale. Monte Carlo simulations will be employed both to verify the validity of the expressions for the uncertainty due to the shot noise derived in section 3.4, and to investigate limiting factors in the optical design of phase quadrature coating based interferometers.

Throughout this chapter, several assumptions will be made. Firstly a homodyne interferometer design including an intensity reference output will be assumed, such that fluctuations in the laser source power are compensated for by normalising the quadrature outputs to the intensity reference. The interfering light will be assumed to be quasi-monochromatic, coherent and fully polarised. The interfering beams will be taken to be detected by photodiodes operating in photovoltaic mode (with no bias voltage applied), in order to minimise dark currents. Measurements will be assumed to take place in air, under environmental conditions suitable for precision displacements measurements, at approximately 20 °C following ISO 1 [90], although small fluctuations in environmental parameters around this point will be considered. Additionally, a differential plane mirror interferometer measuring a mirror that is displaced along a single axis will be assumed, and therefore the effects of out of plane motion of the target mirror and the mechanical stability of the reference mirror will not be considered. Finally, only static measurements will be considered, that is measurements in which a displacement is measured between two static locations of the measurement mirror. As a result, data age uncertainties will not be considered.

3.2 Displacement measurement uncertainty budget

3.2.1 The measurement problem

Displacement measuring optical interferometry first measures a change in interferometric phase, as will be described in section 3.3, then converts this phase change to a displacement with, accounting for the dead path correction, but neglecting wavefront errors, the following equation [49]:

$$\Delta x = \frac{\lambda_1}{2\pi N_{\text{pass}}} \Delta\varphi + \frac{\lambda_1 - \lambda_0}{\lambda_0} x_0, \quad (3.1)$$

where λ_0 and λ_1 are the optical wavelengths at the start and end of the measurement respectively, N_{pass} is the pass number of the interferometer, equal to two for a standard Michelson configuration, and x_0 is the dead path, the difference in length between the measurement and reference arms of the interferometer at the start of the measurement. It should be noted that following this definition the dead path, x_0 , is independent of the pass number of the interferometer, and is instead the distance that would be measured with a ruler between the reference and measurement mirrors of a differential plane mirror interferometer, assuming balanced paths within the interferometer optics. For the purposes of uncertainty propagation, it is helpful to expand λ_0 and λ_1 into optical frequency and refractive index components, leading to

$$\Delta x = \frac{c}{2\pi N_{\text{pass}} n_1 f_1} \Delta\varphi + \left(\frac{n_0 f_0}{n_1 f_1} - 1 \right) x_0, \quad (3.2)$$

where n_0 and n_1 are the average refractive indices in the beam paths at the start and end of the measurement respectively, and similarly f_0 and f_1 are the optical frequencies at the start and end of the measurement.

Equation 3.2 assumes the interfering wavefronts take the form of plane waves, with the wavelength being constant throughout space for a fixed refractive index value. In reality, interferometers have finite beam sizes, and therefore non-planar wavefronts. This deviation from the plane wave case will introduce an error into the measured displacement if the plane wave phase-displacement relation given by equation 3.2 is applied [58,91]. In many practical cases, as will be described in more detail in section 3.2.3, the effects of non-planar wavefronts can be accounted for by introducing a small (positive) wavefront correction factor $F_\lambda \approx 0$ [91], the form of which is elaborated upon in section 3.2.3, resulting in

$$\Delta x = \frac{(1 + F_\lambda) c}{2\pi N_{\text{pass}} n_1 f_1} \Delta\varphi + \left(\frac{n_0 f_0}{n_1 f_1} - 1 \right) x_0. \quad (3.3)$$

In addition to the effects of refractive index, optical frequency, dead path and wavefront curvature described by equation 3.3, the geometry of the measurement system must be considered in experimental work. The interferometer will measure the change in length of the beam path, whereas the desired measurand is the displacement of a physical object. Whilst in the ideal case of a laser beam reflecting from a perfectly flat mirror that remains perfectly perpendicular to the beam axis, and displaces only along the beam axis, the two are equivalent, any angular or off axis motion of the measurement target may introduce a measurement error. As geometrical errors are not unique to interferometric measurement, and are highly dependent on a specific optical configuration and measurement set up, only the simple cases of isolated cosine and Abbé errors will be considered here, with the aim being to allow the order of magnitude of such errors to be estimated, in order to establish if a more detailed analysis is required on a case-by-case basis. Geometrical errors will therefore not be included in the measurement equation.

3.2.2 Mathematical model

The measured displacement will be determined from a modified version of equation 3.3. In most practical applications, the optical frequency will not be measured during the measurement, and will instead be considered to take a single value at the start and end of the measurement, with an associated uncertainty. As such, the measured displacement will be calculated as

$$\Delta x = \frac{(1 + F_\lambda) c}{2\pi N_{\text{pass}} n_1 f} \Delta\varphi + \left(\frac{n_0}{n_1} - 1 \right) x_0. \quad (3.4)$$

The measurand Δx therefore has the following input quantities:

- $\Delta\varphi$, the measured change in the interferometric phase,
- n_0 , the average refractive index in the dead path at the start of the measurement,
- n_1 , the average refractive index in the dead path, plus the additional path length introduced by the displacement (which may be negative), at the end of the measurement,
- f , the optical frequency,
- x_0 , the dead path,
- F_λ , the wavefront correction factor,
- N_{pass} , the interferometer pass number,
- c , the free space speed of light in the vacuum,

of which all but N_{pass} and c have associated uncertainties.

3.2.3 Contributory variances

Whilst equation 3.4 is used to calculate the displacement, for the purposes of uncertainty propagation equation 3.3 must be used to model the effects of interactions between the optical frequency uncertainty and the dead path. Propagating uncertainties through equation 3.3 using the partial derivative uncertainty propagation technique [89] and then setting $f = f_0 = f_1$ and $\sigma_f = \sigma_{f_0} = \sigma_{f_1}$ results in an

expression for the variance of the displacement, given by

$$\begin{aligned}
\sigma_{\Delta x}^2 = & \left(\frac{(1 + F_\lambda) c}{2\pi N_{\text{pass}} n_1 f} \right)^2 \sigma_{\Delta\varphi}^2 + \left(\frac{n_0}{n_1} - 1 \right)^2 \sigma_{x_0}^2 \\
& + \left(\frac{x_0}{n_1} \right)^2 \sigma_{n_0}^2 + \left(\frac{(1 + F_\lambda) c}{2\pi N_{\text{pass}} n_1^2 f} \Delta\varphi + \frac{x_0}{n_1} \right)^2 \sigma_{n_1}^2 \\
& + \left(\left(\frac{n_0}{n_1 f} x_0 \right)^2 + \left(\frac{(1 + F_\lambda) c}{2\pi N_{\text{pass}} n_1 f^2} \Delta\varphi + \frac{n_0}{n_1 f} x_0 \right)^2 \right) \sigma_f^2 \\
& + \left(\frac{c}{2\pi N_{\text{pass}} n_1 f_1} \Delta\varphi \right) \sigma_{F_\lambda}^2. \quad (3.5)
\end{aligned}$$

Further simplifications may be made by solving equation 3.4 for $\Delta\varphi$ and substituting this result into equation 3.5, yielding

$$\begin{aligned}
\sigma_{\Delta x}^2 = & \left(\frac{(1 + F_\lambda) c}{2\pi N_{\text{pass}} n_1 f} \right)^2 \sigma_{\Delta\varphi}^2 + \left(1 - \frac{n_0}{n_1} \right)^2 \sigma_{x_0}^2 + \left(\frac{x_0}{n_1} \right)^2 \sigma_{n_0}^2 + \left(\frac{n_0 x_0}{n_1 f} \right)^2 \sigma_f^2 \\
& + \left(\frac{\Delta x + x_0}{f} \right)^2 \sigma_f^2 + \left(\frac{\Delta x + x_0}{n_1} \right)^2 \sigma_{n_1}^2 + \left(\frac{n_1 (\Delta x + x_0) - n_0 x_0}{(1 + F_\lambda) n_1} \right)^2 \sigma_{F_\lambda}^2, \quad (3.6)
\end{aligned}$$

where the first four terms are independent of displacement, and the final four terms are displacement dependant. Equation 3.6 will be used in this chapter to calculate displacement measurement uncertainties.

Phase uncertainty

The uncertainty in the measured change in interferometric phase may be estimated using the methods that will be described in section 3.3. In cases where the phase uncertainty is not a dominant uncertainty source, the phase dependent uncertainty calculation presented in section 3.3 may not be warranted, and in such cases the procedures described in section 3.3 can instead be utilised to identify the upper limit of the phase uncertainty, which may then be used as a fixed value in equation 3.6. Further discussion of the uncertainty in the measured change in interferometric phase is reserved for section 3.3.

Refractive index uncertainty

The uncertainty in the measured displacement will be affected by the uncertainty in the refractive index both at the start and end of the measurement process for each displacement point measured. The refractive index depends on both the optical frequency and environmental parameters, with the sensitivity of the refractive index to key environmental parameters at 633 nm under standard atmospheric conditions as given in [7] presented in table 3.1. It should be noted when interpreting table 3.1 that the practical influence of each parameter on the refractive index is determined

Table 3.1: Standard values of atmospheric parameters and corresponding refractive index sensitivities evaluated at 633 nm, as given in [7].

Influence parameter	Value at standard conditions	Refractive index sensitivity coefficient
Temperature	20 °C	$-9.2 \times 10^{-7} \text{ K}^{-1}$
Pressure	101 325 Pa	$+2.7 \times 10^{-9} \text{ Pa}^{-1}$
CO ₂ content	400 ppm	$+1.4 \times 10^{-10} (\text{ppm})^{-1}$
Relative humidity	50 %	$-8.7 \times 10^{-9} (\% \text{RH})^{-1}$
Dew point	9.27 °C	$-1.5 \times 10^{-8} \text{ K}^{-1}$
Water vapour pressure	1168 Pa	$-1.8 \times 10^{-10} \text{ Pa}^{-1}$

not only by the sensitivity coefficient, but also by the stability of each environmental parameter, and therefore in the temperature stabilised environments typically employed for precision dimensional metrology, variations in the air pressure may be the largest influence upon the refractive index. Most commonly in the field of precision interferometry, refractive index is calculated using modern improvements to the Edlén equations [92], a number of which have been published [93–97].

The most recent equations for the refractive index of air [97] suggest uncertainties of 1×10^{-8} are possible under ideal conditions, although the authors note that such low uncertainties may not be realised in practical measurements where environmental parameters may be measured with lower accuracy. In addition, localised variations in the environmental parameters due to air turbulence may be a limiting factor as environmental parameters are measured at a finite number of locations. Uncertainties in the environmental parameters may be propagated through the refractive index correction [97], however, the possibility of localised variations in the refractive index, and the uncertainties in the fitted numerical values used in the correction equations must also be taken into account on a case-by-case basis.

Alternative approaches to refractive index measurement include refractometry [98] and multiple wavelength [88] based approaches, however these are less commonly employed, and are beyond the scope of this work.

In addition to changes in the refractive index of air, optical path length changes in the portion of the beam path that propagates within the interferometer optics may take place, primarily due to thermal changes. Precision interferometry will generally take place in thermally stable environments, however, it is still desirable to design the interferometer optics such that the path lengths in each material used in the interferometer optics are balanced between the measurement and reference arms, minimising the effect of any thermal drifts, which may otherwise be challenging to quantify. If path lengths in glass are not balanced between the arms of the interferometer, this must be accounted for when calculating the refractive index, which, following the definitions used in this work, is averaged across the dead path, which may include materials other than air.

Optical frequency uncertainty

In precision interferometry the frequency uncertainty will typically be determined from the calibration certificate of the stabilised laser employed. Alternatively, standard values and suggested uncertainties are provided by the International Bureau of Weights and Measures (BIPM) [7] for a number of laser sources commonly employed in displacement measurement.

Relative standard uncertainties of 2.1×10^{-11} are achieved by iodine stabilised HeNe lasers operating at 632.8 nm [7], with the more commonly encountered polarisation stabilised HeNe lasers achieving uncertainties parts in 10^9 [99].

Dead path uncertainty

The dead path, defined here as the difference in length between the measurement and reference arms of the interferometer, such that the pass number of the interferometer does not alter the dead path, will contribute to the measurement uncertainty by increasing the sensitivity of the instrument to changes in refractive index and optical frequency.

In addition to this effect, the uncertainty in the dead path length will also directly contribute to the measurement uncertainty as described by the second term of equation 3.6. The magnitude of this contribution will depend on the relative change in refractive index over the course of the measurement, with the significant implication that even if the refractive index is known exactly, if there is any uncertainty in the dead path length any changes in refractive index will increase the measurement uncertainty.

Wavefront uncertainty

If the interfering wavefronts in the interferometer are assumed to take the form of plane waves, an error will be introduced into the displacement measurement [58,91]. If this error is not corrected, it will contribute to the measurement uncertainty, however, in the majority of cases a straight forward correction as applied in equation 3.6 is sufficient. The wavefront correction can be treated as an effective correction to the wavelength, with a fractional error in the effective wavelength introduced by wavefront effects for interference between Gaussian TEM₀₀ beams given by [91]

$$\left| \frac{\Delta\lambda}{\lambda} \right| = \left(\frac{\lambda}{2\pi w_0} \right)^2, \quad (3.7)$$

where w_0 is the beam waist radius, defined as the distance from the beam centre at which the intensity falls to $1/e^2$ of the maximum value, and assuming that the interfering beams have travelled approximately equal distances from the beam waist, and that the full beam area collected by the detector. This result has been confirmed a number of times [100–107]. For a wavelength of 632.8 nm and a beam waist radius of 1 mm wavefront effects will therefore, in the case of ideal TEM₀₀ beams, introduce a fractional error in the measured displacement of 1×10^{-8} if uncorrected. This constitutes an error of the same order of magnitude as refractive index uncertainty when applying an Edlén style refractive index correction [97]. As already proposed

in equation 3.3, to correct for this effect the wavelengths in equation 3.1 may be replaced with corrected effective wavelengths given by [91]

$$\lambda_{\text{ef.}} = \frac{\lambda}{1 - \left(\frac{\lambda}{2\pi w_0}\right)^2}, \quad (3.8)$$

with the validity of this correction restricted to the case

$$\Delta x \ll \frac{\pi w_0^2}{\lambda}, \quad (3.9)$$

a displacement range of approximately 5 m for a wavelength of 632.8 nm and a beam waist radius of 1 mm. This restriction is satisfied for all interferometric measurements presented in this thesis. Equation 3.8 again assumes that the interfering beams have travelled equal distances from the beam waist, implying the interferometer starts measurement with no dead path, however, the equation remains valid for small deviations from this case. This may be demonstrated by applying the full form of equations derived in [91], accounting for the interfering beams having travelled different distances from the beam waist. Using the equations of [91], it can be shown that a fractional error of 9×10^{-12} is introduced into the effective wavelength by applying equation 3.8 with a 1 m path length imbalance. In cases where extremely large path length imbalances are present the exact equations may be applied [91], however, as will be discussed later in this chapter, in the presence of such large dead paths, refractive index uncertainty is likely to dominate over wavefront effects for measurements made in air.

Wavefront distortions from the ideal Gaussian TEM₀₀ mode will also alter the effective wavelength. Where higher order TEM modes are present the divergence of the interfering beams can be used to estimate the effective wavelength [106], although this has been shown to give an overestimate [107]. Fully correcting for wavefront effects requires the wavefront profiles of both interfering beams to be known, and as such it has been suggested that in precision interferometry well aligned TEM₀₀ mode beams should be ensured, for example by spatial filtering [107]. Where possible, relatively small increases in beam radius can minimise the effects of errors in the wavefront correction by reducing the scale of the correction required, for example by increasing the beam waist radius to 1.5 mm the fractional error introduced can be reduced to 4.5×10^{-9} for a wavelength of 632.8 nm, and any errors in the correction introduced by small deviations from the ideal TEM₀₀ mode will also be at the parts in 10^9 level.

As previously described, updating equation 3.1 to include the effective wavelength given by equation 3.8 results in equation 3.3, where the wavefront correction factor F_λ is given by

$$F_\lambda = \frac{1}{1 - \left(\frac{\lambda_1}{2\pi w_0}\right)^2} - 1, \quad (3.10)$$

$$= \frac{c^2}{(2\pi n_1 f w_0)^2 - c^2}. \quad (3.11)$$

Whilst numerical uncertainties in the beam waist radius, optical frequency and refractive index could be propagated through equation 3.11, uncertainties in the deviations from the ideal TEM₀₀ wavefronts assumed by equation 3.11 are likely to dominate [105]. As such, the fractional uncertainty in F_λ will be large unless precise measurements of the interfering wavefronts are made, with a fractional uncertainty of 15 % assigned even when careful wavefront measurements were made [105]. As the change in refractive index approaches zero, the final term of equation 3.6 will approximate Δx , and therefore a 50 % fractional uncertainty in F_λ will contribute a fractional uncertainty of approximately 5×10^{-9} to the measured displacement for a wavelength of 632.8 nm and a beam waist radius of 1 mm.

Angular misalignments and shearing between the beams are an additional source of wavefront errors. In general, in precision interferometry, angular misalignments between the beams will be minimised, as will beam shear. Where significant angular misalignments between the interfering beams are present, or where beam shearing is unavoidable, exact solutions for the effect on the phase of the integrated interference pattern are available [104], with the restriction that the full beams must be detected. The authors suggest detectors three times larger than the beam radii are required for this condition to be fulfilled [104].

Geometrical uncertainty

Geometrical cosine and Abbé errors are not directly related to the topic of interferometry, however, where practical displacement measurements are to be made, they must be considered.

Cosine errors occur where there is a misalignment between the measurement axis and the axis along which the displacement is intended to be measured, and, in the simple case considered in figure 3.1 **A**, introduce an error into the measured displacement of

$$\Delta x_{\cos} = (1 - \cos(\theta)) \Delta x, \quad (3.12)$$

where Δx is the measured displacement and θ is the angular misalignment between the measurement and displacement axes. The exact form of cosine errors in displacement measurements made with optical interferometry will depend on the optical configuration employed, for example, whether the interferometer employs a plane mirror or corner cube retroreflector as the measurement target. In the plane mirror case, two potential sources of cosine error must be considered: misalignments between the laser beam axis and the mirror normal, and misalignments between the intended measurement axis and the mirror normal.

Abbé errors arise where there is a spatial offset between the measurement and displacement axes, and where rotations take place about a point on the displacement axes, as may be the case if for example a displacement stage pitches during a measurement of a target placed on top of the stage. The Abbé error is given by

$$\Delta x_{\text{Abbé}} = d \tan(\alpha), \quad (3.13)$$

for an Abbé offset d and rotation error α .

These two forms of geometrical errors are compared in figure 3.1, where it can

be seen that for equal angular errors the Abbé error may be significantly larger in magnitude where an Abbé offset exists. As a result, the Abbé error may often be the largest uncertainty contribution in dimensional metrology, unless extreme care is taken to eliminate them [108].

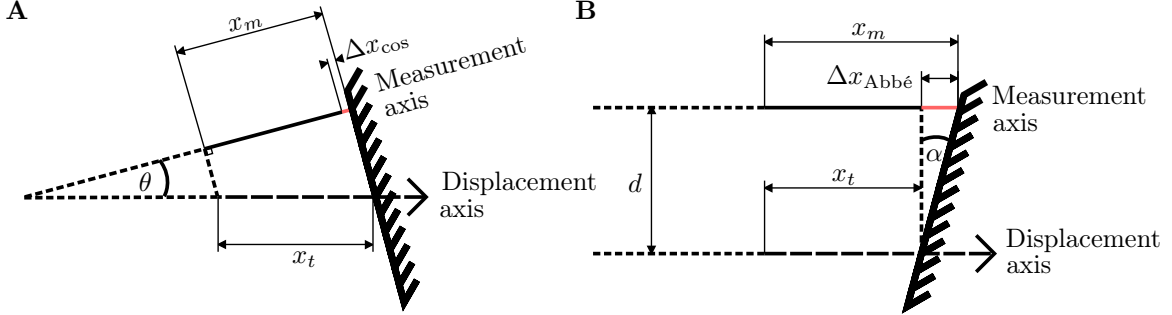


Figure 3.1: Geometric representation of cosine (A) and Abbé (B) errors. In both cases the mirror is displaced a distance x_t along the displacement axis and a displacement x_m is measured along the measurement axis. Angles θ and α are equal for comparison. The effects of reflection from the mirror at non-normal incidence are neglected here for ease of comparison.

Methods for minimising both cosine [109] and Abbé errors in interferometry have been described in the literature [108], and should be employed where equations 3.12 and 3.13 indicate the magnitude of these error sources may be significant relative to the desired measurement uncertainty.

A final source of errors that will be grouped with geometrical errors in this work are errors due to mechanical drift or vibration in the reference arm of the interferometer, which are indistinguishable in interferometric measurements from displacements in the measurement arm. These are grouped with cosine and Abbé error terms here as they do not represent an error in the interferometric displacement measurements but an error in the relationship between that measurement and the desired displacement measurement. In precision applications such problems can be avoided with the use of differential interferometer designs (see, for example, [50, 54]). In many cases it is possible to design the measurement system such that the measurand is the differential displacement between two mirrors, eliminating errors due to unwanted motion of the reference mirror. Where this is not possible, the mechanical stability of the reference arm of the interferometer must be greater than the desired measurement accuracy.

3.2.4 Combined standard uncertainty

The combined standard uncertainty can be calculated from the square root of the variance of the displacement, which may be calculated, neglecting geometrical errors, using equation 3.6. Geometrical errors must be calculated on a case-by-case basis, however, equations for estimating the magnitude of such errors are given in the previous section, intended only to provide an indication of whether a more in depth analysis of the geometrical errors is warranted.

Although some contributory variances may follow a non-Gaussian distribution, notably the uncertainty in the measured change in interferometric phase in cases where significant non-linearities are present, due to the large number of contributory variances the distribution of errors in the measured displacement will approximate the Gaussian due to the Central Limit Theorem [89]. As such, expanded uncertainties may be calculated with standard techniques, with expanded uncertainties with a 95 % level of confidence being given by multiplying the standard uncertainty by a coverage factor of 1.960 [89].

3.3 Phase measurement uncertainty budget

3.3.1 The measurement problem

This section will outline the phase measurement process for homodyne interferometry. Some information that will be presented in section 3.4 is repeated here to permit the uncertainty budget to be utilised in isolation, and to place the work presented in sections 3.4 in context.

Homodyne optical interferometers aim to measure a change in interferometric phase, however, the phase is not measured directly. Instead, the interferometer, under the assumptions considered in this work, produces three optical outputs: two quadrature signals (P_x and P_y), and an intensity reference signal P_r . Ideally the quadrature outputs will take the form of sinusoids with equal amplitudes, minima of 0 W, and a 90° phase difference. The reference output P_r will ideally be a fixed percentage of the input optical power.

These optical outputs, following the assumptions made in this chapter, are incident on photodiode detectors, which produce photocurrents (i_x , i_y and i_r). The photocurrents are then amplified and converted to voltages (V_x , V_y and V_r), which are digitised, and recorded.

In order to minimise the effects of fluctuations in the output power of the illuminating source, the quadrature voltage signals V_x and V_y are normalised to V_r . For this normalisation process to be effective, any offsets in the voltage signals that are present in the absence of laser illumination (dark offsets) must be eliminated. Throughout this chapter dark offsets, the voltage offsets that are present in the absence of deliberate illumination, in the voltage signals are assumed to have been eliminated by measuring the voltage signals with the illuminating source blocked, and subtracting the measured dark voltages from the voltage signals. Normalising by V_r then yields the normalised quadrature signals S_x and S_y , given by

$$S_x = \frac{V_x}{V_r}, \quad (3.14)$$

$$S_y = \frac{V_y}{V_r}. \quad (3.15)$$

The normalised quadrature signals S_x and S_y will be offset from the origin as a result of both the dark offset subtraction process, and any offset in the optical signals due to imperfect fringe contrast. The quadrature phase difference between the

signals may also deviate from the ideal case, and the amplitudes of the normalised quadrature signals may not be identical. Such deviations will, if uncorrected, introduce errors into the measured interferometric phase [21]. The ideally sinusoidal quadrature signals may also be distorted by the presence of unwanted harmonics due to multiple reflections, polarisation leakage within the optics of the interferometer [85], and electronic effects, which will introduce further errors into the phase measurement. All these error sources are periodic in nature; the error will repeat with some harmonic of the illuminating wavelength if the source of the error remains constant. As a result, these error sources do not accumulate with displacement. Such errors are commonly described as non-linearities, and the term non-linearities will be considered in this work to refer only to errors in the measured phase that are periodic with some harmonic of the wavelength.

The effects of non-ideal amplitudes, offsets and quadrature phase difference in the normalised quadrature signals may be corrected for with an ellipse fitting non-linearity correction [21, 87], commonly termed the Heydemann correction. The effects of harmonic distortions, in contrast, could not be corrected for in a purely optical interferometric system prior to the work presented in chapter 6 of this thesis. Such corrections are beyond the scope of this uncertainty budget, and discussion of these techniques is reserved for chapter 6.

The application of an ellipse fitting Heydemann correction will, however, be considered here, and this is achieved by fitting a conic section of the form [21, 87]

$$AS_x^2 + BS_y^2 + CS_xS_y + DS_x + ES_y = 1 \quad (3.16)$$

to the normalised quadrature signals, where A to E are the ellipse fit coefficients, as discussed in detail in section 6.2.1. Denoting, for clarity, individual data points of the quadrature signals S_x and S_y as x_i and y_i , for i in the range 0 to N , equation 3.16 may be written in matrix form as

$$\begin{pmatrix} x_0^2 & y_0^2 & x_0y_0 & x_0 & y_0 \\ x_1^2 & y_1^2 & x_1y_1 & x_1 & y_1 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ x_N^2 & y_N^2 & x_Ny_N & x_N & y_N \end{pmatrix} \begin{pmatrix} A \\ B \\ C \\ D \\ E \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{pmatrix}, \quad (3.17)$$

where the matrix equation may be solved for the fit coefficients. Left multiplying both sides of equation 3.17 by the transpose of the quadrature signal matrix,

following the ordinary least squares technique, results in

$$\begin{pmatrix} x_0^2 & x_1^2 & \dots & x_N^2 \\ y_0^2 & y_1^2 & \dots & y_N^2 \\ x_0 y_0 & x_1 y_1 & \dots & x_N y_N \\ x_0 & x_1 & \dots & x_N \\ y_0 & y_1 & \dots & y_N \end{pmatrix} \begin{pmatrix} x_0^2 & y_0^2 & x_0 y_0 & x_0 & y_0 \\ x_1^2 & y_1^2 & x_1 y_1 & x_1 & y_1 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ x_N^2 & y_N^2 & x_N y_N & x_N & y_N \end{pmatrix} \begin{pmatrix} A \\ B \\ C \\ D \\ E \end{pmatrix} = \begin{pmatrix} x_0^2 & x_1^2 & \dots & x_N^2 \\ y_0^2 & y_1^2 & \dots & y_N^2 \\ x_0 y_0 & x_1 y_1 & \dots & x_N y_N \\ x_0 & x_1 & \dots & x_N \\ y_0 & y_1 & \dots & y_N \end{pmatrix} \begin{pmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{pmatrix}, \quad (3.18)$$

$$\begin{pmatrix} \Sigma x_i^4 & \Sigma x_i^2 y_i^2 & \Sigma x_i^3 y_i & \Sigma x_i^3 & \Sigma x_i^2 y_i \\ \Sigma x_i^2 y_i^2 & \Sigma y_i^4 & \Sigma x_i y_i^3 & \Sigma x_i y_i^2 & \Sigma y_i^3 \\ \Sigma x_i^3 y_i & \Sigma x_i y_i^3 & \Sigma x_i^2 y_i^2 & \Sigma x_i^2 y_i & \Sigma x_i y_i^2 \\ \Sigma x_i^3 & \Sigma x_i y_i^2 & \Sigma x_i^2 y_i & \Sigma x_i^2 & \Sigma x_i y_i \\ \Sigma x_i^2 y_i & \Sigma y_i^3 & \Sigma x_i y_i^2 & \Sigma x_i y_i & \Sigma y_i^2 \end{pmatrix} \begin{pmatrix} A \\ B \\ C \\ D \\ E \end{pmatrix} = \begin{pmatrix} \Sigma x_i^2 \\ \Sigma y_i^2 \\ \Sigma x_i y_i \\ \Sigma x_i \\ \Sigma y_i \end{pmatrix}, \quad (3.19)$$

$$\mathbf{H}_M \begin{pmatrix} A \\ B \\ C \\ D \\ E \end{pmatrix} = \mathbf{H}_V \quad (3.20)$$

where the sums are over all data points to be fitted. Equivalently, the sums may be replaced by averages of the products of powers of the quadrature signals. The vector of fit coefficients may then be solved for with standard matrix techniques. Estimates of offsets in the quadrature signals, the gain ratio between the quadrature signals and the quadrature phase may then be calculated from the fitted coefficients. These signal parameters are given by [87]

$$\alpha_x = \frac{2BD - EC}{C^2 - 4AB}, \quad (3.21)$$

$$\alpha_y = \frac{2AE - DC}{C^2 - 4AB}, \quad (3.22)$$

$$G_{xy} = \sqrt{\frac{B}{A}}, \quad (3.23)$$

$$\delta = \cos^{-1} \left(\frac{C}{\sqrt{4AB}} \right) - \pi, \quad (3.24)$$

where α_x and α_y are the offsets of S_x and S_y respectively, G_{xy} is the amplitude ratio of S_x to S_y , and δ is the quadrature phase shift between the signals. The quadrature

phase error, the deviation from the ideal 90° value, is given by

$$\delta' = \sin^{-1} \left(\frac{C}{\sqrt{4AB}} \right). \quad (3.25)$$

Following the original methodology of [87], corrected signals C_x and C_y are then given by

$$C_x = S_x - \alpha_x, \quad (3.26)$$

$$C_y = \frac{1}{\cos(\delta')} \left((S_x - \alpha_x) \sin(\delta') + G_{xy} (S_y - \alpha_y) \right), \quad (3.27)$$

with the instantaneous interferometric phase calculated from the corrected signals using the four quadrant arctangent function as

$$\varphi = \text{atan2} \left(\frac{C_y}{C_x} \right). \quad (3.28)$$

Alternatively, following the methodology of [110], the phase may be calculated directly from the fit coefficients as

$$\varphi = \text{atan2} \left(\frac{\sqrt{4AB - C^2} (E + CS_x + 2BS_y)}{2B(D + 2AS_x) - C(E + CS_x)} \right), \quad (3.29)$$

where in this case the fractional representation of the four quadrant arctangent function has been employed for convenience of notation. The methodology of [110] will ultimately be followed here, permitting uncertainties in the interferometric phase to be calculated directly, without the need to evaluate the covariance between the corrected quadrature signals.

The four quadrant arctangent function produces an instantaneous phase value in the interval $(-\pi/2, \pi/2]$, and therefore in order to measure phase changes over a range greater than 2π , fringes must be ‘counted’, commonly referred to as phase unwrapping [22]. Fringe counting is achieved by tracking the integer number of fringes (2π periods) covered, incrementing the total count N_F each time the instantaneous phase crosses from $-\pi/2$ to $\pi/2$, and decrementing N_F on each crossing from $\pi/2$ to $-\pi/2$. As long as the spatial sampling rate of the interferometer is sufficiently high for no phase change of greater than π to take place between two measurement points, no additional uncertainty is introduced by this counting process.

Finally, in order to measure a phase change, the difference between two interferometric phase values recorded at positions $N_{F,0} + \varphi_0$ and $N_{F,1} + \varphi_1$ must be taken, resulting in final measured phase change of

$$\Delta\varphi = N_{F,1} + \varphi_1 - N_{F,0} - \varphi_0. \quad (3.30)$$

3.3.2 Mathematical model

For the purposes of uncertainty propagation, as no uncertainty is associated with the integer fringe count, $N_{F,0}$ and $N_{F,1}$ may be neglected. Uncertainties in the

ellipse fit coefficients will also not be considered here. Instead, uncertainty arising from residual non-linearities will be included in the phase measurement uncertainty as a separate term accounting for both elliptical and higher order non-linearity effects, with a fixed value of the peak-to-peak non-linearity established through a combination of experimental measurement and consideration of the interferometer design. Additionally, it will be assumed that the uncertainties in the normalised quadrature signals at different displacements are uncorrelated.

Following these assumptions, the uncertainties in individual instantaneous phase measurements are uncorrelated, and the uncertainty in a phase difference measurement is determined by

$$\sigma_{\Delta\varphi} = \sqrt{\sigma_{\varphi_0}^2 + \sigma_{\varphi_1}^2}. \quad (3.31)$$

As a result, this uncertainty budget will consider the uncertainty in an individual measurement of the instantaneous phase, φ , determined by

$$\varphi = \text{atan2} \left(\frac{\sqrt{4AB - C^2} (E + CS_x + 2BS_y)}{2B(D + 2AS_x) - C(E + CS_x)} \right) + e_{\text{NL}}, \quad (3.32)$$

where

$$S_x = \frac{V_x}{V_r}, \quad (3.33)$$

$$S_y = \frac{V_y}{V_r}, \quad (3.34)$$

and e_{NL} is a dummy error variable introduced to represent the uncertainty contribution of the non-linearity, and has a mean value of zero. The measurand φ therefore has the following input quantities:

- V_x , the cosine output voltage measurement,
- V_y , the sine output voltage measurement,
- V_r , the intensity reference voltage measurement,
- e_{NL} , the peak-to-peak phase non-linearity,
- A to E , the ellipse fit coefficients,

with all input values other than the ellipse fit coefficients contributing to the measurement uncertainty.

3.3.3 Contributory variances

The uncertainty in a measurement of the instantaneous interferometric phase may be considered to have two components: a statistical uncertainty, arising from noise in the voltage signals (having a number of both optical and electronic origins), and an additional uncertainty term arising as a result of uncorrected (systematic) non-linearity errors. For the estimation of the statistical uncertainty, two cases will be

considered here. Firstly, the case where repeated phase measurements may be made, and a type A [89] uncertainty estimate calculated, and secondly, the case where repeat measurement is not possible, and the phase uncertainty in a single data point must be evaluated using type B techniques [89]. In general, the type A approach is preferable [89], and in many metrological applications where an uncertainty budget is required repeat measurement is possible, for example, when calibrating a nanopositioning stage. As such, this is the approach that this work will focus upon. In cases where repeat measurement is not possible, techniques for propagating statistical uncertainties in the voltage signals to an uncertainty in the measurand, the instantaneous interferometric phase, will be considered. The equations derived in the type B case will also be used later in this chapter to investigate the ultimate performance limits of homodyne interferometry in the shot noise limited case.

In the type A case, the statistical variance in the instantaneous phase measurement may be calculated from repeated measurements using standard techniques [89]. The variance of the interferometric phase may then be calculated as

$$\sigma_{\varphi}^2 = \sigma_{\varphi, \text{Stat.}}^2 + \sigma_{\text{NL}}^2, \quad (3.35)$$

where $\sigma_{\varphi, \text{Stat.}}^2$ is the statistical variance in the measured instantaneous phase, and σ_{NL}^2 is the estimate variance in the instantaneous phase due to residual non-linearity errors.

In the type B case, variances and covariances in the measured voltage signals must be estimated, and propagated through the mathematical model of the measurement. As the input quantities V_x , V_y and V_r are not employed directly in equation 3.32, but are instead used to calculate S_x and S_y through equations 3.33 and 3.34, expressions for variances and covariances in S_x and S_y , in terms of variances and covariances in the voltage signals, will be derived. Following equation 3.32, the variance in the instantaneous phase measurement, including the effects of variances in and covariances between the ellipse fit coefficients, is derived in appendix B. The propagation of non-linearities through equation 3.32 is not new, having previously been described in [110, 111]; it is repeated in this work for consistency of notation. Neglecting the uncertainty due to the fit coefficients, and including an additional term (σ_{NL}^2) to account for the variance introduced by residual non-linearities, results in the expression for the variance in the instantaneous phase measurement that will

be used in the type B case:

$$\begin{aligned}\sigma_\varphi^2 = & (4AB - C^2) \left(\sigma_\varphi^2|_{S_x, S_x} \sigma_{S_x}^2 + \sigma_\varphi^2|_{S_y, S_y} \sigma_{S_y}^2 + \right. \\ & \left. \sigma_\varphi^2|_{S_x, S_y} \sigma_{S_x, S_y} \right) / 4 \left(4AB^2 S_y^2 \right. \\ & \left. + B \left(4AS_y(CS_x + E) + (2AS_x + D)^2 - C^2 S_y^2 \right) \right. \\ & \left. - (CS_x + E)(C(AS_x + CS_y + D) - AE) \right)^2 + \sigma_{\text{NL}}^2,\end{aligned}\quad (3.36)$$

$$\sigma_\varphi^2|_{S_x, S_x} = (C(CS_y + D) - 2A(2BS_y + E))^2, \quad (3.37)$$

$$\sigma_\varphi^2|_{S_y, S_y} = (C(CS_x + E) - 2B(2AS_x + D))^2, \quad (3.38)$$

$$\begin{aligned}\sigma_\varphi^2|_{S_x, S_y} = & -2(2B(2AS_x + D) - C(CS_x + E)) \\ & (2A(2BS_y + E) - C(CS_y + D)).\end{aligned}\quad (3.39)$$

Equivalently, uncertainties may be propagated through equations 3.26, 3.27 and 3.28. The variances and covariance of the corrected quadrature signals are given in terms of the signal parameters, whilst neglecting the uncertainty in the signal parameters, by

$$\sigma_{C_x}^2 = \sigma_x^2, \quad (3.40)$$

$$\sigma_{C_y}^2 = \tan^2(\delta') \sigma_x^2 + \frac{G_{xy}^2}{\cos^2(\delta')} \sigma_y^2 + 2G_{xy} \frac{\sin(\delta')}{\cos^2(\delta')} \sigma_{x,y}, \quad (3.41)$$

$$\sigma_{C_x, C_y} = \tan(\delta') \sigma_x^2 + \frac{G_{xy}}{\cos(\delta')} \sigma_{x,y}. \quad (3.42)$$

Equations 3.40 and 3.41 were again derived with the partial derivative method [89]. Equation 3.42 was derived by applying the definition of the covariance [112] and evaluating expectation values of C_x , C_y , and the product of C_x and C_y , assuming a bivariate Gaussian distribution [52] for S_x and S_y . This derivation is presented in appendix A. The corresponding variance in the instantaneous phase measurement is then given, again following the partial derivative technique [89], by

$$\sigma_\varphi^2 = \sigma_{C_x}^2 \left(\frac{\partial \varphi}{\partial C_x} \right)^2 + \sigma_{C_y}^2 \left(\frac{\partial \varphi}{\partial C_y} \right)^2 + 2\sigma_{C_x, C_y} \frac{\partial \varphi}{\partial C_x} \frac{\partial \varphi}{\partial C_y} + \sigma_{\text{NL}}^2 \left(\frac{\partial \varphi}{\partial e_{\text{NL}}} \right)^2 \quad (3.43)$$

$$= \frac{\sigma_{C_x}^2 C_y^2 + \sigma_{C_y}^2 C_x^2 - 2\sigma_{C_x, C_y} C_x C_y}{(C_x^2 + C_y^2)^2} + \sigma_{\text{NL}}^2. \quad (3.44)$$

Equations 3.36 to 3.39 will be employed in this work to avoid the intermediate calculation of the variances and covariances in the corrected quadrature signals, however, equations 3.40 to 3.42, along with equation 3.44, are also reported here as they are beneficial when investigating the scaling effects of the signal parameters G_{xy} and δ' on the phase uncertainty.

This section will first consider the estimation of the statistical variance in the

instantaneous phase, with the discussion broken down into type A and type B approaches. Following this, methods for estimating the variance due to uncorrected non-linearity errors will be discussed.

Statistical uncertainty

Type A evaluation

In the type A case the statistical variance in the instantaneous phase is evaluated using standard techniques [89], from repeated measurements of the instantaneous phase. By evaluating the statistical variance in this manner, the effects of mechanical vibrations and air turbulence will also be included in the uncertainty estimate. The type A approach is particularly appropriate when the outputs of two displacement measuring sensors with differing bandwidths must be compared, or where two sensors must be compared and synchronous measurement is not possible.

Type B evaluation

Uncertainties in the measured voltage signals will arise as a result of statistical noise, quantisation errors introduced by the finite resolution of analogue to digital converters, and non-linearity in the photodiode response, current to voltage amplification, and digitisation electronics. In a well designed measurement system quantisation errors will be below the noise floor, and as such this effect will not be considered here. Non-linearity in the measurement electronics will contribute to the overall phase measurement non-linearity, therefore these effects will also not be considered. Scale errors will not result in phase measurement uncertainty, as any such errors will be compensated for by the Heydemann correction. Any offsets in the voltage measurement will be eliminated by the dark voltage measurement subtraction process, as previously described in the phase measurement assumptions. As a result, only uncertainties arising from noise in the quadrature signals will be considered here.

In experimental work, it is desirable to measure the variances in, and covariances between, the three voltage signals (V_x , V_y and V_r) directly, without attempting to separate optical and electronic noise sources. The variances and covariance of the voltage signals may be separated from interference effects by blocking either the measurement or reference arm of the interferometer, such that each output of the interferometer measures a constant optical power. If the input optical power to the interferometer is then stepped through a range of values, the variances and covariances of the three channels can be evaluated directly at each power level. This approach is, however, limited, as by blocking one arm of the interferometer, the maximum optical power at the output is reduced. Additionally, the ratios between the three outputs will remain fixed whilst varying the input optical power, and therefore covariances are only measured for limited combinations of voltage values. If some assumptions can be made about the form of the uncertainties in the voltage signals, it may be possible to extrapolate the voltage dependent variances and covariances from variance and covariance measurements made in this manner, however, the exact form the voltage signal uncertainties will be highly dependent on the measurement electronics, and is therefore beyond the scope of this uncertainty budget.

Assuming variances and covariances in the voltage signals can be estimated for a given interferometer system, the variances of the normalised quadrature signals S_x and S_y can be calculated from equations 3.33 and 3.34 with the partial derivative technique [89], including the effects of the covariance, as

$$\sigma_{S_x}^2 = \left(\frac{1}{V_r}\right)^2 \sigma_{V_x}^2 + \left(\frac{V_x}{V_r^2}\right)^2 \sigma_{V_r}^2 - 2\frac{V_x}{V_r^3} \sigma_{V_x, V_r}, \quad (3.45)$$

$$\sigma_{S_y}^2 = \left(\frac{1}{V_r}\right)^2 \sigma_{V_y}^2 + \left(\frac{V_y}{V_r^2}\right)^2 \sigma_{V_r}^2 - 2\frac{V_y}{V_r^3} \sigma_{V_y, V_r}. \quad (3.46)$$

The covariance between S_x and S_y can be approximated applying the definition of the covariance [112], and evaluating (approximate) expectation values. This derivation is presented in appendix C and, assuming a multivariate Gaussian [52], distribution results in

$$\begin{aligned} \sigma_{S_x, S_y} = \frac{1}{V_r^6} & \left(V_r^4 \sigma_{V_x, V_y} - V_r^3 (V_x \sigma_{V_y, V_r} + V_y \sigma_{V_x, V_r}) \right. \\ & + V_r^2 (\sigma_{V_r}^2 (V_x V_y + 3\sigma_{V_x, V_y}) + 5\sigma_{V_x, V_r} \sigma_{V_y, V_r}) \\ & - 5V_r \sigma_{V_r}^2 (V_x \sigma_{V_y, V_r} + V_y \sigma_{V_x, V_r}) + \sigma_{V_r}^4 (2V_x V_y + 3\sigma_{V_x, V_y}) \\ & \left. + 12\sigma_{V_r}^2 \sigma_{V_x, V_r} \sigma_{V_y, V_r} \right). \quad (3.47) \end{aligned}$$

It should be noted that as a result of equation 3.47 being derived based upon the assumption that V_x , V_y and V_r follow a trivariate Gaussian distribution, the analysis is limited to the case $V_r \gg 0$, as S_x and S_y are asymptotic as the reference voltage approaches zero. As an additional consequence of this assumption, the results may be invalid in cases where the V_x or V_y approach zero, and the voltage noise is shot noise limited, as shot noise follows a Poisson distribution [113]. In practical measurement, this is unlikely to be a limitation, as for small voltage signals the noise floor of the detection and amplification electronics is likely to dominate, and due to the large number of contributing sources to this noise floor, it is likely to follow a Gaussian distribution.

Non-linearity uncertainty

This section will consider the non-linearity of homodyne interferometers from three perspectives. Firstly, the effects of elliptical non-linearities, which are introduced by the previously described quadrature phase, DC offset and gain ratio errors, will be considered. This group of non-linearities distort the Lissajous figure formed by the two quadrature signals from the ideal origin centred circular form to an ellipse, or an offset circle [21]. Secondly, higher order non-linearities, introduced by the presence of unwanted harmonics [85] will be considered. Finally, methods for the practical evaluation of non-linearities will be discussed, along with the distribution of non-linearity errors, so as to allow the discussed effects to be included in the phase measurement uncertainty budget.

Elliptical non-linearities

Elliptical non-linearities are corrected for by the Heydemann correction [87], a subject area that has been extensively studied in the literature [21, 83, 87, 111], and that is discussed in detail in section 6.2.1 of this thesis. Such non-linearities will therefore not typically directly affect the measurement uncertainty. Instead, the uncertainty implications of ellipse fitting corrections will be considered from the perspective of the scaling effect of the correction on the existing noise in the quadrature signals, and from the perspective of the additional uncertainty contribution from the uncertainties in the fit coefficients. As previously stated, it is often preferable to neglect the uncertainty in the ellipse fit coefficients, as if higher order non-linearities are present a measurement of residual non-linearity must be included in the uncertainty budget. If both a measurement of residual non-linearities and the uncertainties in the ellipse fit coefficients are included in the phase measurement uncertainty calculation, the contribution of the uncertainty in the ellipse fit coefficients will be accounted for twice. Nevertheless, as the effects of uncertainties in the ellipse fit coefficients are discussed in the literature [110, 111], this contribution will be reviewed briefly here.

The scaling of uncertainties in the quadrature signals by the Heydemann correction may be observed by inspecting equations 3.40 to 3.42. The first conclusion that may be drawn is that even if the quadrature signals were uncorrelated, if there is a quadrature phase error to be corrected for, the corrected quadrature signals will be correlated to an extent. This provides further evidence that the covariance of the quadrature signals in homodyne interferometry should not be neglected. Secondly, it can be seen from equations 3.40 to 3.42 that the signal offsets do not affect the measurement uncertainty if an ellipse fitting correction is applied. Similarly, the absolute values of the amplitudes of each signal do not directly effect the measurement uncertainty after applying an ellipse fitting correction, although the ratio of the amplitudes does have a scaling effect. It is important to note that whilst the signal offsets do not appear in equations 3.40 to 3.42, optical offsets introduced by poor fringe contrast have the effect of reducing the signal to noise ratios of the measured quadrature signals, an effect that is in this case already accounted for in the uncertainty in the normalised quadrature signals.

If an interferometer design can be guaranteed not to be affected by higher order non-linearities, it may be desirable to estimate residual non-linearity based upon fit coefficient uncertainties. Variances and covariances in the fit coefficients may be estimated according to [110] as

$$\begin{pmatrix} \sigma_{A,A} & \sigma_{A,B} & \sigma_{A,C} & \sigma_{A,D} & \sigma_{A,E} \\ \sigma_{A,B} & \sigma_{B,B} & \sigma_{B,C} & \sigma_{B,D} & \sigma_{B,E} \\ \sigma_{A,C} & \sigma_{B,C} & \sigma_{C,C} & \sigma_{C,D} & \sigma_{C,E} \\ \sigma_{A,D} & \sigma_{B,D} & \sigma_{C,D} & \sigma_{D,D} & \sigma_{D,E} \\ \sigma_{A,E} & \sigma_{B,E} & \sigma_{C,E} & \sigma_{D,E} & \sigma_{E,E} \end{pmatrix} = \frac{\sum (Ax_i^2 + By_i^2 + Cx_iy_i + Dx_i + Ey_i - 1)^2}{n - 5} \mathbf{H}_M^{-1}, \quad (3.48)$$

where the sum gives the sum of the squares of the fit residuals. Equation 3.48 is an

approximation that does not account for uncertainties in the quadrature signals, and in cases where only a small number of points are included in the fit, equation 3.48 may not correctly describe the variances and covariances of the ellipse fit parameters. In such cases, a more rigorous ellipse fitting approach may be required. Ellipse fitting, and the propagation of uncertainty through ellipse fitting algorithms is a complex subject area, with a variety of methodologies possible. Discussion of these techniques is therefore reserved for chapter 6.2.1, where a comparison of a range of approaches proposed in the literature is made, considering both the performance of the fitting algorithm, and the accuracy of any proposed uncertainty propagation methodology.

However they are estimated, variances and covariances in the fit coefficients A to E , given by equation 3.48, along with the quadrature signals S_x and S_y may be propagated through equation 3.32 using the equations presented in appendix B. As an alternative to the equations presented in appendix B, an approximation for calculating the phase uncertainty whilst accounting for the uncertainty in the ellipse fit coefficients has also been proposed in the literature [110]. This approximation is valid for quadrature signals with equal amplitudes and equal uncertainties. In homodyne interferometry, the uncertainties in the quadrature signals are phase dependent, and the amplitudes of the quadrature signals may not be equal, therefore this approximation is not used in this work.

Higher order non-linearities

Higher order non-linearities are not corrected for by ellipse fitting corrections, and will therefore effect the phase uncertainty directly [85]. Such non-linearities are highly variable in form, often arising as a result of unwanted reflections within the interferometer optics, the alignment of which may change with both time and displacement. Higher order non-linearities may also arise from distortion introduced by the measurement electronics, although with a detection, amplification and digitisation system that is well matched to the desired accuracy of the interferometer system this should not be a significant uncertainty source. As a result of the displacement and time sensitivity of such non-linearities, a general analysis is not possible, although sources can be identified for individual interferometers [59–61, 63, 65, 84, 114, 115], and a modelling technique is developed in chapter 5 designed to analyse the sources of higher order non-linearities when designing interferometers [85].

Whilst a general analysis is not possible, one important point that has not previously been realised in the literature should be addressed; where higher order non-linearities are present, phase distortions can occur without a corresponding radial

distortion to the Lissajous figure. This may be demonstrated with simulated quadrature signals of the form

$$S_x = \sum_{i=0}^{\infty} \frac{1}{4^i} \cos((i+1)\varphi), \quad (3.49)$$

$$= \frac{30}{17 - 8 \cos(\varphi)} - 2, \quad (3.50)$$

$$S_y = \sum_{i=0}^{\infty} \frac{1}{4^i} \sin((i+1)\varphi), \quad (3.51)$$

$$= \frac{16 \sin(\varphi)}{17 - 8 \cos(\varphi)}. \quad (3.52)$$

Maxima and minima may be found by differentiation, and occur at

$$\max(S_x) = +\frac{4}{3}, \quad (3.53)$$

$$\min(S_x) = -\frac{4}{5}, \quad (3.54)$$

$$\max(S_y) = +\frac{16}{15}, \quad (3.55)$$

$$\min(S_y) = -\frac{16}{15}. \quad (3.56)$$

Zero centred ± 1 normalised signals can then be obtained as

$$C_x = \frac{8 - 17 \cos(\varphi)}{8 \cos(\varphi) - 17}, \quad (3.57)$$

$$C_y = \frac{15 \sin(\varphi)}{17 - 8 \cos(\varphi)}. \quad (3.58)$$

In polar coordinates C_x and C_y can be seen to describe a unit radius circle as

$$C_x^2 + C_y^2 = 1, \quad (3.59)$$

with the phase given by

$$\text{atan2}\left(\frac{17 \cos(\varphi) - 8}{15 \sin(\varphi)}\right) \neq \varphi. \quad (3.60)$$

This effect is demonstrated for equations 3.57 and 3.58 in figure 3.2, where the distortion introduced by the presence of higher order terms can clearly be seen when C_x and C_y are plotted against φ , whilst the Lissajous figure remains circular. The loss constant of 4 used in this demonstration was selected to give a clearly visible effect in figure 3.2, however similar results may be found for larger values, corresponding to more strongly attenuated higher order terms.

An important consequence of this effect is that higher order non-linearities cannot be corrected for based only upon the measured quadrature signals, as the signals

shown in figure 3.2 are indistinguishable from undistorted quadrature signals following a non-linear phase profile introduced by a non-linear displacement profile. This result also demonstrates that higher order non-linearities can not be evaluated by considering the radial magnitudes of the quadrature signals alone, an approach that has previously been proposed in the literature [50].

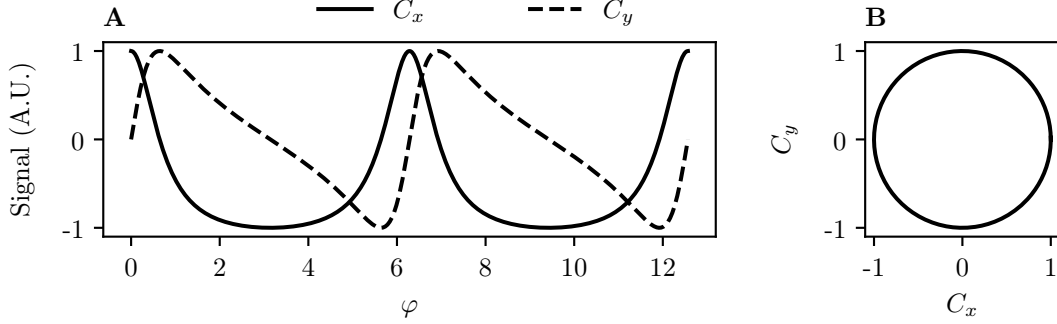


Figure 3.2: Demonstration of phase non-linearities introduced by higher order non-linearities without radial distortion. C_x and C_y are plotted as calculated with equations 3.57 and 3.58.

Evaluation of non-linearity

After applying an ellipse fitting correction, residual non-linearities will often still be present in the interferometric phase measurement, introduced both by errors in the fitting process, and by higher order distortions to the Lissajous figure that arise as a result of the presence of unwanted harmonics in the quadrature signals. Such higher order harmonics will tend to be highly sensitive to the alignment of the interferometer optics, and therefore some caution must be used when applying an individual measurement of residual non-linearities to the same interferometer at a later date. An additional factor that must be considered when evaluating residual non-linearity is that the performance of the ellipse fitting correction is dependent on the number of points included in the ellipse fit. The number of points required to reach a given uncertainty will be dependent on the noise in the quadrature signals, and the ellipse fitting algorithm employed, as demonstrated in section 6.2.1, however, in general several hundred data points in the fit have been suggested to be sufficient for the ellipse fitting algorithm not to introduce significant additional uncertainty [110]. Any measurement of the residual non-linearity may therefore only be valid for cases where the number of points included in the fit, and the spatial sampling rate, are similar to those employed during the displacement measurement. As a result, in-situ non-linearity analysis is desirable, both to provide a check of residual non-linearity after first aligning an interferometer, and to ensure non-linearities do not unexpectedly increase over time due to drift in the alignment of an interferometer set up.

One method of in-situ non-linearity analysis proposed in the literature involves analysing the radial error in the Lissajous figure, and making use of the fact that

for ideally sinusoidal quadrature signals, the Fourier components of the phase error and the radial error divided by the average of the radius vector around the Lissajous figure are equal in magnitude [50]. This technique relies upon the quadrature signals being pure sinusoids, with no harmonic content, and fails to detect higher order non-linearities. As a result, this technique should not be relied upon to estimate non-linearities.

An alternative approach that is capable of resolving non-elliptical non-linearities is to directly compare the measured interferometric displacement to another quantity related to the displacement with a different spatial-frequency space error distribution. This approach has been demonstrated for the evaluation (and correction) of non-linearities by measuring the same displacement with two interferometer operating at two separate wavelengths [69], however, the technique is not restricted to two interferometric measurements. Where the task is to evaluate, rather than correct, for non-linearity errors, the secondary displacement measurement can, for example, be a stage position read out, or the voltage supplied to a piezoelectric element. Whilst such displacement measurements will have a lower accuracy than the interferometric measurement, and may themselves be non-linear or hysteretic relative to displacement, non-linearities in the interferometric measurement can often be resolved by taking the Fourier transform of the difference between the interferometric displacement and the secondary displacement measurement, and identifying peaks at harmonics off the illuminating wavelength, as any errors in the non-optical second displacement measurement are unlikely to be periodic with harmonics of the interferometer wavelength. By taking the sum of the Fourier amplitudes of each identified peak, an upper limit can be placed on the amplitude of the non-linearities. This sum gives an upper limit as it is implicitly assumed that the harmonic components of the non-linearity add constructively, which may not be the case. Alternatively, the Fourier transform can be windowed around the non-linearity harmonics as described in [69], and the windowed transform inverted in order to recover an estimate of the non-linearities that accounts for the phase-dependent value of the non-linearity. Some care must be taken when evaluating non-linearities with Fourier techniques as non-linearity components that are highly variable with time or displacement may not be resolved if the amplitude and phase of the non-linearity component change significantly during the measurement. As such, the results of in-situ techniques should be compared to more reliable direct non-linearity measurements where possible, to build up a detailed picture of the non-linearity of a given interferometer system.

Direct measurements of non-linearities may be made against a more linear displacement measure. Following the recent adoption of the lattice parameter of silicon as a secondary realisation of the metre [8], traceable non-linearity measurements may be made by comparing optical interferometer and X-ray interferometer measurements [54]. In some cases non-linearities can be reduced by iteratively adjusting the optical alignment, however such processes should be avoided when measuring non-linearities for the purpose of creating an uncertainty budget for a given interferometer design. Instead, a fixed alignment routine should be followed, and if variations in the non-linearity magnitudes that cannot be explained by deviations from the fixed alignment routine are observed, the maximum value of the non-linearities

should be taken as the performance limit. Comparison to a more linear interferometer is also possible, however when exploring the performance limits of optical interferometry, a more linear system may not be available.

Having obtained a measurement of the peak-to-peak non-linearity, this must then be related to the interferometric phase uncertainty. A sinusoidal error will follow an arcsine distribution, a special case of the beta distribution [52], with a probability density function given in this case by

$$f(e_{\text{NL}}) = \frac{1}{\pi \sqrt{\left(\frac{\Delta\varphi_{\text{NL}}}{2}\right)^2 - e_{\text{NL}}^2}}, \quad (3.61)$$

where e_{NL} is the non-linearity error in the instantaneous phase and $\Delta\varphi_{\text{NL}}$ is the peak-to-peak non-linearity error. Where non-linearities at multiple spatial periods are present, assuming an arcsine distribution based upon the peak-to-peak non-linearity will tend to overestimate the measurement uncertainty, as the peak-to-peak non-linearity will be dominated by the largest amplitude harmonic present, and the presence of smaller amplitude terms will weight the probability distribution towards the centre (zero error). For the exact distribution of the errors to be known, the non-linearity at each phase value must be known exactly, in which case they would be corrected for, and as such an arcsine distribution will be assumed here as a compromise.

Assuming an arcsine distribution as described by equation 3.61, the variance arising from a non-linearity error with an unknown form, but a known peak-to-peak value of $\Delta\varphi_{\text{NL}}$ is given by [52]

$$\sigma_{\text{NL}}^2 = \frac{\Delta\varphi_{\text{NL}}^2}{8}. \quad (3.62)$$

It should be noted that by assuming that only the peak-to-peak non-linearity is known, and the form is not, the periodic nature of the non-linearity is not exploited. For measurements of a phase difference corresponding to an integer number of wavelengths, the non-linearity error is zero if the non-linearities remain fixed over the course of the measurement. The restriction to phase differences covering an integer number of wavelengths, as opposed to an integer number of optical fringes, arises as a result of the potential for non-linearities with a once every other fringe period to arise in double pass interferometer designs. Exploiting the periodic nature of the non-linearity to reduce measurement uncertainty may be beneficial in cases where a spatial resolution of one optical fringe, typically greater than 100 nm, is sufficient. As previously discussed, however, non-linearities arising as a result of higher order effects can be highly sensitive to the exact alignment of the interferometer optics, and may therefore change in both phase and amplitude whilst measuring a phase difference. In this case, the non-linearity error for a phase difference measurement corresponding to an integer number of wavelengths is not zero, and as such the stability of the non-linearities of individual interferometer designs must be considered on a case-by-case basis when exploiting the periodic nature of the non-linearity.

3.3.4 Combined standard uncertainty

In summary, two approaches to estimating the variance in an instantaneous phase measurement have been described: a type A statistical approach, and a type B approach based on propagation of uncertainties in the measured voltage signals through to an uncertainty in the calculated change in the interferometric phase. In the type A case, the statistical variance in the instantaneous interferometric phase is estimated from repeat measurements following standard techniques. In the type B case, variances and covariance in the measured voltage signals must be obtained. Possible approaches for determining variances and covariance in the measured voltage signals have been discussed, however, the exact methods required will be dependent on the measurement electronics employed. Having obtained the required variances and covariances in the voltage signals, variances and covariances in the normalised quadrature signals S_x and S_y may be calculated using equations 3.45 to 3.47. Equations 3.36 to 3.39 may then be used to calculate the statistical variance in the instantaneous interferometric phase.

In both the type A and type B cases, the statistical variance must be combined with the estimated variance due to the phase non-linearity, which may be calculated from a measurement of the peak-to-peak non-linearity using equation 3.62. Having obtained the variance in the instantaneous phase measurement, the standard deviation in a measurement of the interferometric phase change $\Delta\varphi$ can then be calculated from the sum of two such variances, following equation 3.31.

In general, the probability distribution of the error in the phase difference measurement will not be Gaussian, due to the arcsine distribution of the non-linearity errors. In principle it is possible to calculate the form of the resulting probability distribution by convolving the Gaussian distribution of the statistical uncertainty contributions with the arcsine distribution of the non-linearity contributions for each of the instantaneous phase points to be compared [89], and then convolving the resulting distributions of the errors in each instantaneous phase measurement with each other. As the arcsine distribution of the non-linearity error is only approximate, with non-linearities commonly consisting of multiple harmonics, the Central Limit Theorem will instead be used in this work to estimate coverage factors for expanded uncertainties, and the associated levels of confidence, following the suggestions of [89]. In particular, when considering the displacement uncertainty, rather than the phase difference uncertainty in isolation, the large number of contributory variances suggest that approximations made with the Central Limit Theorem are valid. As the uncertainty in a measurement of the phase difference will rarely be employed in isolation, a detailed analysis of the expanded uncertainty will not be performed here, with the standard deviation being sufficient to allow uncertainties in quantities calculated from the phase difference to be estimated.

3.4 Shot noise limited homodyne interferometry

In this section, the sources and form of shot noise in photovoltaic homodyne detection will be discussed. This will be followed by an analysis of the effects of shot

noise on the intensity normalisation process. Finally, equations describing the variances and covariance of the normalised quadrature signals S_x and S_y in terms of the optical properties of a homodyne interferometer will be derived.

Whilst in practical applications the shot noise may not be a limiting factor, with, even in the case of an ideal optical interferometer, a number of other noise sources arising from the measurement electronics [30], it does provide a fundamental limit to the performance of displacement measuring interferometry at a given optical power and measurement bandwidth. As such, an exploration of the shot noise limit is beneficial to provide a performance baseline when considering other aspects of an interferometer design. The limits derived in this section will also be used throughout this thesis to provide non-arbitrary levels of noise when modelling the outputs of homodyne interferometers, with physically valid covariances. In contrast to the approach employed in section 3.3, where the goal was to allow a practical uncertainty budget to be produced, the goal of this section is to provide equations that may be used to guide the optical design of homodyne interferometers.

3.4.1 Detection noise

In order to measure displacements with optical interferometry optical intensities must be measured. One of the most fundamental limits on interferometric measurements is therefore shot noise, arising from the quantisation of light into photons. In cases where displacements are confined to sufficiently small fractions of a fringe, of order of nanometres or less, squeezed state techniques may be employed to reduce the shot limit [116], however such conditions will usually prove too restrictive for displacement measuring interferometry and as such are beyond the scope of this work.

Assuming conventional laser illumination is used, the shot noise in the detected interferograms will follow the Poisson distribution, as will the shot noise in the photocurrent produced by the photodiode. The Root Mean Squared (RMS) noise introduced by shot noise is therefore determined purely by the number of electrons in the photocurrent [113] in a time window determined by the inverse of the measurement bandwidth Δf , given by

$$N_{\text{electron}} = \frac{\chi_{\lambda} P + i_{\text{dark}}}{q \Delta f}, \quad (3.63)$$

where i_{dark} is the dark current produced by the photodiode in the absence of any illumination, q is the charge on an electron, P is the illuminating optical power and χ_{λ} is the responsivity of the photodiode at the illuminating wavelength, defined as the ratio of the photocurrent to the illuminating optical power. Alternatively the responsivity may be defined in terms of the quantum efficiency of the photodiode as

$$\chi_{\lambda} = \frac{hc}{\lambda q} \text{QE}_{\lambda}, \quad (3.64)$$

where the quantum efficiency is the fraction of incident photons that are detected, allowing the responsivity of an ideal photodiode ($\text{QE}_{\lambda} = 1$) to be calculated. Although photon statistics follow the Poisson distribution, in the large N limit, as

will typically be the case at the optical powers used in optical interferometry, the Poisson distribution approximates the Gaussian [52]. The ratio of the mean of the photocurrent, μ_i , to the standard deviation (due to the shot noise), $\sigma_{i,\text{shot}}$, is then given by [113]

$$\frac{\mu_i}{\sigma_{i,\text{shot}}} = \sqrt{N_{\text{electron}}}, \quad (3.65)$$

resulting in a shot noise standard deviation approximated by

$$\sigma_{i,\text{shot}} = \sqrt{q\Delta f (\chi_\lambda P + i_{\text{dark}})}. \quad (3.66)$$

Thermal, or Johnson, noise also arises in photodiodes as a result of the shunt resistance of real-world diodes. This noise contribution is temperature dependant, and the resulting standard deviation is given by [117]

$$\sigma_{i,\text{thermal}} = \sqrt{\frac{4k_{\text{B}}T\Delta f}{R_{\text{shunt}}}} \quad (3.67)$$

where k_{B} is the Boltzmann constant, T is temperature (in Kelvin) and R_{shunt} is the shunt resistance of the diode.

The total detection noise resulting from these contributions is given by

$$\sigma_{i,\text{detection}} = \sqrt{q\Delta f (\chi_\lambda P + i_{\text{dark}}) + \frac{4k_{\text{B}}T\Delta f}{R_{\text{shunt}}}}. \quad (3.68)$$

3.4.2 Intensity normalisation

All optical sources will have a degree of noise in the intensity of the output, along with a degree of longer term drift, with the noise in real world laser sources being significantly higher than the shot noise limit [118]. As previously stated, for the purposes of this work it will be assumed that an optical intensity reference output will be used to minimise the effects of both longer term drifts and noise in the illuminating optical intensity. In order for this process to be effective, the reference and signal intensities must be detected synchronously and in phase, and the detected signals must be suitably low pass filtered so as to avoid aliasing effects. Whilst in practice this normalisation process may not completely eliminate the effects of laser noise, the ideal shot noise limited case will be considered here.

Even in the ideal case, some correlated noise will remain in the normalised quadrature signals due to the presence of noise in the reference signal. As this normalisation process will take place after the signals have been detected and amplified, a general form of the quadrature and reference optical signals will be used to analyse the effects of the normalisation process, given by

$$P_x = P_0 (\alpha_x + \beta_x \cos(\varphi)), \quad (3.69)$$

$$P_y = P_0 (\alpha_y + \beta_y \cos(\varphi + \delta)), \quad (3.70)$$

$$P_r = P_0 \beta_r, \quad (3.71)$$

for

$$0 \leq \alpha_{x/y} \leq \beta_{x/y}, \quad (3.72)$$

$$0 \leq \beta_r, \quad (3.73)$$

where the offsets α , gain coefficients β and quadrature phase shift δ will be determined by the optical design of the interferometer. In the ideal case

$$\alpha_x = \beta_x = \alpha_y = \beta_y, \quad (3.74)$$

$$\delta = 90^\circ. \quad (3.75)$$

The optical powers will then be detected as voltage signals taking the form

$$V_x = G_x \left(P_x + \gamma_{s,x} \sqrt{P_x} + \gamma_{d,x} \right), \quad (3.76)$$

$$V_y = G_y \left(P_y + \gamma_{s,y} \sqrt{P_y} + \gamma_{d,y} \right), \quad (3.77)$$

$$V_r = G_r \left(P_r + \gamma_{s,y} \sqrt{P_y} + \gamma_{d,y} \right), \quad (3.78)$$

where $G_{x/y/r}$ are gain coefficients and the noise coefficients $\gamma_{s/d,x/y/r}$ are uncorrelated random variables with mean zero and are introduced to model the effects of shot and dark noise terms. As previously discussed, it is assumed that a measurement of the dark level has been made, and subtracted from the signals such that when no laser illumination is present the quadrature and reference voltage signals are centred on zero. Following the derivations of section 3.4.1 and assuming noise introduced by the amplification and digitisation electronics is negligible relative to the detection noise, the standard deviations of the noise coefficients will be given by

$$\sigma_{\gamma_{s,x/y/r}} = \sqrt{q\Delta f}, \quad (3.79)$$

$$\sigma_{\gamma_{d,x/y/r}} = \sqrt{\frac{1}{\chi_\lambda} \left(q\Delta f i_{\text{dark}} + \frac{4k_B T \Delta f}{R_{\text{shunt}}} \right)}, \quad (3.80)$$

where the responsivity has been absorbed into the optical power to voltage gain as implied by equations 3.76 to 3.78.

Dividing by V_r gives the normalised quadrature signals following equations 3.33 and 3.34. Variances and covariances in the normalised signals may then be determined from the definition of the variance and covariance [112] as

$$\sigma_{S_x}^2 = \langle S_x^2 \rangle - \langle S_x \rangle^2, \quad (3.81)$$

$$\sigma_{S_y}^2 = \langle S_y^2 \rangle - \langle S_y \rangle^2, \quad (3.82)$$

$$\sigma_{S_x, S_y} = \langle S_x S_y \rangle - \langle S_x \rangle \langle S_y \rangle, \quad (3.83)$$

where angle brackets denote the expectation value of the contained expression. Expectation values can be calculated by integrating the product of the expression and the joint probability distribution of the random variables contained in the expression, with respect to each of the random variables [119]. In the case of the normalised

signals, the random variables are the source power P_0 , shot noise coefficients $\gamma_{s,x/y/r}$ and dark noise coefficients $\gamma_{d,x/y/r}$, and all can be assumed to be independent, such that the joint probability distribution is given by the product of the probability distributions of the individual variables [52]. Taking the first quadrature signal as an example,

$$S_x = \frac{G_x \left(P_0 S'_x + \gamma_{s,x} \sqrt{P_0 S'_x} + \gamma_{d,x} \right)}{G_r \left(P_0 \beta_r + \gamma_{s,r} \sqrt{P_0 \beta_r} + \gamma_{d,r} \right)}, \quad (3.84)$$

where

$$S'_x = (\alpha_x + \beta_x \cos(\varphi)), \quad (3.85)$$

and correspondingly

$$S'_y = (\alpha_y + \beta_y \cos(\varphi + \delta)). \quad (3.86)$$

The expectation value of S_x is then given by

$$\langle S_x \rangle = \int S_x f(\gamma_{s,x}) f(\gamma_{s,r}) f(\gamma_{d,x}) f(\gamma_{d,r}) f(P_0) d\gamma_{s,x} d\gamma_{s,r} d\gamma_{d,x} d\gamma_{d,r} dP_0, \quad (3.87)$$

where the integral with respect to each random variable is over the full range that variable may take. Although in the case of the source power only positive values are possible, a Gaussian distribution will be assumed, given by [52]

$$f(x) = \frac{1}{\sqrt{2\pi}\sigma_x} e^{-\frac{1}{2}\left(\frac{x-\mu_x}{\sigma_x}\right)^2}, \quad (3.88)$$

for all variables, limiting this analysis to the case

$$\sigma_{P_0} \ll \mu_{P_0}. \quad (3.89)$$

For the variables $\gamma_{s,x}$ and $\gamma_{d,x}$ this integral may be evaluated analytically, however, for the other terms this is not possible. As noise contributions are likely to be small, a series expansion around the mean value is appropriate, and taking second order Taylor expansions about μ_{P_0} for P_0 , and about zero for $\gamma_{s,r}$ and $\gamma_{d,r}$ and integrating results in

$$\begin{aligned} \langle S_x \rangle = \frac{G_x}{G_r} & \left(\beta_r^3 \mu_{P_0}^5 + \beta_r^2 \mu_{P_0}^2 \sigma_{\gamma_{s,r}}^2 (\mu_{P_0}^2 + \sigma_{P_0}^2) \right. \\ & \left. + \beta_r \mu_{P_0} \sigma_{\gamma_{d,r}}^2 (\mu_{P_0}^2 + 3\sigma_{P_0}^2) + 6\sigma_{\gamma_{d,r}}^2 \sigma_{\gamma_{s,r}}^2 (\mu_{P_0}^2 + 6\sigma_{P_0}^2) \right) \frac{S'_x}{\beta_r^4 \mu_{P_0}^5}. \end{aligned} \quad (3.90)$$

A further simplification may be made by eliminating terms where the combined exponent of the standard deviation terms is greater than two, once again assuming that the noise terms are small, resulting in

$$\langle S_x \rangle = \frac{G_x}{G_r} \frac{\beta_r \mu_{P_0} \left(\beta_r \mu_{P_0} + \sigma_{\gamma_{s,r}}^2 \right) + \sigma_{\gamma_{d,r}}^2}{\beta_r^3 \mu_{P_0}^2} S'_x. \quad (3.91)$$

Applying the same techniques, approximate expressions may also be found for the expectation values of the second quadrature signal along with the squares and product of the quadrature signals, given by

$$\langle S_y \rangle = \frac{G_y}{G_r} \frac{\beta_r \mu_{P_0} \left(\beta_r \mu_{P_0} + \sigma_{\gamma_{s,r}}^2 \right) + \sigma_{\gamma_{d,r}}^2}{\beta_r^3 \mu_{P_0}^2} S'_y, \quad (3.92)$$

$$\langle S_x^2 \rangle = \frac{G_x^2}{G_r^2} \frac{\beta_r^2 \left(\sigma_{\gamma_{d,x}}^2 + \mu_{P_0} S'_x \left(\mu_{P_0} S'_x + \sigma_{\gamma_{s,x}}^2 \right) \right) + 3\beta_r \mu_{P_0} S_x'^2 \sigma_{\gamma_{s,r}}^2 + 3\sigma_{\gamma_{d,r}}^2 S_x'^2}{\beta_r^4 \mu_{P_0}^2}, \quad (3.93)$$

$$\langle S_y^2 \rangle = \frac{G_y^2}{G_r^2} \frac{S_y'^2 \left(\beta_r \mu_{P_0} \left(\beta_r \mu_{P_0} + 3\sigma_{\gamma_{s,r}}^2 \right) + 3\sigma_{\gamma_{d,r}}^2 \right) + S_y' \beta_r^2 \mu_{P_0} \sigma_{\gamma_{s,y}}^2 + \beta_r^2 \sigma_{\gamma_{d,y}}^2}{\beta_r^4 \mu_{P_0}^2}, \quad (3.94)$$

$$\langle S_x S_y \rangle = \frac{G_x G_y}{G_r^2} \frac{\beta_r \mu_{P_0} \left(\beta_r \mu_{P_0} + 3\sigma_{\gamma_{s,r}}^2 \right) + 3\sigma_{\gamma_{d,r}}^2}{\beta_r^4 \mu_{P_0}^2} S'_x S'_y. \quad (3.95)$$

$$(3.96)$$

Substituting these expressions into equations 3.81 to 3.83 and again eliminating combined exponents greater than two yields

$$\sigma_{S_x}^2 = \frac{G_x^2}{G_r^2} \frac{1}{\beta_r^4 \mu_{P_0}^2} \left(\beta_r^2 \sigma_{\gamma_{d,x}}^2 + \beta_r^2 \mu_{P_0} \sigma_{\gamma_{s,x}}^2 S'_x + \left(\beta_r \mu_{P_0} \sigma_{\gamma_{s,r}}^2 + \sigma_{\gamma_{d,r}}^2 \right) S_x'^2 \right), \quad (3.97)$$

$$\sigma_{S_y}^2 = \frac{G_y^2}{G_r^2} \frac{1}{\beta_r^4 \mu_{P_0}^2} \left(\beta_r^2 \sigma_{\gamma_{d,y}}^2 + \beta_r^2 \mu_{P_0} \sigma_{\gamma_{s,y}}^2 S'_y + \left(\beta_r \mu_{P_0} \sigma_{\gamma_{s,r}}^2 + \sigma_{\gamma_{d,r}}^2 \right) S_y'^2 \right), \quad (3.98)$$

$$\sigma_{S_x S_y} = \frac{G_x G_y}{G_r^2} \frac{1}{\beta_r^4 \mu_{P_0}^2} \left(\beta_r \mu_{P_0} \sigma_{\gamma_{s,r}}^2 + \sigma_{\gamma_{d,r}}^2 \right) S'_y S'_x. \quad (3.99)$$

These equations approximate the variances and covariance of the quadrature signals in the case that the noise contributions are small, as should be the case in precision optical interferometry.

Table 3.2: Properties of a Hamamatsu S2386-18K/L photodiode, taken as an example of a typical photodiode, required to calculate the standard deviation of the photocurrent, as given in the manufacturer data sheet [120].

Parameter	Value	Notes
632.8 nm responsivity	0.43 A W ⁻¹	
Dark current	0.6 pA	Estimated at 20 °C.
Shunt resistance	100 GΩ	Typical value.

3.4.3 Monte-Carlo simulation

In order to test the validity of the approximations made in equations 3.97 to 3.99 Monte-Carlo simulations [121] were employed. Monte-Carlo simulations allow random uncertainties to be propagated through the measurement model through repeated random sampling, without the need for analytical uncertainty propagation techniques. Monte-Carlo simulations were carried out with $M = 10000$ Monte-Carlo repeats and $N = 1000$ phase points. $M \times N$ random laser powers P_0 were generated following a Gaussian distribution [52] with a 1 mW mean and 0.01 mW standard deviation, corresponding to 1 % relative intensity noise. Idealised outputs P_x , P_y and P_r were then generated modelling a phase-quadrature interferometer according to a modified version of equations 4.8 and 4.9, a derivation of which is given in appendix D. Modifying these equation to include a lossless intensity reference pick off with a reflectivity β_r results in

$$P_x = (1 - \beta_r) P_0 2\Re\Im(1 + \cos(\varphi)), \quad (3.100)$$

$$P_y = (1 - \beta_r) P_0 (\Re^2 + \Im^2 + 2\Re\Im \cos(\varphi + \delta)), \quad (3.101)$$

$$P_r = \beta_r P_0, \quad (3.102)$$

where $\Re$ and $\Im$ are the power reflectivity and transmissivity of the phase quadrature coating, as explained in detail in section 4.2. The optical signals were calculated taking $\beta_r = 0.1$, $\Re = \Im = 0.3$, and $\delta = 90^\circ$ (corresponding approximately to values found for NPL interferometers [54]), with phase values φ uniformly distributed in the range $[0, 2\pi]$ in $N = 1000$ steps. Voltages were then calculated using equations 3.76 to 3.78 with Gaussian noise added according to equations 3.79 and 3.80 using numerical values from table 3.2 for the photodiode properties, and assuming a 100 kHz bandwidth. Signal gains were set to give each modelled voltage output a maximum value of 1 V, with the maximum calculated across all phase values and Monte-Carlo repeats. S_x and S_y were then calculated according to equations 3.33 and 3.34, and variances and covariances determined from the $M = 10000$ Monte-Carlo repeats. The resulting variances and covariances are presented in figure 3.3, along with analytic solutions calculated using equations 3.79, 3.80 and 3.97 to 3.99. Good agreement can be seen between the Monte-Carlo simulation results and the analytical approximation. More accurate approximations may be found if the higher order terms are not eliminated, however, as these equations are only likely to be used to guide the design of interferometric systems, rather than to produce estimates of the uncertainty in practice, the shortened form is preferred here.

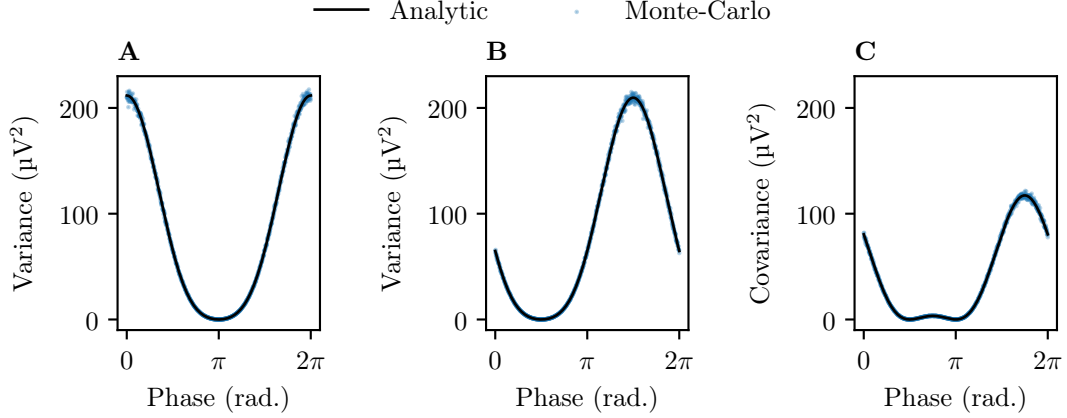


Figure 3.3: Variances (**A** and **B**) and covariance (**C**) of the quadrature signals for shot and dark noise coefficients calculated with equations 3.68 and 3.67 using numerical values from table 3.2 as described in the text. Values have been calculated both with equations 3.97 to 3.99 and with a Monte Carlo approach using equations 3.69 to 3.78 for 10000 repeats, as described in the text.

3.5 Discussion and comparison of limiting factors

This section aims to identify factors that limit the performance of homodyne interferometry, with a particular focus on performance in the nanoscale regime. Limiting factors for displacement measurement, other than the interferometric phase change uncertainty, will first be discussed. Transition points between the dominant uncertainty contributions for short range measurements, where fixed uncertainty contributions dominate, and long range measurements, where displacement dependent uncertainty sources dominate, will be identified. Following this, the factors that limit the uncertainty in the measurement of an interferometric phase change will be considered. A particular focus will be placed on the properties of phase quadrature coating based interferometers, which are the focus of this thesis.

3.5.1 Displacement measurement uncertainty

The variance in an interferometric measurement of displacement is given, according to equation 3.6, by

$$\begin{aligned} \sigma_{\Delta x}^2 = & \left(\frac{(1 + F_\lambda) c}{2\pi N_{\text{pass}} n_1 f} \right)^2 \sigma_{\Delta \varphi}^2 + \left(1 - \frac{n_0}{n_1} \right)^2 \sigma_{x_0}^2 + \left(\frac{x_0}{n_1} \right)^2 \sigma_{n_0}^2 + \left(\frac{n_0 x_0}{n_1 f} \right)^2 \sigma_f^2 \\ & + \left(\frac{\Delta x + x_0}{f} \right)^2 \sigma_f^2 + \left(\frac{\Delta x + x_0}{n_1} \right)^2 \sigma_{n_1}^2 + \left(\frac{n_1 (\Delta x + x_0) - n_0 x_0}{(1 + F_\lambda) n_1} \right)^2 \sigma_{F_\lambda}^2. \quad (3.6) \end{aligned}$$

Table 3.3: Typical values for parameters affecting interferometer uncertainties assuming a thermally stabilised single frequency HeNe operating at 632.8 nm with a beam waist radius of 1 mm. The values given are approximate, and are intended only to be used to give rough estimates of the measurement uncertainty.

Influence parameter	Symbol	Value	Standard uncertainty	Fractional uncertainty
Optical frequency	f	473.6 THz	3 MHz	6×10^{-9}
Refractive index	n	1.00027	1×10^{-8}	1×10^{-8}
Wavefront correction	F_λ	1×10^{-8}	5×10^{-9}	0.5

From equation 3.6 it is apparent that for short range displacements the first four terms will tend to dominate the measurement uncertainty, with the last three displacement dependent terms becoming dominant over larger displacement ranges, although this distinction can be blurred in the case of large dead paths. The point of transition between these groups of uncertainty contributions will be dependant on the exact configuration and operating environment of the interferometer system in question, however, some general conclusions can be drawn. Some typical values for an interferometer used for precision displacement measurements are given in table 3.3, assuming a thermally stabilised HeNe laser operating at approximately 632.8 nm, a beam waist radius of 1 mm and environmental parameters as given in table 3.1. A single refractive index value is given in table 3.3, however, depending on both environmental stability and measurement time, the refractive index will generally vary during measurement. Assuming a temperature change of -0.1°C , an air pressure change of 1 hPa a relative humidity change of -1% and constant CO_2 content, with initial values as given in table 3.1, where signs have been chosen such that the changes act constructively, the refractive index will change by approximately 1×10^{-6} . Substituting these values into equation 3.6, taking the initial refractive index to be unity and F_λ to be zero for simplicity, and assuming a double pass interferometer with $N_{\text{pass}} = 4$, results in the uncertainty contributions given in table 3.4, which, when added in quadrature, result in the total uncertainty in the displacement measurement in this simplified case. The same limits are shown graphically in figure 3.4, with the limits imposed by a number of commonly employed laser stabilisation methods shown, along with the effects of various peak-to-peak non-linearities and dead path uncertainties.

For the conditions assumed here, it can be seen from table 3.4 and figure 3.4 that in the large displacement limit the refractive index uncertainty will dominate for measurements made in air where stabilised HeNe lasers are employed. As a result, even when the refractive index can be determined up to the accuracy of the improved Edlén equations [97], which may not always be the case, the frequency stability of polarisation stabilised HeNe lasers is sufficient for interferometry in air, and the reduced frequency uncertainty of iodine stabilisation is not required. For interferometers operating in the vacuum, greater frequency stability would be beneficial. A further implication is that the large fractional uncertainty associated with

Table 3.4: Approximate uncertainty contributions to the measured displacement assuming an initial refractive index of unity, a change in refractive index during the measurement of 1×10^{-6} , a wavefront correction factor of zero, and other parameters as defined in table 3.3.

Uncertainty source	Symbol	Uncertainty contribution
Interferometric phase	$\Delta\varphi$	$(3 \times 10^{-8}) \sigma_{\Delta\varphi}$
Dead path	x_0	$(1 \times 10^{-6}) \sigma_{x_0}$
Refractive index	n	$(1 \times 10^{-8}) \sqrt{(\Delta x + x_0)^2 + x_0^2}$
Optical frequency	f	$(6 \times 10^{-9}) \sqrt{(\Delta x + x_0)^2 + x_0^2}$
Wavefront correction	F_λ	$(5 \times 10^{-9}) \Delta x$

the wavefront correction factor, resulting from the simplified assumption of Gaussian beam interference, is not a limiting factor, and for measurements in air, more advanced wavefront corrections are also not required.

In the small displacement limit, uncertainties may be dominated by phase uncertainty, dead path uncertainty, or again by the refractive index uncertainty. Taking a typical value of the interferometric phase uncertainty to be 4×10^{-4} rad, equivalent to a displacement of approximately 10 pm, the phase uncertainty will dominate if the uncertainty in the dead path is below approximately 12 μm , and the dead path is below approximately 1 mm. Larger changes in refractive index (corresponding to larger changes in environmental parameters) during the measurement would require a lower uncertainty in the dead path, and where measurements are made in unstable environments, or over longer time scales, this may become a limiting factor, even if the refractive index is measured accurately.

From table 3.4 it can also be seen that, if the previous restriction on the dead path uncertainty is met, and assuming zero dead path, a 10 pm equivalent phase uncertainty will dominate over the refractive index uncertainty for displacements of less than 1.2 mm, a condition that is often fulfilled in dimensional nanometrology, for example when characterising a nanopositioning stage [8], or carrying out AFM measurements.

3.5.2 Phase measurement uncertainty

Phase measurement uncertainties will be introduced both by noise and non-linearities in homodyne interferometers. The fundamental limit in the phase uncertainty for a given bandwidth and optical power is given by the shot noise, as described in section 3.4, and this limit will be considered first. For ease of comparison, uncertainties in the change in interferometric phase will be converted to displacements in this section, assuming zero dead path, unity refractive index, a wavefront correction factor of zero, a wavelength of exactly 632.8 nm, and a double pass interferometer

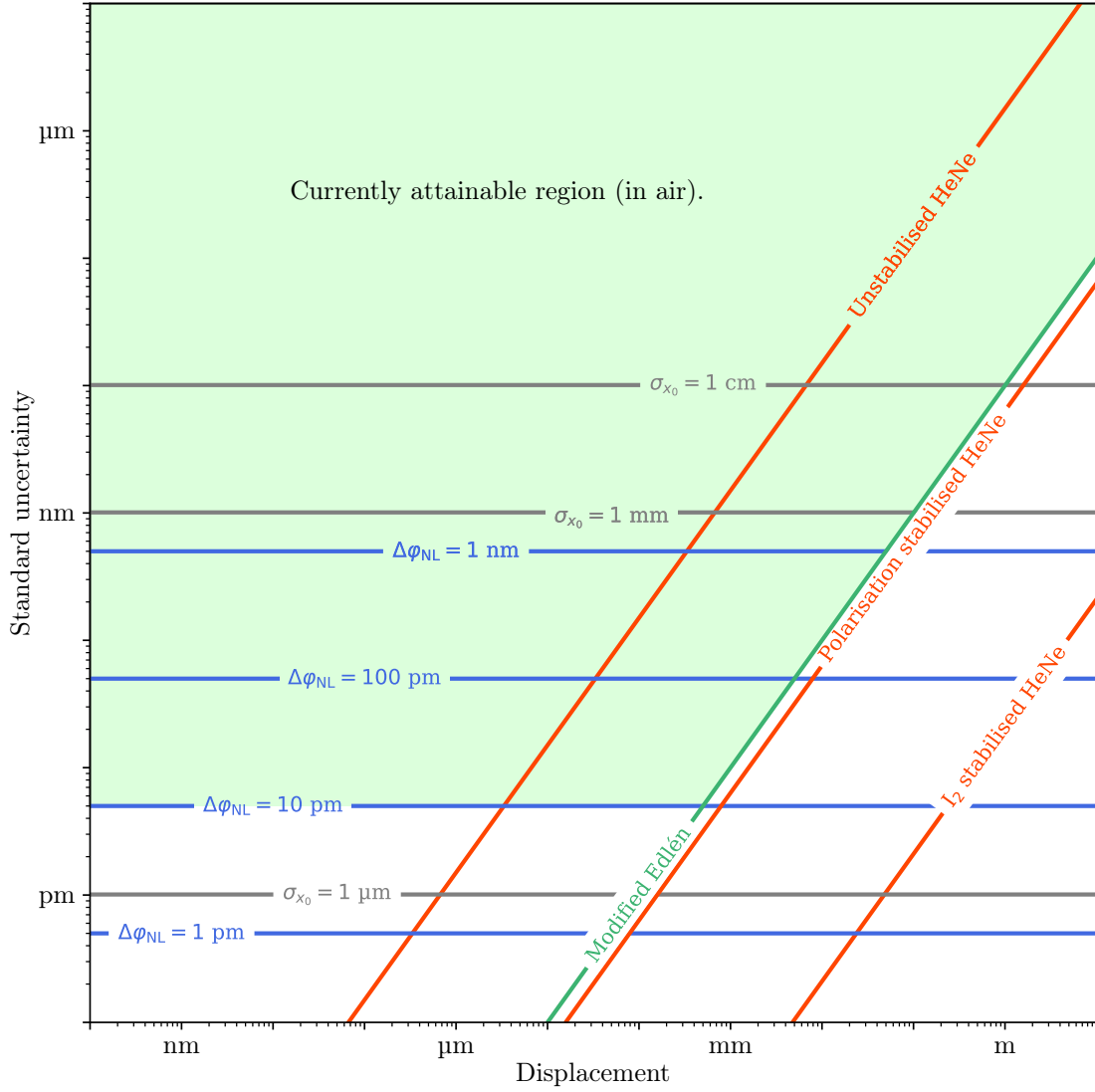


Figure 3.4: Graphical illustration of uncertainty limits in homodyne interferometry. The limits imposed by a number of laser frequency stabilisation methods, dead path standard uncertainties (σ_{x_0}) and peak-to-peak non-linearities ($\Delta\varphi_{NL}$) are shown, along with the refractive index limit imposed for measurements in air by the uncertainty in the improved Edlén equations. Uncertainties are calculated based upon the parameter values given in table 3.3, with a change in refractive index of 1×10^{-6} over the course of the measurement, and zero dead path.

configuration with $N_{\text{pass}} = 4$, using the following equation:

$$\sigma_x = \frac{632.8 \times 10^{-9}}{8\pi} \sigma_{\Delta\varphi}. \quad (3.103)$$

Taking the shot noise limited current uncertainty, given by equation 3.68, and propagating this through the quadrature and reference channels using the analytical

expressions derived in sections 3.4 and 3.3 allows a lower limit to be placed on displacement measurement uncertainties with a given homodyne interferometer design. This thesis focusses on phase-quadrature coating based interferometers, a topic that will be discussed in detail in later chapters, and as such the phase uncertainties in a typical phase-quadrature interferometer will, as an example, be considered here.

The output powers of a phase-quadrature based interferometer are given by equations 3.100 to 3.102. The uncertainty in the displacement arising from an interferometer following these equations may be calculated both with the analytic expressions presented in section 3.3, and with an extension of the Monte-Carlo simulation process described in section 3.4.3. In order to validate the analytic expressions, a Monte-Carlo simulation was once again carried out following the procedure described in section 3.4.3 to produce simulated normalised quadrature signals S_x and S_y with $M = 10000$ Monte-Carlo repeats and $N = 1000$ phase steps. A Heydemann correction was then applied to each Monte-Carlo repeat separately, fitting equation 3.16. The instantaneous interferometric phase was calculated with equation 3.32, and a fringe counting algorithm as described in section 3.3 applied. Analytic uncertainties were also calculated by applying equations 3.36 to 3.39 to the analytic uncertainties in the normalised quadrature signals calculated in section 3.4.3. The results of this process are shown in figure 3.5 **A**. The Monte-Carlo process described in section 3.4.3 was then modified to simulate interferometers with $\Re = \Im = 0.3$ and $\delta = 110^\circ$, and $\Re = 0.4$, $\Im = 0.2$ and $\delta = 90^\circ$, and the process repeated with these non-ideal parameters. The modified results are shown in figure 3.5 **B** and **C** respectively. In all cases the uncertainty plotted is for a displacement starting at an instantaneous phase of zero, with S_x at a maxima, and S_y at midscale for the zero quadrature phase error case.

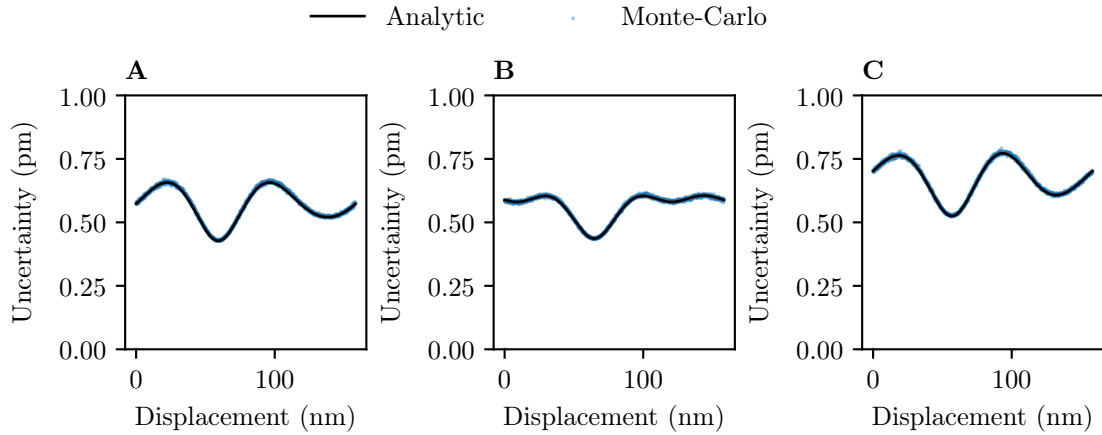


Figure 3.5: Displacement uncertainties in an idealised shot noise limited phase-quadrature interferometer calculated with both a Monte-Carlo approach as described in the text, and analytical expressions, for three different conditions. **A** shows results from an interferometer with $\Re = \Im = 0.3$ and $\delta = 90^\circ$. **B** shows results from an interferometer with $\Re = \Im = 0.3$ and $\delta = 110^\circ$. **C** shows results from an interferometer with $\Re = 0.4$, $\Im = 0.2$ and $\delta = 90^\circ$.

Good agreement between the Monte-Carlo simulation and analytic expressions can be seen for all results presented in figure 3.5. This result validates both the analytical expressions given in section 3.3, and the statement that where large numbers of points are included in the ellipse fit, the fit coefficient uncertainty is negligible [110]. The analytical expressions derived in section 3.4 are also further validated.

Following the same conditions as for figure 3.5 **A**, but varying the measurement bandwidth and illuminating optical powers allows the fundamental performance limits of homodyne interferometry to be explored. Figure 3.6 shows the variation in displacement uncertainty (evaluated at the maximum within a fringe) for a range of bandwidths and illuminating powers, with the 1 pm contour indicated.

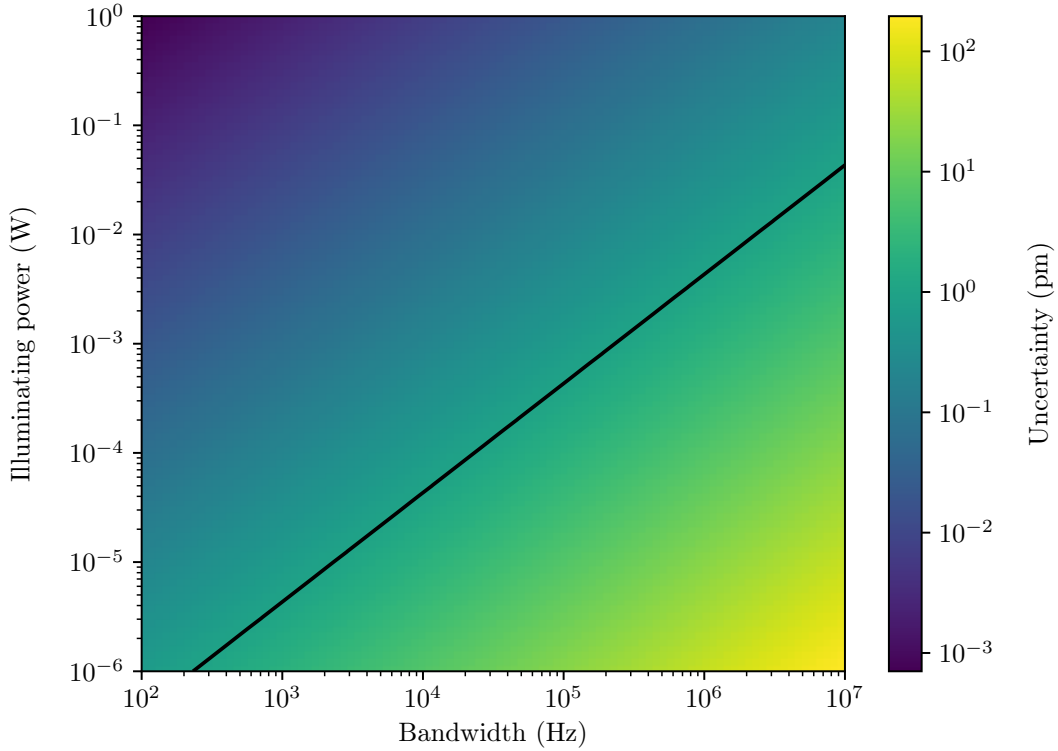


Figure 3.6: Variation of the maximum displacement uncertainty with measurement bandwidth and illuminating optical power for an interferometer with parameters as described in the text. The 1 pm contour line is indicated.

Whilst the 10% intensity reference pick off reflectivity used in 3.5 has been commonly employed in NPL interferometers, the analytical expressions can be used to identify an optimal value in a shot noise limited interferometer with $\Re = \Im = 0.3$ and $\delta = 90^\circ$, as shown in figure 3.7. A minima occurs at $\beta_r = 0.33$, although for small deviations around this point the increase in uncertainty is relatively minor. This minima arises as noise in the reference channel is coupled into the normalised sine and cosine channels by the normalisation process, and as such small reference

intensities (which will have higher noise levels relative to the signal magnitude) result in higher noise levels in the phase measurement. Similarly, very weak sine and cosine signals caused by an excessively large reference pick off reflectivity will have higher noise levels, relative to the signal amplitudes. The precise location of this minima will be dependant on the noise distribution in an individual interferometer system, however, in general equalising the average optical powers at the reference, sine and cosine channels will give good performance.

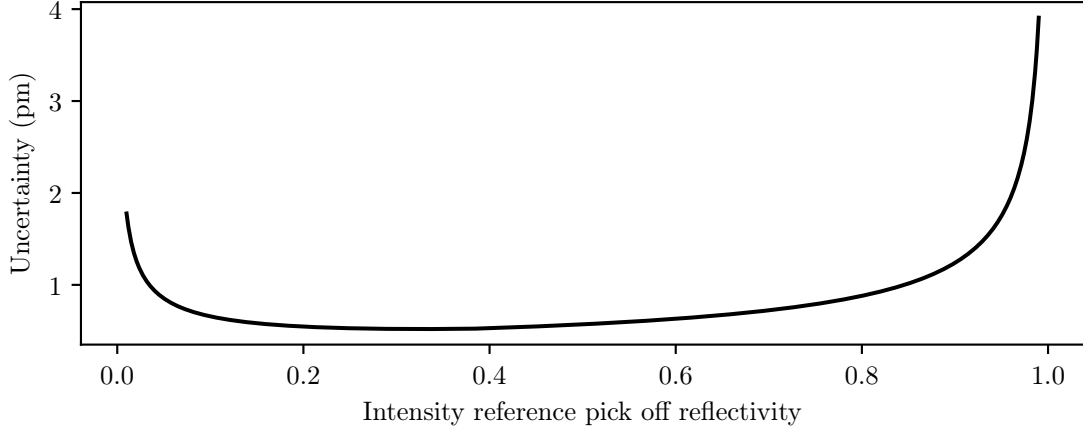


Figure 3.7: Variation of the maximum displacement uncertainty with the intensity reference pick off reflectivity β_r under conditions as described in the text

The analytical expressions can also be used to test the effects of varying optical parameters on the interferometer performance, where throughout this subsection it will be assumed that the non-linearities introduced by any non-ideal optical parameters have been fully corrected for. The effects of varying the quadrature phase for an interferometer with $\beta_r = 0.1$ and other parameters as described for figure 3.7 are shown in figure 3.8, where an idealised ellipse fitting style correction has again been applied using the known ellipse parameters. In this case, a minima occurs at a quadrature phase shift of approximately 70° , a result that may be surprising as a 90° quadrature phase shift is generally considered ideal. This deviation of the uncertainty minima from 90° can be explained as a result of the scaling effects of the ellipse fitting correction; at some points on the ellipse the correction stretches the phase spread of any intensity noise distribution, whereas at other points it is compressed. This effect is demonstrated for an extreme example in figure 3.9. In figure 3.9 **A** the phase-distribution for each data point can be seen to be approximately equal for all data points, whereas in 3.9 **C** the phase distribution can be seen to vary depending on the location on the Lissajous figure. The sharp corner seen at the minima of figure 3.8 can also be explained as a result of the stretching and compressing effects of the ellipse fitting correction; the minima arises as a result of the intersection between the phase uncertainties at the compressed and stretched points on the Lissajous figure. This effect will be highly dependent on noise distribution, but in general small increases in noise level may be achieved with quadrature phase

shifts other than 90° . The uncertainty scaling effects of the quadrature phase shift can however be seen in figure 3.8 to be relatively minor compared to the non-linearities present in even high performance interferometer designs [11], indicating that the exact value of the quadrature phase shift is not a critical parameter in interferometer design.

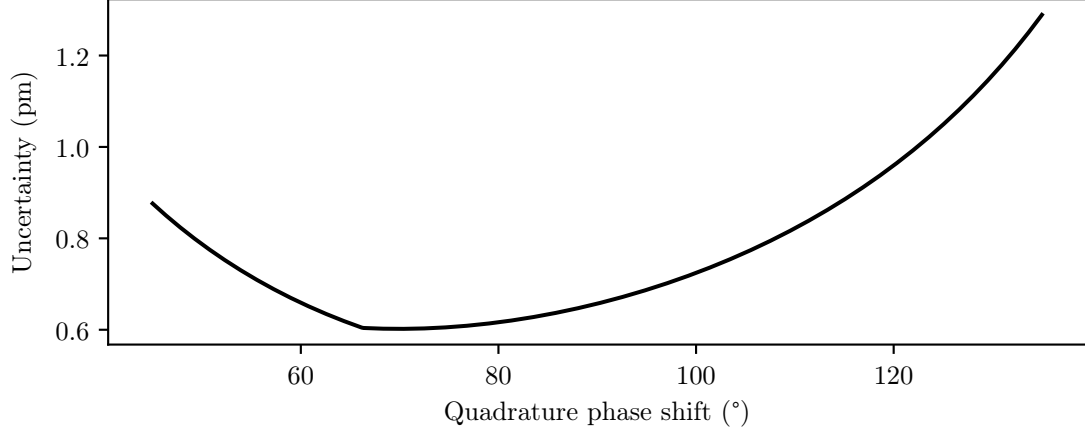


Figure 3.8: Variation of the maximum displacement uncertainty with quadrature phase shift δ under conditions as described in the text.

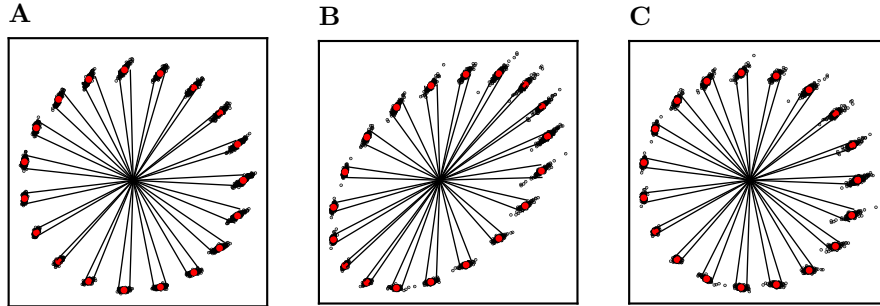


Figure 3.9: Demonstration of the phase distribution scaling effects of ellipse fitting corrections - radial lines indicate 95 % confidence limits for the phase distribution of each data point. **A** shows data with a 90° quadrature phase shift. **B** shows data with identical intensity noise to **A**, but a 50° phase shift. **C** shows the result of applying an ellipse fitting correction to **B**.

Similarly, the effect of varying reflectivity to transmission ratio of the quadrature coating on the measurement uncertainty can be investigated. The results of varying the reflectivity to transmission ratio for a constant absorption of 0.4, and reflectivities varying from 0.1 to 0.5 are shown in figure 3.10. In this case the minima occurs at a reflectivity to transmission ratio of unity, and this remains the case for any noise

distribution likely to be encountered in a real device. Exceptions may occur where significant noise is introduced into one channel electronically, however this situation should not be encountered in practical measurements.

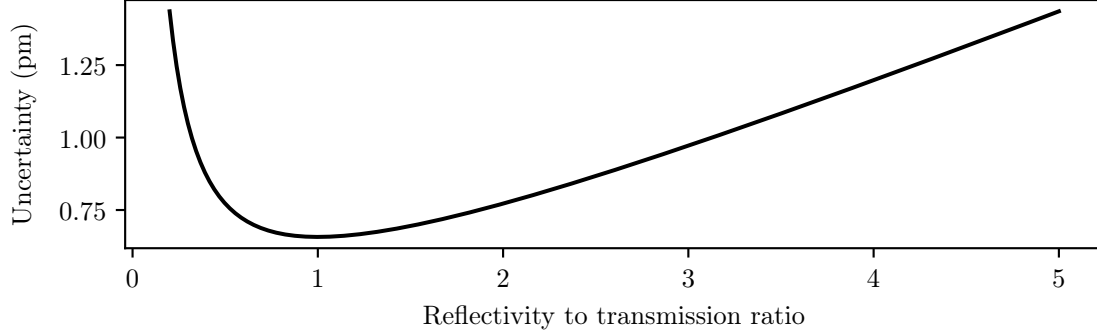


Figure 3.10: Variation of the maximum displacement uncertainty with reflectivity to transmission ratio under conditions as described in the text.

The effects of non-linearities that cannot be corrected for are more challenging to make general statements about than noise sources affecting the phase measurement, however, from the review of existing interferometer designs presented in chapter 2, it is apparent that uncorrected non-linearities will generally have magnitudes greater than 10 pm unless extreme care is taken to eliminate them. The effects of uncorrected non-linearities must be evaluated on a system by system basis when constructing an uncertainty budget for a given optical interferometer, as described in section 3.3. A general conclusion that can be drawn, however, is that considering non-linearities of the orders of tens of picometres as typical for a precision interferometer, non-linearities will typically dominate over phase uncertainties introduced by noise in the quadrature signals if the shot noise limit is approached, unless only very low optical powers are available, or very high bandwidths are required (see figure 3.6 for limiting cases).

3.6 Summary

In this chapter, general uncertainty budgets have been presented both for displacement measurements made with homodyne interferometry, and for the measurement of a change in the interferometric phase. These uncertainty budgets are intended to give a clear overview of potential uncertainty sources in homodyne interferometry, rather than to be directly applicable to a specific interferometer system. In addition to the uncertainty budgets, in order to provide a performance limit to target when considering other uncertainty sources, the shot noise limit in homodyne interferometry has been explored. This limit has then been used to investigate the effects of various key parameters in phase-quadrature coating based homodyne interferometry, although the results are equally applicable to other homodyne interferometer results.

Limiting factors in homodyne interferometry have been identified, and are shown graphically in figure 3.4, and are tabulated under typical conditions in table 3.4. For short range measurements, either non-linearity or dead path uncertainty have been shown to limit the displacement measurement uncertainty, with a dead path uncertainty of less than $12\text{ }\mu\text{m}$ required to reach a displacement measurement uncertainty of 10 pm , assuming zero dead path, and a change in the refractive index of 1×10^{-6} during the measurement. In the long range case, for measurements made in air with stabilised HeNe lasers, the refractive index uncertainty has been shown to be the dominant uncertainty contribution, with the displacement uncertainty contributed by the refractive index uncertainty reaching 10 pm for a displacement of 1.2 mm , assuming zero dead path.

Whilst this chapter has primarily reviewed uncertainty sources already discussed in the literature, some new results have been presented. A key conclusion of this chapter, demonstrated in figure 3.2, is that higher order non-linearities cannot be corrected for based upon the quadrature signals alone, as it is possible to produce a circular Lissajous figure with a non-linear phase change. This result has not previously been demonstrated in the literature, and has significant implications, both for the correction, and for the measurement of non-linearities in homodyne interferometers. Non-linearity estimation based upon the amplitude deviations of Lissajous figure formed by the quadrature signals from the ideal circular form has been proposed in the literature [50]. The work presented here has demonstrated that this technique may underestimate non-linearity where higher order effects are present. As higher order effects are almost exclusively responsible for the non-linearities that remain after applying a Heydemann correction, the likelihood of underestimating residual non-linearities with the amplitude deviation technique is extremely high, and this technique should therefore not be employed. When considering possible non-linearity correction approaches, this finding has the significant implication that more information is required to characterise and correct for higher order effects, with this conclusion driving much of the work presented in chapter 6.

Another novel result is the fact that a 90° quadrature phase shift may not be ideal (assuming an ellipse fitting correction is applied), as demonstrated in figure 3.8. The mechanism responsible for this effect has also been demonstrated in figure 3.9. Whilst the 90° quadrature phase shift commonly employed in homodyne interferometry has been shown to be non-optimal, with the optimal value dependent on the noise distribution in the quadrature signals, a more practical conclusion of section 3.5.2 is that the exact values of the quadrature phase shift and reflectivity to transmission ratio in a phase-quadrature coating based interferometer are not critical to the interferometer performance, as non-linearities are likely to dominate the phase-measurement uncertainty relative to the noise scaling effects of non-ideal quadrature phase shift and reflectivity to transmission ratio values. Although demonstrated in this chapter for an idealised non-linearity correction using known parameters, this conclusion is further reinforced in section 6.2.1, where the effects of varying reflectivity to transmission ratio and quadrature phase shift on the performance of various ellipse fitting algorithms are investigated, and found to be in agreement with the results presented here.

Chapter 4

Phase-quadrature coatings for homodyne interferometry

4.1 Introduction

Two primary methods have emerged for introducing the required 90° quadrature phase shift in homodyne interferometry. The first and more commonly employed technique involves the use of polarising optics as detailed in section 2.2.1. The second, and the focus of this chapter, makes use of necessarily lossy (as demonstrated in section 4.2) thin film optical coatings designed such that suitable phase changes are introduced upon reflection from and transmission through the phase-quadrature coating. This approach was first demonstrated in 1953 [24] with single layer aluminium coatings, and has been further refined for plate [122] and cube [123] beam splitters with a single layer metallic coating, and for plate beam splitters with multi-layer metallic coatings [25]. The fabrication of multi-layer metallic coatings for cube beam splitters, however, has not previously been described in the literature.

Interferometers employing phase-quadrature coatings are designed such that at each of the two quadrature outputs the interfering beams from the measurement and reference arms of the interferometer have undertaken different numbers of reflections and transmissions through the phase-quadrature coating [24]. The principle of operation of a phase-quadrature coating based interferometer is demonstrated in figure 4.1, with the total phase change at outputs a and b being given by

$$\varphi_a = (\varphi_t + \varphi_r) - (\varphi_r + \varphi_t) \quad (4.1)$$

$$= 0 \quad (4.2)$$

$$\varphi_b = (\varphi_t + \varphi_t) - (\varphi_r + \varphi_r) \quad (4.3)$$

$$= 2(\varphi_t - \varphi_r), \quad (4.4)$$

where φ_r and φ_t are the phase changes upon reflection from (figure 4.1 **B**) and transmission through (figure 4.1 **A**) the phase-quadrature coating respectively, and a symmetrical coating design has been assumed. From these equations it can be seen that by designing the phase-quadrature coating to produce appropriate values

of φ_r and φ_t the quadrature condition can be achieved, that is

$$2(\varphi_t - \varphi_r) = 90^\circ. \quad (4.5)$$

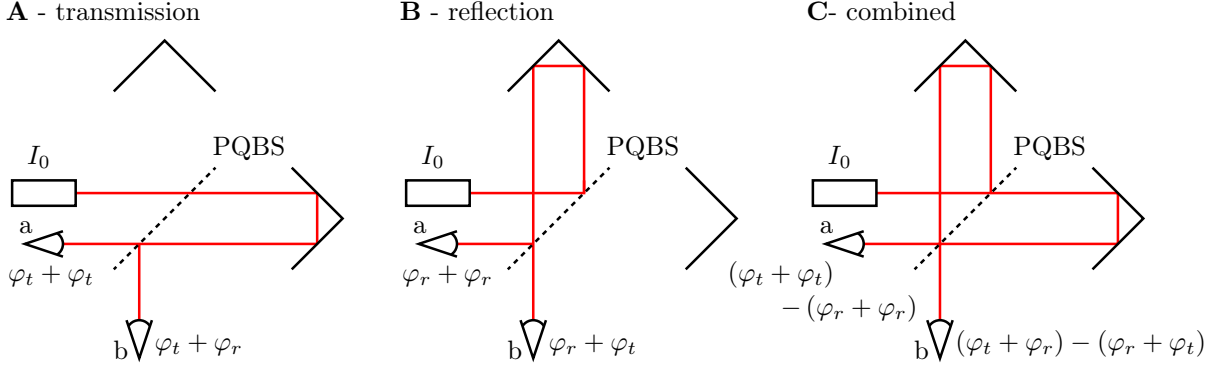


Figure 4.1: Demonstration of the principle of operation of phase-quadrature coatings. I_0 represents optical source intensity, PQBS the phase-quadrature beam splitter. φ_r and φ_t are the phase changes introduced by reflection from and transmission through the phase-quadrature coating respectively, and the resulting phase differences at the outputs a and b are indicated on the schematic.

The values of the phase changes φ_r and φ_t are determined by the complex argument of the amplitude reflectivity and transmissivity coefficients of the phase-quadrature coating, given via the four quadrant arctangent function by

$$\varphi_r = \text{atan2}(\text{Im}(\mathbf{r}), \text{Re}(\mathbf{r})), \quad (4.6)$$

$$\varphi_t = \text{atan2}(\text{Im}(\mathbf{t}), \text{Re}(\mathbf{t})), \quad (4.7)$$

and are therefore coupled to the intensity reflectivity and transmissivity coefficients given by the complex magnitude of $\mathbf{r}$ and $\mathbf{t}$ (the amplitude reflectivity and transmissivity coefficients). As a unity intensity reflectivity to transmissivity ratio is desirable for phase-quadrature coatings employed in interferometry, this places a further constraint on the design of such a coating. In addition to this, the phase coefficients exhibit dependencies on both the polarisation and angle of incidence of the illuminating light.

The following sections detail the theoretical limits of phase-quadrature coatings, along with a design methodology for achieving the desired 90° phase shift and unity intensity reflectivity to transmissivity ratio. Practical considerations and techniques required to fabricate such coatings are then discussed, along with methods of verifying the properties of the fabricated coating.

4.2 Conservation of energy considerations

Considering the restrictions placed upon the outputs of quadrature homodyne interferometers by the principle of conservation of energy allows some general limits to

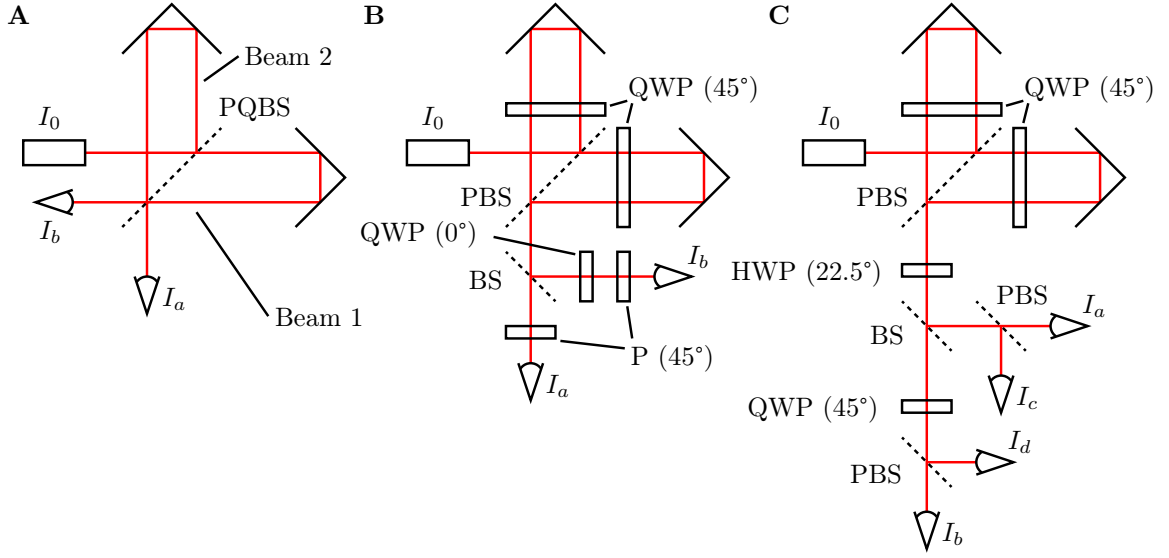


Figure 4.2: Schematic of three possible corner cube retro-reflector based Michelson interferometers employing phase-quadrature coatings (**A**), two output polarisation based quadrature detection (**B**) and four output polarisation based quadrature detection (**C**). I_0 , input intensity; PQBS, phase-quadrature beam splitter; PBS, polarising beamsplitter; BS, non-polarising beamsplitter; QWP, quarter waveplate; HWP, half waveplate; P, polariser; I_a , I_b , I_c and I_d output intensities.

be placed upon the performance of any phase-quadrature coating, without a more detailed assessment of the coating design being made. Although the general forms of the outputs of a quadrature homodyne interferometer are well known (see, for example, [21]), physical quantities are often absorbed into constants of proportionality, clouding any analysis of conservation of energy. As such, a re-derivation of the outputs of a quadrature homodyne interferometer is given in appendix D, from first principles.

Three possible morphologies for Michelson interferometers employing corner cube retro-reflectors are presented in figure 4.2: a phase-quadrature interferometer (**A**), along with two (**B**) and four (**C**) output polarisation based homodyne interferometers for comparison. In order to apply conservation of energy principles, the outputs I_a and I_b , and in the case of the four output interferometer I_c and I_d , must be found in terms of the input intensity I_0 .

Taking first the phase-quadrature coating based interferometer (figure 4.2 **A**), the outputs, as derived in appendix D, are given by

$$I_a = 2\Re\Im I_0(1 + \cos(\varphi)), \quad (4.8)$$

$$I_b = I_0(\Re^2 + \Im^2 + 2\Re\Im \cos(\varphi + \delta)), \quad (4.9)$$

where $\Re$ and $\Im$ are the power reflectivity and transmission coefficients of the phase quadrature coating, φ is the interferometric phase, and δ is the quadrature phase shift introduced by the phase-quadrature coating.

For an ideal quadrature interferometer δ is 90° , and the reflectivity to transmissivity ratio is unity, giving outputs that vary as sine and cosine with changing interferometric phase φ , with the maximum possible signal amplitudes for a given input intensity

$$I_a^{\text{ideal}} = 2\Re^2 I_0(1 + \cos(\varphi)), \quad (4.10)$$

$$I_b^{\text{ideal}} = 2\Re^2 I_0(1 + \sin(\varphi)). \quad (4.11)$$

For the two output polarisation based quadrature interferometer (figure 4.2 **B**), the outputs will be of a similar form. With ideal alignment of the polarisation axes of the various components the effective value of R and T will be

$$\Re = \Im = \sqrt{\frac{1}{8}}, \quad (4.12)$$

accounting for the 50 % intensity loss taking place at the 45° orientated output polarisers. The quarter waveplate in the path of the I_b output introduces the quadrature phase shift, resulting in the ideal case in

$$I_a^{\text{ideal}} = \frac{1}{4} I_0(1 + \cos(\varphi)), \quad (4.13)$$

$$I_b^{\text{ideal}} = \frac{1}{4} I_0(1 + \sin(\varphi)). \quad (4.14)$$

Applying conservation of energy considerations to a two output quadrature interferometer, and accounting for losses on transmission through the phase-quadrature beamsplitter coating or output polarisers, the sum of the outputs must be less than the input intensity

$$I_a + I_b \leq I_0. \quad (4.15)$$

Taking the sum of the output intensities to be

$$I_{\text{total}} = I_a + I_b \quad (4.16)$$

this total intensity is given by

$$I_{\text{total}} = \left((\Re + \Im)^2 + 2\Re\Im (\cos(\varphi) + \cos(\varphi + \delta)) \right) I_0, \quad (4.17)$$

for the phase-quadrature interferometer. Extrema of this total will occur when

$$\frac{dI_{\text{total}}}{d\varphi} = 0, \quad (4.18)$$

$$\Rightarrow \sin(\varphi) = -\sin(\varphi + \delta). \quad (4.19)$$

The quadrature phase δ can be restricted to the range $-\pi$ to π with no loss of generality, and applying this restriction leads to values of φ at the extrema of I_{total} of

$$\varphi = \pm \frac{\delta}{2} + 2\pi n \quad (4.20)$$

for integers n . As the intensity is periodic in φ with period 2π , the value of n is arbitrary. Maxima will occur where the previous extrema condition is met, and

$$\frac{d^2 I_{\text{total}}}{d\varphi^2} < 0. \quad (4.21)$$

Maxima in the sum of the output intensities therefore occur at

$$\varphi = 2\pi n - \frac{\delta}{2}. \quad (4.22)$$

Evaluating the sum of the output intensities at the maximal point yields

$$I_{\text{total}}|_{\text{max}} = \left((\Re + \Im)^2 + 4\Re\Im \cos\left(\frac{\delta}{2}\right) \right) I_0, \quad (4.23)$$

and therefore the conservation of energy restriction implies

$$(\Re + \Im)^2 + 4\Re\Im \cos\left(\frac{\delta}{2}\right) \leq 1, \quad (4.24)$$

$$\cos\left(\frac{\delta}{2}\right) \leq \frac{1 - (\Re + \Im)^2}{4\Re\Im}. \quad (4.25)$$

For an interferometer with a unity reflectivity to transmissivity ratio this restricts the maximum value of the intensity reflectivity (and therefore transmissivity) coefficient to the range

$$0 < \Re \leq \frac{1}{2} \left(1 + \cos\left(\frac{\delta}{2}\right) \right)^{-\frac{1}{2}}, \quad (4.26)$$

or evaluated numerically at quadrature

$$0 < \Re \lesssim 0.38. \quad (4.27)$$

In the four output case (figure 4.2 **C**) two pairs of outputs are produced, with the 50 % losses at the output polarisers of the two output instead redirected to additional detectors. An additional phase difference Δ will be present between the pairs of outputs, in the ideal case of 180° , resulting in

$$I_a^{\text{ideal}} = \frac{1}{4} I_0 (1 + \cos(\varphi)), \quad (4.28)$$

$$I_b^{\text{ideal}} = \frac{1}{4} I_0 (1 + \sin(\varphi)), \quad (4.29)$$

$$I_c^{\text{ideal}} = \frac{1}{4} I_0 (1 - \cos(\varphi)), \quad (4.30)$$

$$I_d^{\text{ideal}} = \frac{1}{4} I_0 (1 - \sin(\varphi)). \quad (4.31)$$

Imposing conservation of energy considerations results in

$$I_a + I_b + I_c + I_d \leq I_0. \quad (4.32)$$

Taking the sum of the output intensities to be

$$I_{\text{total},4} = I_a + I_b + I_c + I_d, \quad (4.33)$$

yields, for an interferometer with an effective reflectivity of R and a unity effective R to T ratio,

$$I_{\text{total},4} = 2\Re^2 I_0 (4 + \cos(\varphi) + \cos(\varphi + \delta) + \cos(\varphi + \Delta) + \cos(\varphi + \delta + \Delta)). \quad (4.34)$$

Due to the lossless nature of the four output design, at least in the case of idealised optics as considered here, $I_{T,4}$ is constrained by conservation of energy to be equal to the input intensity. Evaluating this equality at the quadrature point ($\delta = \pi/2$) for an ideal effective reflectivity of

$$\Re = \sqrt{\frac{1}{8}} \quad (4.35)$$

results in

$$I_{\text{total}} = \frac{1}{4} I_0 (4 + \cos(\varphi) + \sin(\varphi) + \cos(\varphi + \Delta) + \sin(\varphi + \Delta)), \quad (4.36)$$

a condition that is only satisfied for all values of φ by a 180° secondary phase difference Δ . Selecting this value for Δ permits any value of $\Re$ satisfying the restriction

$$\Re \leq \sqrt{\frac{1}{8}} \quad (4.37)$$

to satisfy conservation of energy requirements.

An important consequence of the limit imposed by equation 4.26 is that the commonly employed two output polarisation based quadrature interferometer design (figure 4.2 **B**) does not maximise the intensity outputs of the interferometer due to the 50% losses imposed by the 45° orientation of the output polarisers. This limitation is removed by employing Polarising Beam Splitter (PBS) rather than polarisers at the outputs, and utilising both the reflected and transmitted output beams, as is done for the four output interferometer (figure 4.2 **C**), however, this is at the cost of introducing a significant number of additional optical components. In principle, it may be possible to design a phase-quadrature coating that does reach the limits imposed by conservation of energy, and therefore exceeds the output intensities of a polarisation based two output quadrature interferometer, resulting in an increase in signal to noise ratio. An additional implication of equation 4.26 is that a phase-quadrature coating exhibiting an ideal unity $\Re$ to $\Im$ ratio must always be absorbing.

The effects of the increased losses in the phase-quadrature coating based and two output polarisation based interferometer designs may be investigated by applying

Table 4.1: Displacement uncertainty due to shot noise contributions for the interferometer designs described in section 4.2 assuming an illuminating power of 1 mW at 633 nm with idealised unity quantum efficiency photodiodes and a 1 MHz measurement bandwidth.

Interferometer design	Displacement uncertainty
Two output polarising	± 3.30 pm
Four output polarising	± 2.33 pm
Ideal phase-quadrature ($\Re = 0.38$)	± 3.05 pm
Existing phase-quadrature ($\Re = 0.29$)	± 4.02 pm

the equations derived in chapter 3. The output powers of the phase-quadrature interferometer may be modelled with equations 4.8 and 4.9, the two output polarising interferometer with equations 4.13 and 4.14 and the four output polarising interferometer with equations 4.28 to 4.31. Applying equation 3.68 to each output with an illuminating power of 1 mW, a wavelength of 633 nm, idealised unity quantum efficiency photodiodes and a 1 MHz measurement, followed by applying equation 3.44 directly to the resulting signals and uncertainties results in the maximum displacement uncertainties given in table 4.1. The maximum uncertainties have been calculated by multiplying the maximum uncertainty, within a fringe, in the instantaneous phase by $\sqrt{2}$, equivalent to a displacement from the highest noise location on a fringe, to the same location on the next fringe, therefore giving the worst case value.

4.3 Thin film modelling

In order to design a phase-quadrature coating a modelling approach capable of handling the effects of absorbing optical materials must be employed. Transfer matrix methods are widely utilised in thin film modelling [20,124–129], and a transfer matrix modelling methodology based upon that of Yeh [130] was employed in this work. The goal of the transfer matrix modelling approach is to construct a transfer matrix that links the electric field on one side of a multi-layer thin film stack to the other. This is achieved by assuming the electric field at any point in the multi-layer stack may be described as the sum of forward and backward propagating plane waves at that point, where the backward propagating component is a result of back reflections of the forward propagating wave from the layer interfaces, as shown in figure 4.3 for a stack consisting of N layers. As transfer matrix modelling is a well understood technique the details of the implementation used in this work are given in appendix E, with the only major differences from the implementation described in [130] being the explicit treatment of complex refractive indices that arise in absorbing media, without the use of non-physical complex valued angles, and the extension to a 4×4 matrix to allow simultaneous modelling of orthogonal polarisation components.

Following the derivation given in appendix E a global transfer matrix $\mathbf{T}$ from

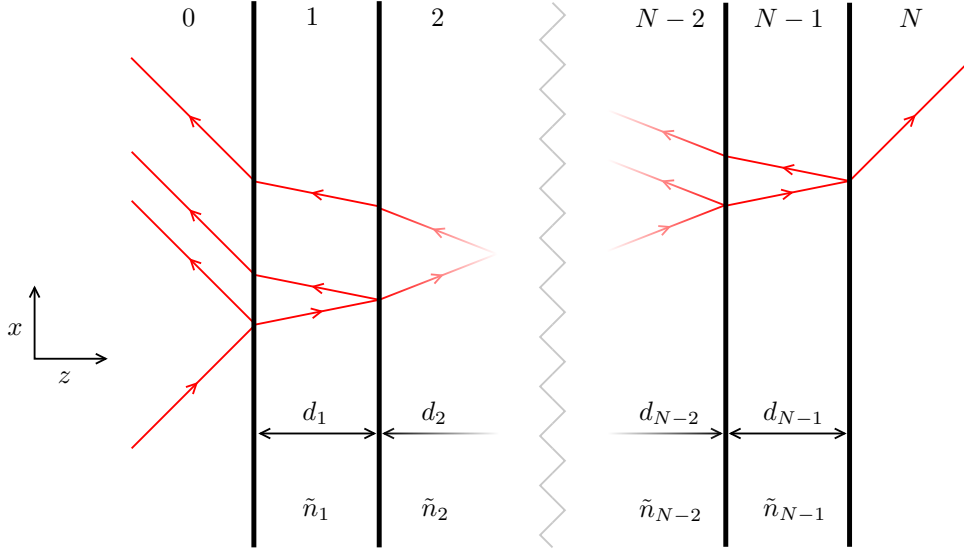


Figure 4.3: Geometry of the thin film transfer matrix modelling methodology. Light is incident from the left hand side, with material layers m labelled 0 to N , layer thicknesses d_m and complex refractive indices $\tilde{n}_m$. For clarity, only first order reflections are shown.

layer N to layer 0 may be defined in terms of the individual transfer matrices for each layer m , $\mathbf{P}_m$, and propagation matrices between layers m and $m+1$, $\mathbf{T}_{m,m+1}$, as

$$\begin{pmatrix} \tilde{E}_F^{(S)}(0) \\ \tilde{E}_B^{(S)}(0) \\ \tilde{E}_F^{(P)}(0) \\ \tilde{E}_B^{(P)}(0) \end{pmatrix} = \mathbf{T}_{0,1} \left(\prod_{m=1}^{N-1} \mathbf{P}_m \mathbf{T}_{m,m+1} \right) \begin{pmatrix} \tilde{E}_F^{(S)}(N) \\ \tilde{E}_B^{(S)}(N) \\ \tilde{E}_F^{(P)}(N) \\ \tilde{E}_B^{(P)}(N) \end{pmatrix}, \quad (4.38)$$

$$= \mathbf{T} \begin{pmatrix} \tilde{E}_F^{(S)}(N) \\ \tilde{E}_B^{(S)}(N) \\ \tilde{E}_F^{(P)}(N) \\ \tilde{E}_B^{(P)}(N) \end{pmatrix}, \quad (4.39)$$

where the complex electric field component values in the zeroth and N^{th} layers, $\tilde{E}(0)$ and $\tilde{E}(N)$ respectively, have polarisations indicated by superscripts and forward or backward directions indicated by subscripts.

Whilst in some cases obtaining the global transfer matrix is the goal, as may be the case if the multi-layer stack is to be included as part of a larger model (as is the case for the modelling technique presented in chapter 5), more commonly reflectivity and transmission coefficients of the stack are desired. Amplitude reflectivity and

transmission coefficients defined for light incident in the positive and negative z directions are given by

$$\begin{aligned}
\mathbf{r}_+^{(S)} &= \frac{\mathbf{T}_{10}}{\mathbf{T}_{00}}, & \mathbf{r}_+^{(P)} &= \frac{\mathbf{T}_{32}}{\mathbf{T}_{22}}, \\
\mathbf{t}_+^{(S)} &= \frac{1}{\mathbf{T}_{00}}, & \mathbf{t}_+^{(P)} &= \frac{1}{\mathbf{T}_{22}}, \\
\mathbf{r}_-^{(S)} &= -\frac{\mathbf{T}_{10}}{\mathbf{T}_{00}}, & \mathbf{r}_-^{(P)} &= -\frac{\mathbf{T}_{32}}{\mathbf{T}_{22}}, \\
\mathbf{t}_-^{(S)} &= \mathbf{T}_{11} - \frac{\mathbf{T}_{10}\mathbf{T}_{01}}{\mathbf{T}_{00}}, & \mathbf{t}_-^{(P)} &= \mathbf{T}_{33} - \frac{\mathbf{T}_{32}\mathbf{T}_{23}}{\mathbf{T}_{22}},
\end{aligned} \tag{4.40}$$

where $\mathbf{T}_{ij}$ is the matrix element in the i^{th} row and j^{th} column. The corresponding intensity coefficients are then given by

$$\mathfrak{R}_+^{(S/P)} = \left| \mathbf{r}_+^{(S/P)} \right|^2, \tag{4.41}$$

$$\mathfrak{T}_+^{(S/P)} = \text{Re} \left(\frac{k_{z,N}}{k_{z,0}} \right) \left| \mathbf{t}_+^{(S/P)} \right|^2, \tag{4.42}$$

$$\mathfrak{R}_-^{(S/P)} = \left| \mathbf{r}_-^{(S/P)} \right|^2, \tag{4.43}$$

$$\mathfrak{T}_-^{(S/P)} = \text{Re} \left(\frac{k_{z,0}}{k_{z,N}} \right) \left| \mathbf{t}_-^{(S/P)} \right|^2, \tag{4.44}$$

where in the case that layer N is absorbing, the power transmission values are valid only for intensities defined on the high z side of the $N - 1$ to N interface. Finally, the phase changes upon reflection and transmission are given [20] by

$$\varphi_{r\pm}^{(S/P)} = \text{Arg} \left(\mathbf{r}_{\pm}^{(S/P)} \right), \tag{4.45}$$

$$\varphi_{t\pm}^{(S/P)} = \text{Arg} \left(\mathbf{t}_{\pm}^{(S/P)} \right). \tag{4.46}$$

These forward and backward reflectivity, transmission and phase coefficients may then be applied, as described in the following section, to design the desired coating.

4.4 Phase-quadrature coating design

Previous work at the NPL has identified a three layer chromium-gold-chromium coating [25] as a suitable phase-quadrature coating design, capable of attaining a unity reflectivity to transmission ratio and a quadrature phase difference whilst keeping absorption losses acceptably low and being resistant to oxidation of the coating materials. This existing design consists of a 16 nm layer of gold between two 4 nm layers of chromium, and is optimised for use with plate beam splitters at 632.8 nm.

Table 4.2: Literature values of refractive index of materials used in phase-quadrature coatings at 632.8 nm.

Material	Real component (n)	Imaginary component (κ)	Source
Chromium (Cr)	3.1395	3.3152	[131]
Gold (Au)	0.183 44	3.4332	[132]
Schott N-BK7	1.5151	0.0000	[133]

In order to utilise a phase-quadrature coating in a more flexible cube beam splitter design, with glass on both sides of the coating, this design must be modified. The response of a chromium-gold-chromium coating fully encased by glass to varying layer thickness is shown in figure 4.4, modelled with the previously described transfer matrix technique. Literature values of the refractive index have been used for each material, and are given in table 4.2. Intersections between the lines of ideal quadrature angle and reflectivity to transmission ratio occur at the chromium and gold layer thicknesses presented in table 4.3. From these values, it can be seen that the ideal losses are lower in the S polarised case.

The sensitivity of these ideal coating designs to changes in angle of incidence are shown in figure 4.5. From figure 4.5 it can be seen that for S polarised light the quadrature response is insensitive to changes in angle (figure 4.5 A), whereas the variation in reflectivity to transmission ratio with respect to angle of incidence (figure 4.5 B) is greater than for P polarised light. The sensitivity of the designs to changes in the complex refractive index of the metallic layers is shown in figure 4.6, with minimal difference in the magnitude of the sensitivity for the S and P polarised designs.

Sensitivities to all modelled coating parameters are summarised in figure 4.7, with similar sensitivities found in most cases for S and P polarisations. A significant difference between polarisations can be seen in the sensitivity to non-uniform refractive indices of the incident and transmitted media. As glass-coating-glass phase-quadrature beam splitters are deposited onto one glass surface, and cemented to another, this may present a problem in the S polarised case if the refractive index of the optical cement varies significantly from that of the glass. Due to the lower absorption of the S polarised coating design, leading to a higher signal to noise ratio, and generally comparable sensitivities to coating parameters, the S polarised design was selected for fabrication.

The literature values used for the metallic materials were gathered from thermally evaporated films with thicknesses of 40.3 nm and 50.9 nm for chromium [131] and 30.4 nm and 37.5 nm for gold [132], combined with additional measurements made with opaque films, and aim to give a bulk value for the refractive index. Due to the extremely thin nature of the metallic films required for phase-quadrature coatings, with thicknesses of the order of a few atomic lattice spacings, some deviation from the bulk values may be expected. In particular, island formation in thin gold films has been reported in the literature [134, 135], and literature values for the

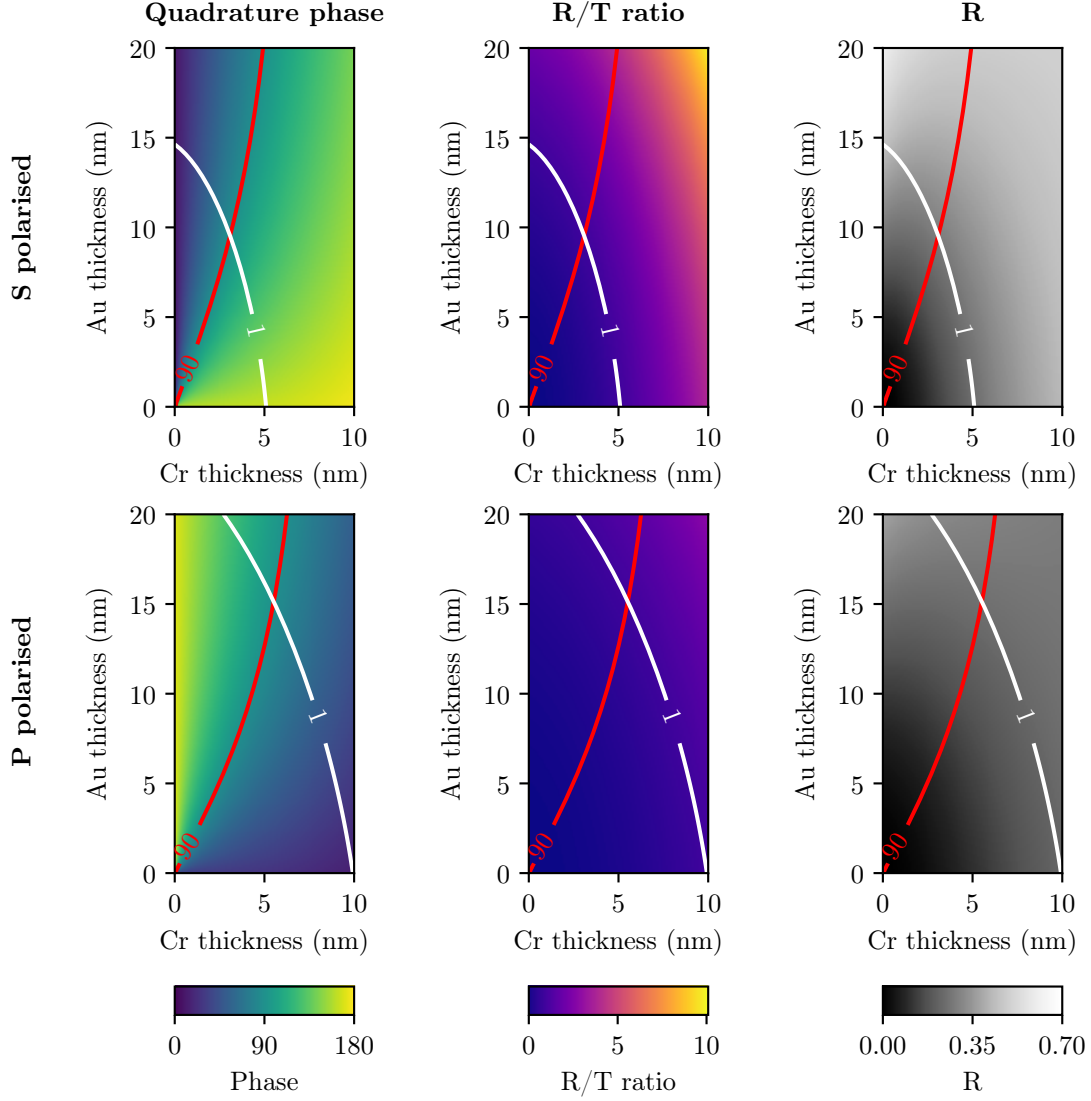


Figure 4.4: The effects of varying layer thicknesses on symmetrical chromium-gold-chromium phase-quadrature coating properties. Marked contours show lines of ideal quadrature angle (red) and reflectivity to transmission ratio (white). A N-BK7 medium is assumed on both sides of the coating, with a 45° angle of incidence.

Table 4.3: Ideal thicknesses of chromium and gold layers and resulting absorption coefficients in a phase-quadrature coating operating at 632.8 nm, using literature values of refractive index, as calculated with the previously described thin film modelling technique.

Polarisation	Cr thickness (nm)	Au thickness (nm)	Absorption
S	3.09	9.54	0.42
P	5.52	15.11	0.48

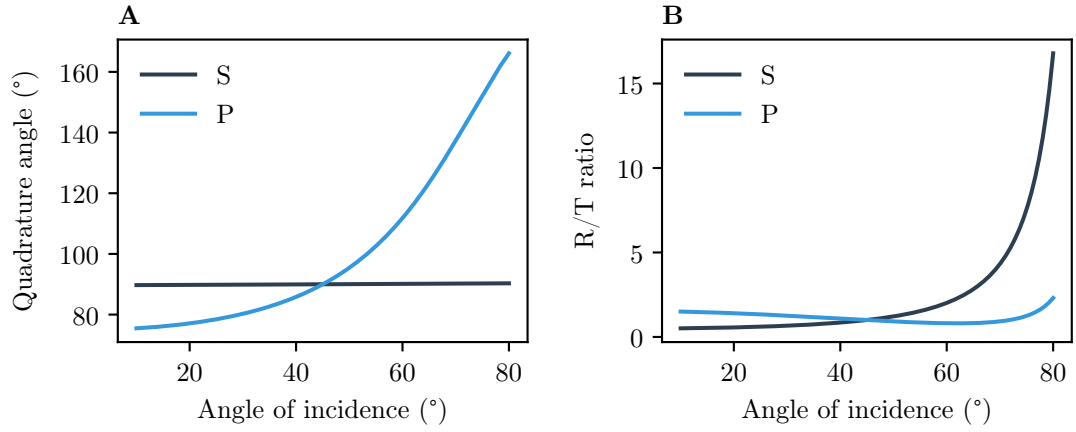


Figure 4.5: Angular sensitivity of quadrature angle (A) and reflectivity to transmission ratio (B) for an ideal chromium-gold-chromium phase-quadrature coating for both S and P polarisations.

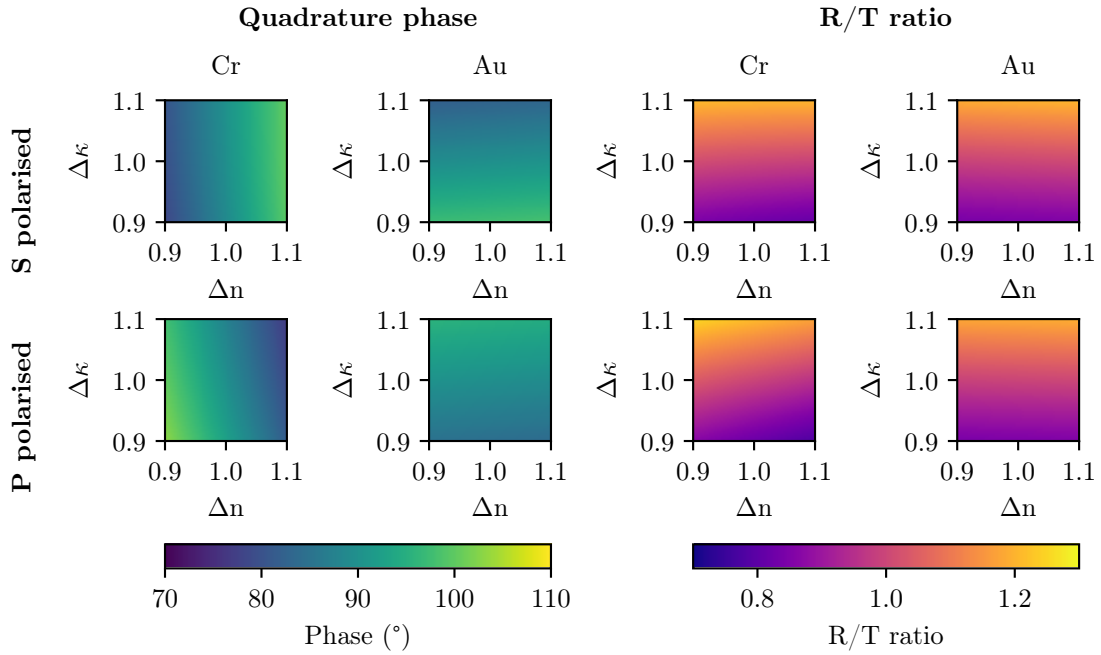


Figure 4.6: The effects of varying complex refractive index on symmetrical chromium-gold-chromium phase-quadrature coating properties. A N-BK7 medium is assumed on both sides of the coating, with a 45° angle of incidence. Changes in refractive index are given fractionally, relative to the literature values presented in table 4.2

refractive index of thin gold films [126, 136, 137], shown in figure 4.8, display an increase in the real component of the refractive index, and decrease in the imaginary

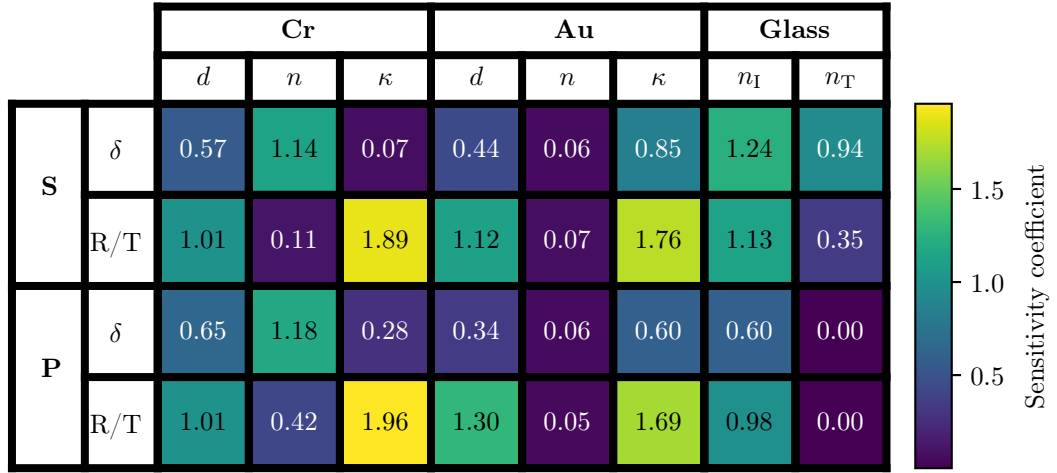


Figure 4.7: Sensitivity coefficients of phase-quadrature coating properties with respect to varying layer thickness, complex refractive index, and incident and transmitted media refractive index. Sensitivities are expressed as the magnitude of the percentage change in the parameter from the ideal (rows) resulting from a 1 % change in relevant variable (columns). Coatings following the designs given in table 4.3 are assumed. Sensitivities have been calculated from partial derivatives approximated with a forward finite difference method. Colour shading is provided as a visual aid, with the colour scale running linearly across the range of values present.

component, as would be expected for a discontinuous gold film in a non-metallic medium. As reported refractive indices for thin gold films are dependant on the substrate, the presence of any adhesion layers, and the deposition conditions, the literature value for bulk gold was used as a starting point, with the expectation that further refinement of the design would be required experimentally.

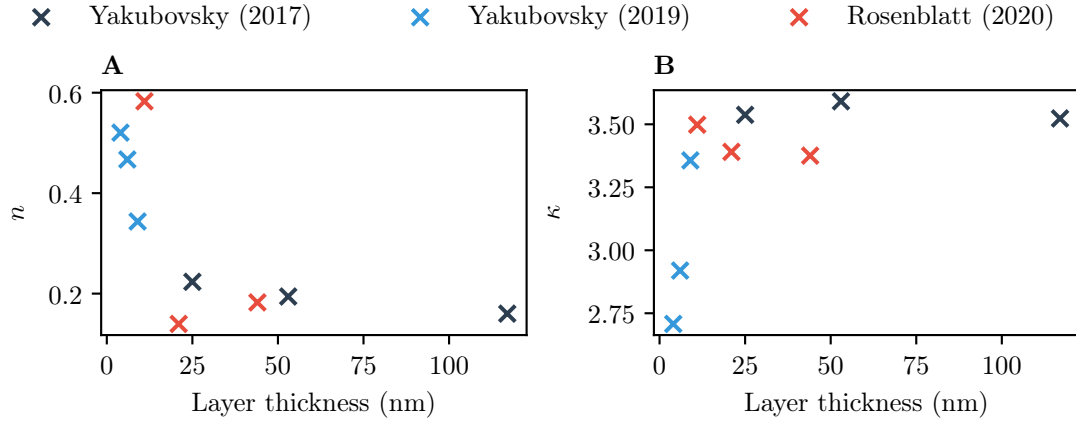


Figure 4.8: Literature values of the real (A) and imaginary (B) components of the complex refractive index of thin gold films as reported by Yakubovsky [126,136] and Rosenblatt [137].

4.5 Coating fabrication

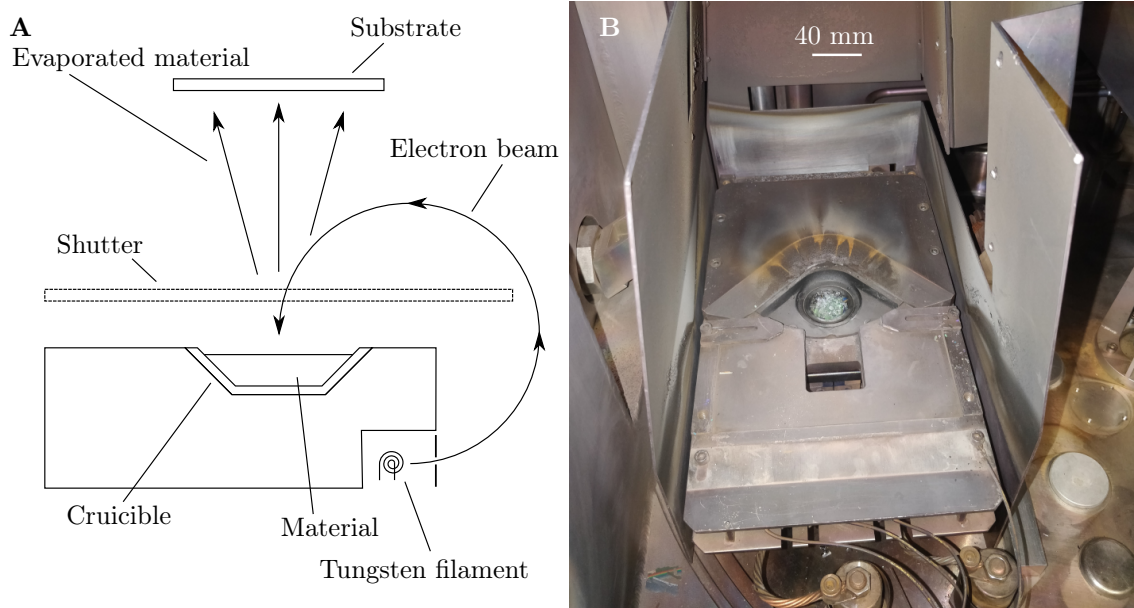


Figure 4.9: Schematic demonstrating the principle of operation of an E-beam gun for thermal evaporation (A) and photograph of the Telemark Industries Model 271 E-beam gun used in this work (B).

Thin metallic films may be produced by a variety of methods falling into two broad categories, chemical and physical deposition [138]. Chemical deposition methods may again be broken down into processes taking place in the liquid and gaseous

phases, and are generally defined as deposition processes in which the primary mechanism of film growth is chemical or electrochemical in nature. Physical methods involve the deposition of a vapour onto the surface to be coated by mechanical means, and as such typically take place under vacuum. The most commonly encountered physical deposition techniques for optical coatings are thermal and sputtering based techniques, with thermal techniques involving evaporation of the source material which then condenses onto the substrate, whereas sputtering relies upon the ejection of particles of the coating material from a solid mass by energetic plasma or gas.

Many coating techniques discussed in the literature [138,139] are suitable for the deposition of nanometre scale layers of chromium and gold, however, due to availability of equipment, thermal evaporation, specifically E-beam deposition, was selected for the fabrication of phase-quadrature coatings. E-beam deposition employs an electron beam to heat the material to be evaporated, with the evaporated material being deposited on the substrate via a line of sight process, as outlined in figure 4.9. A Telemark Industries Model 271 E-beam gun installed in a Nordiko 5050 vacuum coating plant was used to produce the coatings used in this work. This E-beam gun employs a magnetic field transverse to the plane of propagation of the electron beam in figure 4.9 to steer the beam through 270°, preventing contamination of the evaporation material or substrate by filament material. Additional electromagnets act to scan the beam over the material surface, allowing uniform heating over the crucible area. Four crucible pockets allow for deposition of four different materials in a multi-layer design without breaking vacuum.

The deposition process was monitored by a INFICON Quartz Crystal Monitor (QCM) and controlled by a Telemark model 861 thin film deposition controller. This system monitors the change in resonant frequency of the QCM crystal [127] and calculates an estimate of the deposited layer thickness from user programmed acoustic properties of the deposition material. The current supplied to the E-beam gun was regulated by the deposition controller based on a PID loop driven by feedback from the QCM, to control the rate of deposition, along with use of the mechanical shutter shown in figure 4.9 to start and stop the deposition process. This quartz crystal monitor system allowed the deposition process to be controlled to a nominal layer thickness resolution of 0.1 nm. Each layer (chromium, gold, chromium) was deposited sequentially without vacuum being broken.

An oil free scroll pump (Scrollvac SC 60 D) was used for rough evacuation of the vacuum chamber, with a CTI Cryo-Torr 8 cryogenic pump used for pumping the chamber to high vacuum. The coating plant used has been in operation at the NPL since 1991, and there is therefore significant build up of coating material on the chamber walls, limiting pump down speed and ultimate vacuum pressure. Ultimately the chamber pressure was found to reach approximately 1×10^{-7} Torr (approximately 0.01 mPa) after 48 hours of pumping with the cryo pump.

Deposition of phase-quadrature coatings took place at room temperature, and at a target chamber pressure of 2×10^{-5} Torr (approximately 2.67 mPa). For testing purposes, coatings were deposited on 25 mm N-BK7 90° prisms. The substrate surfaces were cleaned with water and detergent followed by acetone and isopropyl alcohol. As a final cleaning step a radio frequency glow discharge was produced in

the vacuum chamber during the pump down process at a pressure of 5×10^{-2} Torr (approximately 6.67 Pa) for 2 minutes. For initial optimisation of the coating design coated prisms were oiled to a second prism with index matched microscope immersion oil, forming a cube beam splitter. This configuration allowed for evaluation of the coating properties in a Michelson interferometer configuration as discussed in the following section. In order to minimise the number of prisms required, the phase-quadrature coatings were removed from the prism surface with chrome etch prior to reuse following the previously described cleaning process. Once the optimisation process was complete, cube beam splitters that were to be used for practical measurements were cemented together with Norland number 65 Ultraviolet (UV) cured optical cement, with a cured refractive index of 1.52 and good adhesion to both glass and metal surfaces.

4.6 Coating optimisation and evaluation

Due to the previously discussed uncertainty in the refractive index of the thin metallic films used in the phase-quadrature coating design, along with the expected non-planar nature of the gold films, the phase-quadrature coating design was optimised experimentally. In order to optimise the optical properties of the coating, namely reflectivity, transmission and induced quadrature phase shift, these properties must first be measured.

4.6.1 Quadrature phase measurement

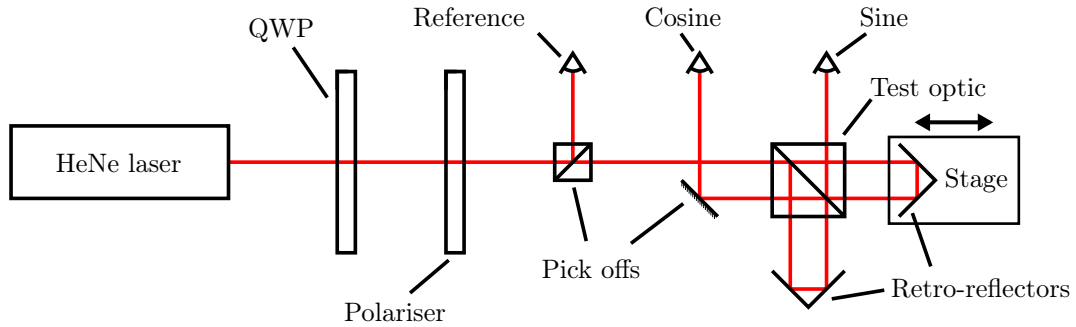


Figure 4.10: Schematic representation of the Michelson interferometer used for quadrature phase angle measurement of phase-quadrature coatings. The moving corner cube retro-reflector in the measurement arm is indicated with arrows. Laser output and QWP aligned to produce circular polarisation at the polariser, allowing rotation of the input polarisation state with minimal intensity fluctuations.

The quadrature phase shift introduced by the phase-quadrature coating was measured by constructing a corner cube based Michelson interferometer, a schematic of which is shown in figure 4.10, with the assembled phase-quadrature beam splitter

cube to be tested. Silvered corner cube retro-reflectors were employed in both interferometer arms in order to minimise the change in polarisation upon reflection. A calibrated Queensgate X-15A nano-positioning stage was used to scan the measurement arm of the interferometer through multiple fringes with nanometric precision. Polarisation was controlled by rotating a Glan-Thompson polariser in the beam with a ThorLabs K10CR1/M rotation stage.

The phase difference between the two interferometer outputs was calculated through two separate techniques. Firstly, an ellipse was fitted to the Lissajous figure formed by the interferometer outputs using a least squares technique [87]. In the ideal case of a 90° quadrature phase shift the Lissajous will take the form of an ellipse with the major and minor axes of the ellipse aligned with the x and y axes of the plot. Any deviation of the quadrature phase will result in distortion of this ellipse, and from the coefficients of the fitted ellipse such deviations can be calculated [87]. Secondly, a Fast Fourier Transform (FFT) was taken of both the sine and cosine outputs with respect to displacement, and the phase difference between the frequency bins corresponding to a period of one half of the illumination wavelength calculated. The Fourier approach has the advantage that higher order harmonics introduced by unwanted reflections from the phase-quadrature beamsplitter surfaces do not effect the measured phase difference, however, this technique was found to be more sensitive to drift in the optical path length difference of the interferometer during a measurement, and was therefore used only as a check on the ellipse fitting technique. Where the two techniques were in disagreement the measurement was repeated, with care taken to eliminate unwanted back reflections where possible.

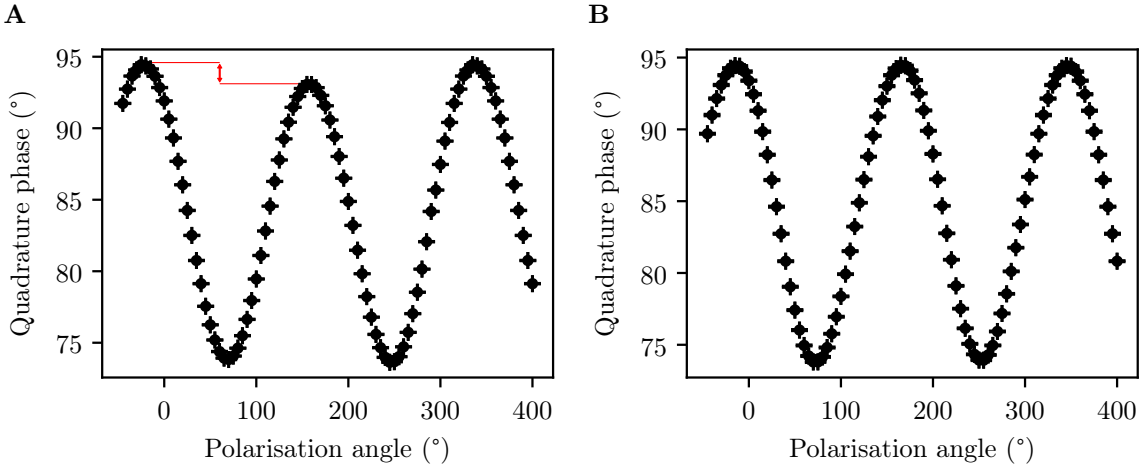


Figure 4.11: Examples of a quadrature phase measurements with (A) and without (B) the unexpected difference between polariser angles separated by 180° , as discussed in the text.

Initially imperfections in the silver coated solid glass corner cube retro-reflectors were found to introduce unexpected differences in the measured quadrature phase shift for polariser rotation angles separated by 180° . An example of this effect is shown in 4.11. A slight wedge in the rotating polariser caused the input beam to the interferometer to precess as the polariser rotated, moving the interfering

beams over the detector area. The 5" manufacturer specified tolerance (Edmund optics #48-606) on the parallelism of the beam reflected from the retro-reflector introduced up to 10" of misalignment between the interfering beams at the detectors, calculated to cause up to approximately 0.23 of an optical fringe to appear across the 3 mm ($1/e^2$ diameter) beam. The combination of the beams moving over the detector area with this angular misalignment resulted in an apparent change in the quadrature phase if the detectors were not identically located with respect to the beam axes. This problem was solved by employing larger area detectors, and using lenses to focus the beams onto the centre of the detector area. Whilst a relatively trivial problem to solve for the interferometer considered here, this effect does have some important implications for interferometers that are required to measure at higher speeds, requiring small area detectors, or for plane mirror interferometers where larger angular deviations of the returned beam may occur. Changes in the quadrature phase that occur within an optical fringe cannot be corrected for by ellipse fitting Heydemann style corrections, and may therefore present a significant uncertainty contribution if this effect is not considered in the interferometer design. The effect may be mitigated by focusing the beams onto the detectors, either with separate lenses as employed here, or using 'bug-eye' detectors with integrated lenses. Where angular deviations in the return beams are expected, minimising the lens focal length, whilst ensuring the lens aperture remains sufficiently large to fully capture the beams, will minimise the movement of the beams in the focal plane, and will therefore also minimise any apparent changes in quadrature phase shift.

When making measurements of the quadrature phase shift both displacement and angular scans through the full 360° range of input polarisation angles were repeated and averaged in order to reduce the effects of noise on the measurement. In order to extrapolate the quadrature phase for the S and P polarised states, sinusoids were fitted to the measured data. Due to the previously described difference between measurements at polariser angles separated by 180° , four separate fits were made, each localised to a range of $\pm 25^\circ$ around the observed extrema in the quadrature phase verses polariser angle data. Averages of the two fitted maxima and minima were taken as the S and P polarised quadrature phase respectively. An expanded uncertainty in the quadrature phase at the extrema was calculated by taking the standard error in the mean of the fitted maxima and minima, and multiplying by 12.71 (derived from the student t distribution with one degree of freedom [89]) to give a 95 % level of confidence. The student t distribution was employed in this case as only two data points were available, to account for the uncertainty in the deviation of the sample mean from the population mean [89]. Whilst uncertainties in the fitted maxima and minima could also be calculated, these values give uncertainties in the fit coefficients of the best fit of a sinusoid to the measured data, and are therefore poor indicators of the uncertainty in the quadrature phase in this case as systematic errors due to the beam precession are present, and the data is not perfectly sinusoidal in nature.

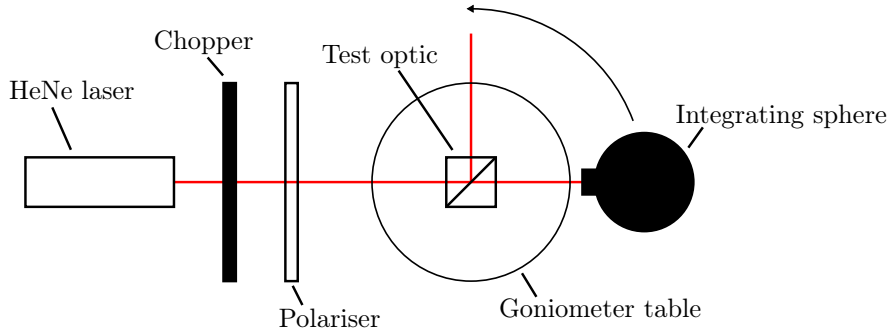


Figure 4.12: Schematic representation of the bench top reflectometer used for reflectivity and transmission measurements of phase-quadrature coatings.

4.6.2 Reflectivity and transmission measurement

Measurement of the reflectivity and transmission properties of the coatings was achieved with the reflectometer shown in figure 4.12, utilising a mechanical chopper and lock-in amplifier to minimise the effects of background lighting on the measurements. This reflectometer had been constructed at NPL prior to this PhD. Reflectivity and transmissivity were measured at 632.8 nm using HeNe laser illumination for both S and P polarisation states, with the polarisation set by rotating a polymer sheet polariser in the incident beam. The polarisation of the incident beam was orientated at approximately 45° to the S and P axes of the test optic. In this application, the goniometer table was used purely to position the integrating sphere in the transmitted and reflected beam paths. An iris at the entrance to the integrating sphere was used to centre the sphere entrance aperture on the relevant beam in an attempt to minimise any effects introduced by the beams taking different paths through the integrating sphere.

In order to measure the absolute reflectivity and transmissivity of the phase-quadrature beam splitter cube, measurements were normalised to the measured laser intensity with no cube in the beam path. Reflectivity and transmission measurements were repeated a minimum of three times, with laser intensity measurements taken before each repeat. Means and standard errors in the means were calculated, along with the corresponding expanded uncertainties. Expanded uncertainties were calculated at the 95 % confidence level, again using the student t distribution with the appropriate number of degrees of freedom [89] due to the relatively small number of repeats.

4.6.3 Optimisation

As can be seen from figure 4.4, for changes in layer thickness on the scales plotted both the quadrature phase and reflectivity to transmission ratios vary smoothly. As a result, the dependence of the quadrature phase and reflectivity to transmission ratio on small changes in layer thickness may be expected to be approximately

linear. This approximation was exploited to optimise the performance of phase-quadrature coatings by repeating coatings with slightly varying layer thicknesses, and fitting planes to the resulting quadrature phases and reflectivity to transmission ratios. Ideal values of the layer thicknesses were then calculated based upon the point at which the fitted planes intercepted the planes of ideal phase-quadrature and reflectivity to transmission ratio. Additional coatings were fabricated at these calculated layer thicknesses, and the process repeated with the measured properties of the new coating included.

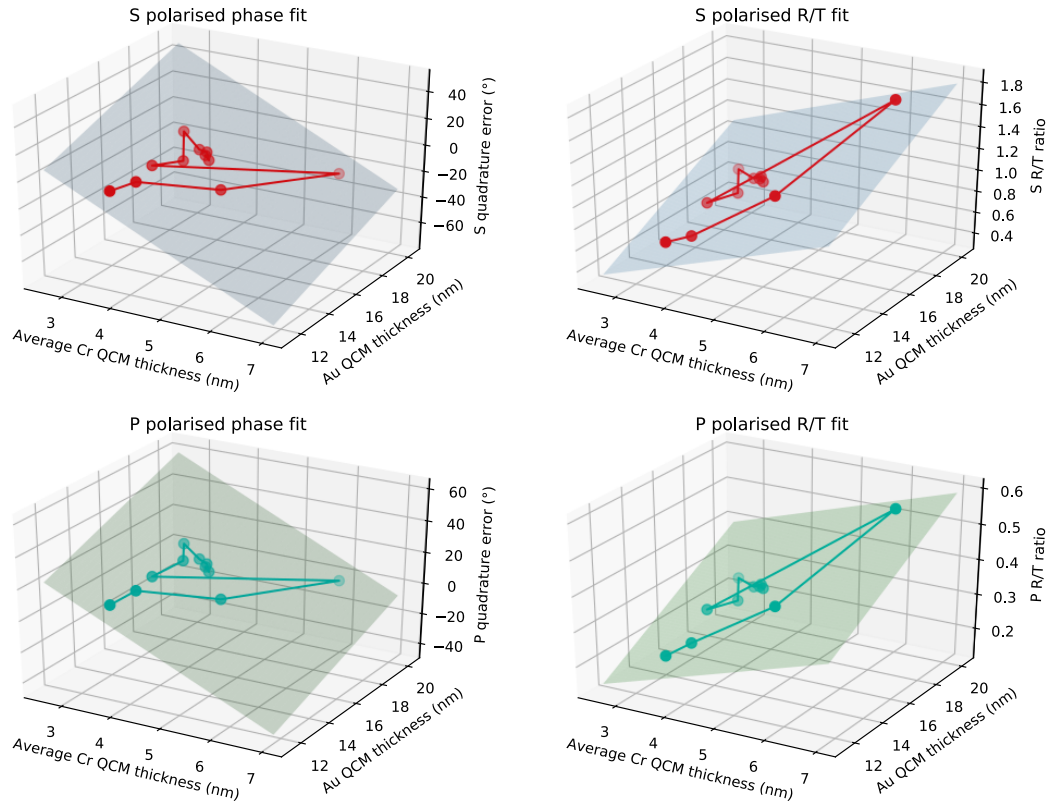


Figure 4.13: Optimisation plane fits to quadrature and reflectivity to transmission ratio measurements for small changes in layer thickness. Phase and reflectivity to transmission ratio data has been plotted against layer thicknesses as measured during deposition with a QCM. In the case of the ideally identical chromium layers, the average thickness of the two layers has been used. A progression of the S polarised data is presented in figure 4.14, with the left most data point corresponding to the first coating in all graphs in this figure.

The final plane fit resulting from this optimisation process is shown in figure 4.13, with the progress of the optimisation process from coating to coating shown in figure 4.14. As can be seen from figures 4.13 and 4.14, this optimisation approach was successful, resulting a coating with the desired 90° phase-quadrature angle and a 1:1 reflectivity to transmission ratio.

Seven phase-quadrature beam splitter cubes were fabricated based upon this

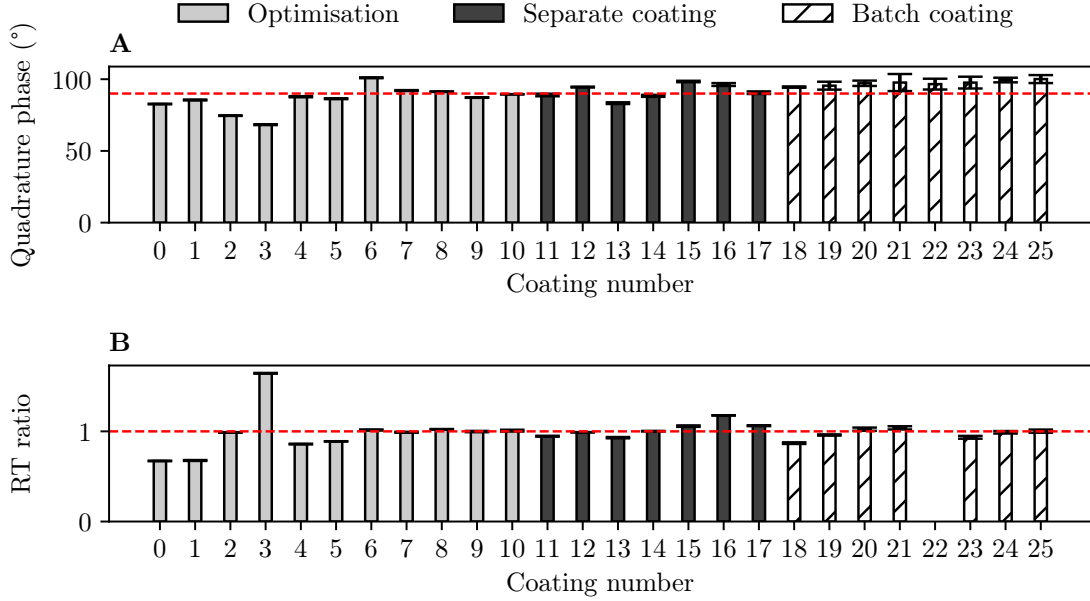


Figure 4.14: Progress of the phase-quadrature coating optimisation process for test optics assembled using microscope immersion oil (optimisation) and repeatability tests assembled using UV cured optical cement (separate coating and batch coating). Separate coatings had phase-quadrature coatings applied in separate coating runs with nominally identical parameters, batch coatings had phase-quadrature coatings applied in the same coating run. Error bars show 95 % confidence limits. Corresponding optimisation plane fits are presented in figure 4.13 for immersion oil measurements. RT ratio data is not shown for coating 22 due to a failed anti-reflection coating.

optimised coating design, with the index matched immersion oil used for the optimisation process replaced with UV cured cement as previously described. These are shown in figure 4.15. Each phase-quadrature coating was applied in a separate coating run to test the repeatability of the coating process. Anti-reflection coatings were applied simultaneously to all beam splitter cubes. Some variation in the properties of the cemented beam splitter cubes was observed, and can be seen for the separate coating data in figure 4.14. In order to investigate whether this lack of repeatability arose from the coating process or the cementing process, a second group of phase-quadrature beam splitter cubes were fabricated with the phase quadrature coatings applied to all prisms to be coated in a single coating run. Anti-reflection coatings were again applied simultaneously to all beam splitter cubes. The data for this batch coated group is also shown in figure 4.14. RT ratio data is not plotted for coating 22 due to a failure in the anti-reflection coating process. The failed anti-reflection coating did not effect the quadrature phase measurement.

The average values of the quadrature phase shift introduced by the coatings for S polarised light, along with the average reflectivity to transmission ratio, for both

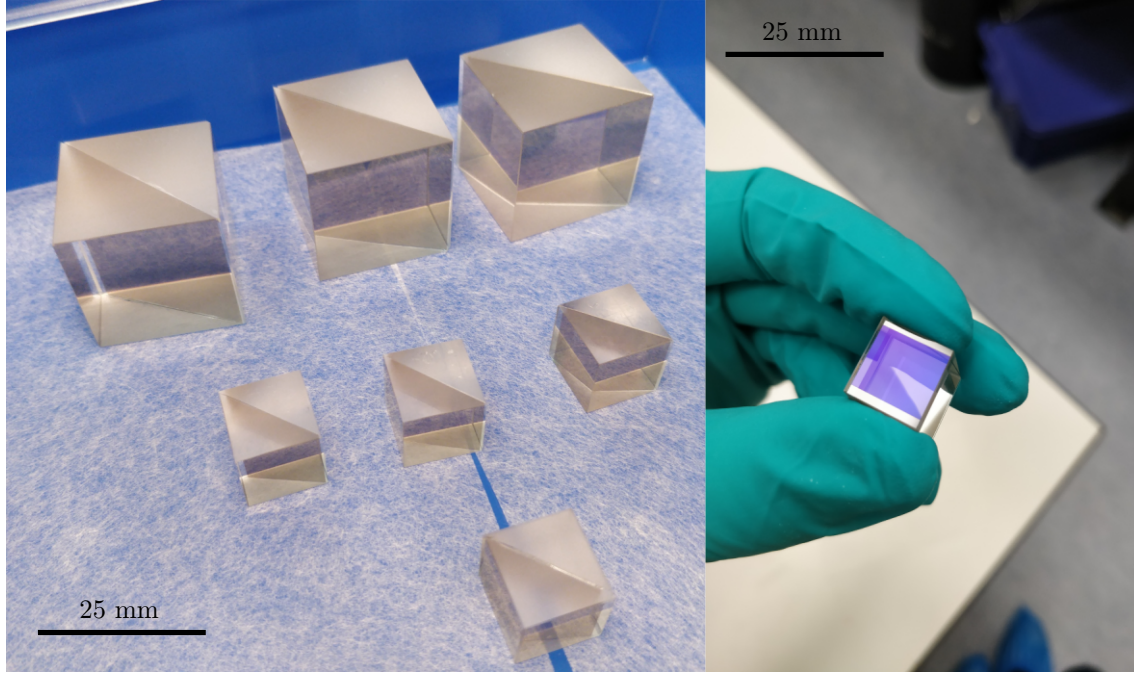


Figure 4.15: Photograph showing a number of the phase-quadrature beam splitter cubes fabricate in the course of this PhD, for applications across the NPL.

the separately coated and batch coated groups are given in table 4.4. Table 4.4 also contains the ranges of quadrature phase shifts and reflectivity to transmission ratios that 95 % of coatings are expected to fall in for each group, to give an indication of the spread of values. The 95 % ranges were calculated assuming a student t distribution with 6 degrees of freedom for the separate coating group, with 7 degrees of freedom for the quadrature phase shift values for the batch coated group, and with 6 degrees of freedom for the reflectivity to transmission ratio values for the batch coated group. The student t distribution was once again employed in this case to calculate the 95 % ranges due to the small sample size, accounting for the uncertainty in the difference between the sample and population means [89].

Table 4.4: Average values and the mean-centred range of values in which 95 % of coatings are expected to fall for the separately and batch coated phase-quadrature cube beam splitter measurements presented in figure 4.14. The range in which 95 % of coatings are expected to fall has been calculated assuming a student t distribution, as described in the text. All values are given for S polarised light.

Coating group	Quadrature phase		RT ratio	
	Average	95 % range	Average	95 % range
Separate	91.5°	$(91.5 \pm 11.7)^\circ$	1.02	1.02 ± 0.19
Batch	97.3°	$(97.3 \pm 4.1)^\circ$	0.97	0.97 ± 0.13

The spread of quadrature phase shift values is smaller for the batch coated group than the separately coated group, indicating that the variability in coating properties is introduced by the coating process, rather than the cementing process. Due to time constraints, the cause of this variability was not investigated. Several potential causes exist, including variable purity of the coating materials, varying coating microstructure, and higher than expected uncertainties in the QCM system. The next step to identify the cause of the coating variability would be to measure the physical thicknesses of the coating layers, for example with AFM measurements. AFM measurements would also give information about the coating microstructure.

For comparison, a survey of existing NPL plane mirror differential interferometers found an average quadrature phase shift of 91.2° , with 95 % of coatings expected to fall in the range $(91.2 \pm 13.2)^\circ$. This suggests a small improvement has been made in coating repeatability, however coatings with unacceptably large quadrature phase shift errors or reflectivity to transmission ratio errors may have been rejected in the past, so a direct comparison is not possible.

4.7 Discussion

For the coating design developed in this work, following the described optimisation process, 95 % of coatings are expected to introduce quadrature phase shifts in the range $(91.5 \pm 11.7)^\circ$ and to have reflectivity to transmission ratios in the range 1.02 ± 0.13 . This shows a slight improvement in repeatability compared to previous NPL designed interferometers, which were found to introduce quadrature phase shifts in the range $(91.2 \pm 13.2)^\circ$. Previous NPL phase-quadrature coatings with undesirable properties may have been rejected, so a direct comparison is not possible. The range of coating properties is relatively large, and no clear cause of the variation between coatings has been identified.

Through further research, it may be possible to improve the repeatability of the coating process, however, the conclusions of chapter 3 indicate that in most practical cases non-linearities are likely to form a significantly larger uncertainty contribution than the additional displacement uncertainty introduced by non-optimal values of quadrature phase shift and reflectivity to transmission ratio. In practical application, the additional uncertainty contribution due to such non-idealities in the quadrature coating is therefore not expected to limit the interferometer performance.

The ultimate performance of phase-quadrature coating based interferometers is limited by the requirement for absorbing optics, and it has been shown that by making use of four optical outputs it is possible for polarisation based quadrature interferometers to achieve higher signal to noise ratios, at the cost of introducing additional optical and electronic complexity. As can be seen from table 4.1, in the shot noise limit the phase-quadrature coating design presented here has higher measurement uncertainties than both polarising interferometer designs, however, these differences, in the range of single digit picometres, are again unlikely to be the limiting factor in practical measurements, where non-linearities will dominate.

The main strength of phase-quadrature coating based interferometry is the small number of optical components required to achieve quadrature outputs. At a minimum, a laser, a phase-quadrature beam splitter, and two corner cube retro-reflectors

are required. This reduces costs, simplifies production and alignment, and most importantly reduces the number of optical surfaces producing unwanted reflections, and therefore potentially introducing higher order non-linearities. As discussed in section 3.3.3, unwanted back reflections may introduce significant non-linearities, typically in the range of tens to hundreds of picometres, and in most practical cases the uncertainty introduced by the presence of such non-linearities will be more significant than the picometre level reductions in displacement uncertainty that are possible with four output polarising interferometers.

As a result of the conclusion of chapter 3, that phase uncertainties will, for the coating designs developed here, be limited by non-linearities rather than signal to noise ratio, the coating design developed following the optimisation process was deemed to be sufficient for the purposes of this work.

4.8 Summary

An overview of the principles behind phase-quadrature coating based interferometry has been given, along with an analysis of the fundamental limits imposed upon the performance of any phase-quadrature coating based interferometer by conservation of energy. In addition, by applying conservation of energy considerations it has been demonstrated that phase-quadrature coatings must be absorbing. A modelling technique has been described that is suitable for designing phase-quadrature coatings made up of such absorbing materials, and the application of this model to the design of a three layer chromium-gold-chromium phase-quadrature coating has been demonstrated.

The methods used to fabricate phase-quadrature coatings have been detailed, along with the optimisation process required to account for experimental deviations from the ideal model, and the measurement approaches used to characterise coating performance. The lack of repeatability in the coating properties has been found not to be a limiting factor in the performance of phase quadrature interferometers. As such, the coating design discussed in this chapter has been used for a variety of applications, and further information of the design and application of phase-quadrature coating based interferometry is given in chapters 7 and 8.

Finally, the consequences of the absorbing nature of phase-quadrature coatings have been investigated. Although four output polarisation based homodyne interferometers maximise signal to noise ratio, the benefits of the reduced number of optical surfaces in phase-quadrature interferometers compared to the polarisation based alternatives appear to outweigh the reduced signal to noise ratio in practical applications, where non-linearities commonly dominate the interferometric phase uncertainty.

Chapter 5

Modelling displacement measuring interferometers

5.1 Introduction

As has already been established, optical interferometers suffer from non-linearities, errors in the measured displacement that are periodic with the wavelength of the illuminating light, introduced by a variety of sources [11], many of which can be corrected using a Heydemann style ellipse fitting correction [21, 87]. Non-linearities that cannot be corrected with ellipse fitting techniques are introduced by the formation of weakly resonant cavities within the optics of interferometer designs [84], breaking the two beam interference assumption typically made when considering displacement measuring interferometer designs. Modelling such cavity formation is therefore desirable if the sources of residual non-linearities in optical interferometers are to be identified, and the development of a model capable of this task is the focus of this thesis chapter.

Complex networks of interconnected cavities may be introduced within the optics of interferometers either by back reflections at material interfaces, or by polarisation leakage in polarising beam splitters, combined with imperfections in other polarising optics. Many displacement measuring interferometer designs [11, 15, 28, 50, 54, 63, 140–144] include polarising beam splitters, and such designs are particularly vulnerable to multiple pass effects due to the combined contributions of polarisation leakage and unwanted back reflections. Polarisation aberrations can be particularly complex to analyse as they may be introduced not only by imperfections in the polarising optics but also by reflections, often from multi-layer or metallic optical coatings, at non-normal angles of incidence, for example within corner cube retro-reflectors [145].

Modelling the effects of these complex networks of cavities with existing techniques commonly employed for interferometer design can be challenging. Jones calculus [146–152], for example, is often applied to modelling optical systems made up of planar polarising optics. Jones calculus models polarising optics by propagating a unidirectional electrical plane wave through a defined series of optical components. Due to this unidirectional treatment all possible optical paths must be modelled separately, with the final electric field amplitudes being combined to

determine the outputs of an optical network. For optical networks containing many interconnected cavities tracing all possible optical paths is an error prone process, and multiple passes must be handled iteratively, with some cut off imposed. An alternative technique is employed by the FINESSE software [153,154], aimed primarily at gravitation wave detector design where optical cavities are a desired feature [155], in which the optical network is considered as a network of nodes, with counter-propagating electric fields coupled together at each node in a similar manner to that employed in thin film transfer matrix modelling techniques [130]. This approach is better suited to the assessment of networks of interconnected cavities (both intentional and unintentional in nature), however the FINESSE software does not account for the effects of polarising optics, and cannot model the effects of reflection from, or transmission through, optically absorbing materials or coatings. As a result of these limitations FINESSE is of limited applicability to displacement measuring interferometer design. Alternative approaches to modelling the effects of multiple passes through optical networks include finite element analysis and ray tracing [156,157] models. Finite element analysis requires a careful mesh design and even then can be impractically computationally intensive as nanoscale resolutions are required over the macroscopic size of the optics to be modelled. Ray tracing approaches must be iterated many times to take into account the effect of cavity formation, and may not produce reliable results, with the effects of cavities of various strengths and optical path lengths being artificially amplified relative to each other depending on the cut-off point imposed on ray propagation. Both ray tracing and finite element analysis approaches require a model of the morphology of the optical components to be constructed, rather than dealing with a relatively simple ‘list’ of optical components as is the case for Jones calculus and FINESSE. Whilst further complexity is added to the model, defining the morphology of the optical components does give finite element analysis and ray tracing techniques the advantage that free-form optical surfaces can be modelled, and finite element techniques have the additional advantage of being able to model diffraction effects. This approach, incorporating orthogonal polarisation states, has, to the knowledge of the author, not previously been described in the literature.

A modelling technique is presented in this chapter that aims to allow the effects of the formation of weakly resonant cavities within networks of planar optical surfaces to be analysed. This approach extends the transfer matrix methodology commonly applied to thin film modelling [20,130,158] to model multiple port devices and the effects of polarisation, combining transfer and scattering matrices to describe an optical network in a manner similar to that of the FINESSE software [153]. This allows the bidirectional propagation of polarised light to be modelled, along with the effects of reflection from and transmission through absorbing (metallic) materials. This is achieved by coupling together pairs of counter-propagating electrical plane waves at nodes in an optical network. By modelling coupled networks of polarising optical components in this way, the effects of polarisation aberrations and unwanted back reflections within the optics of displacement measuring interferometers can be assessed, and used to inform improved interferometer designs.

In this chapter the theoretical background to the technique will be presented first, followed by examples of the scattering and transfer matrices required to model

a range of commonly encountered optical components. The model will then be applied to the simple case of a single back reflection in a polarisation based homodyne interferometer, and compared with an exact analytical solution. A software implementation of this modelling technique, along with much of the work contained in this chapter has already been published [85, 159]. Although an example of the application of this modelling approach will be given in this chapter, the application of the combined scattering and transfer matrix model to the analysis of non-linearity sources in the phase-quadrature coating based interferometers that are the focus of this thesis is reserved for chapter 7, where the results of the modelling process are discussed alongside the optical design of the interferometer, and are used to drive design improvements.

5.2 Theory of the combined scattering and transfer matrix model

The methodology presented here describes a system of optical components with a network of interconnected nodes, with a node at each input/output port of individual optical components, as is the case for the FINESSE software [153]. A complex valued electric field is defined as the sum of two counter-propagating complex valued electrical plane waves travelling along the positive and negative $\hat{z}$ direction [130] at each node m , following the same methodology as the transfer matrix modelling technique described in section 4.3, as

$$\tilde{\mathbf{E}}_m = \tilde{\mathbf{A}}_F e^{i(kz - \omega t + \varphi)} + \mathbf{A}_B e^{i(-kz - \omega t + \varphi)} \quad (5.1)$$

where $\tilde{\mathbf{A}}_F$ and $\tilde{\mathbf{A}}_B$ are the complex amplitudes of the forward and backward propagating components, k is the wavenumber, ω is the angular frequency, t is time and φ is a phase offset. The physical potential of the electrical field at a given time is given by the real component of equation 5.1. These forward and backward propagating components can be further decomposed into orthogonally polarised (along the $\hat{x}$ and $\hat{y}$ directions) counter-propagating components

$$\tilde{E}_{F,x,m} = \tilde{A}_{F,x,m} e^{i(kz - \omega t + \varphi)}, \quad (5.2)$$

$$\tilde{E}_{F,y,m} = \tilde{A}_{F,y,m} e^{i(kz - \omega t + \varphi)}, \quad (5.3)$$

$$\tilde{E}_{B,x,m} = \tilde{A}_{B,x,m} e^{i(-kz - \omega t + \varphi)}, \quad (5.4)$$

$$\tilde{E}_{B,y,m} = \tilde{A}_{B,y,m} e^{i(-kz - \omega t + \varphi)}. \quad (5.5)$$

By making the assumption that we are interested only in steady state (on timescales of the speed of light) solutions, the temporal component of equations 5.2 to 5.5 may be neglected, as any steady state solution must be valid at time zero. The choice of absolute phase is also arbitrary in this case, allowing the phase offset φ to be set to zero. For a given steady state solution the spatial position of each node m is constant, allowing the remaining exponential term to be absorbed into the complex amplitude. A coordinate system may then be chosen such that polarisation components along the $\hat{x}$ and $\hat{y}$ directions represent S and P polarisation states for

any material boundaries at node m , and the ‘forward’ and ‘backward’ components may be denoted as a and b components of the electric field for the sake of brevity of notation. Equations 5.2 to 5.5 can then be rewritten as

$$a_n^{(S)} = \tilde{E}_{F,x,m}, \quad (5.6)$$

$$a_n^{(P)} = \tilde{E}_{F,y,m}, \quad (5.7)$$

$$b_n^{(S)} = \tilde{E}_{B,x,m}, \quad (5.8)$$

$$b_n^{(P)} = \tilde{E}_{B,y,m}. \quad (5.9)$$

The choice of S and P basis states is made for convenience, an arbitrary polarisation state may be decomposed into any orthogonal pair of basis states, and in some cases left and right hand circularly polarised states may be preferable. Following this approach, the vector of electric field components at each node of the model effectively forms a pair of Jones vectors [146], where multiplication of the a or b components by a real valued scalar will change the described power, multiplication by a complex valued scalar will change the phase and power, and multiplication by a complex valued vector may modify the power, phase and polarisation state.

These assumptions along with the choice of coordinate system have several important consequences to the application of the modelling technique. Firstly, by considering only plane wave forms, all optical interfaces are taken to be planar and infinite in extent. As such the effects of diffraction [160] and wavefront curvature [106] cannot be modelled. Secondly by neglecting the temporal component, the model is restricted to monochromatic light, although the results may be readily applied to physical quasi-monochromatic light [20]. An additional result of the plane wave assumption is that all beams are also assumed to be infinite in extent, and as such losses due to misalignment and higher order multiple passes ‘missing’ detectors cannot be modelled. As misalignment cannot be modelled, the results when applied to optical interferometry will tend to give a worst case estimate of non-linearities, which remains useful both as a design tool, and as an indicator of non-linearities that may occur if near perfect alignment is reached. Finally, in order to maintain the physical meaning of the S and P polarised components, changes of coordinate system must take place between optical components where the S and P polarisation states of the material interface planes do not align.

In order to model networks of optical components, the electric field components defined at each node must be linked together. This is achieved in this modelling methodology with a series of scattering and transfer matrix equations. The system of linear equations formed by these scattering and transfer matrix equations is then rewritten in the form of a single network matrix equation, which can be solved to find the values of the electric field components for a given set of boundary conditions.

A scattering matrix is defined here as a matrix linking all outputs of an optical device to all inputs of the same device, which may be generalised to optical devices with any number of ports. A port refers to an input or output of a device; for example, a planar interface between two materials with light normally incident is a two port device and a freespace beam splitter is a four port device. A device with a port number N_p is described by an $2N_p \times 2N_p$ scattering matrix $\mathbf{Q}$, linking input and output S and P polarised components for each port. As such, the scattering

matrix representation of the freespace beam splitter depicted in figure 5.1 is

$$\begin{pmatrix} b_0^{(S)} \\ b_0^{(P)} \\ b_1^{(S)} \\ b_1^{(P)} \\ b_2^{(S)} \\ b_2^{(P)} \\ b_3^{(S)} \\ b_3^{(P)} \end{pmatrix} = \mathbf{Q} \begin{pmatrix} a_0^{(S)} \\ a_0^{(P)} \\ a_1^{(S)} \\ a_1^{(P)} \\ a_2^{(S)} \\ a_2^{(P)} \\ a_3^{(S)} \\ a_3^{(P)} \end{pmatrix}, \quad (5.10)$$

where $\mathbf{Q}$ is an 8×8 scattering matrix, and the components $a_i^{(S)}$, $a_i^{(P)}$, $b_i^{(S)}$ and $b_i^{(P)}$ represent the S and P polarised input and output fields respectively, at the i^{th} port.

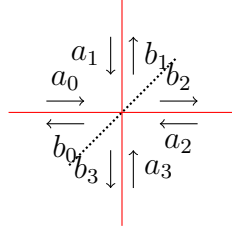


Figure 5.1: Schematic of the electric field components at a four port beam splitter.

Transfer matrices, in contrast to the previously described scattering matrices, are limited to two port devices. A transfer matrix is defined here as the matrix which links the incident and exident fields $a_i^{(S)}$, $a_i^{(P)}$, $b_i^{(S)}$ and $b_i^{(P)}$ at one port of a two port device to the equivalent fields $a_j^{(S)}$, $a_j^{(P)}$, $b_j^{(S)}$ and $b_j^{(P)}$ at the second port. Transfer matrix modelling is a commonly employed technique for multilayer thin film design [20, 130], with transfer matrices used to describe interfaces between materials and propagation matrices to link field components across a single material. For the interfaces depicted in figure 5.2 transfer matrices link the field components through the equations

$$\begin{pmatrix} a_4^{(S)} \\ a_4^{(P)} \\ b_4^{(S)} \\ b_4^{(P)} \end{pmatrix} = \mathbf{T}_{4,5} \begin{pmatrix} a_5^{(S)} \\ a_5^{(P)} \\ b_5^{(S)} \\ b_5^{(P)} \end{pmatrix} \quad (5.11)$$

and

$$\begin{pmatrix} a_6^{(S)} \\ a_6^{(P)} \\ b_6^{(S)} \\ b_6^{(P)} \end{pmatrix} = \mathbf{T}_{6,7} \begin{pmatrix} a_7^{(S)} \\ a_7^{(P)} \\ b_7^{(S)} \\ b_7^{(P)} \end{pmatrix}. \quad (5.12)$$

A special case of the transfer matrix, the propagation matrix, describes the field on

either side of a region made up of a single material with no interfaces. In figure 5.2 a propagation matrix $\mathbf{P}$ links the field amplitudes a_5 and b_5 with a_6 and b_6 as

$$\begin{pmatrix} a_6^{(S)} \\ a_6^{(P)} \\ b_6^{(S)} \\ b_6^{(P)} \end{pmatrix} = \mathbf{P}_{5,6} \begin{pmatrix} a_5^{(S)} \\ a_5^{(P)} \\ b_5^{(S)} \\ b_5^{(P)} \end{pmatrix}. \quad (5.13)$$

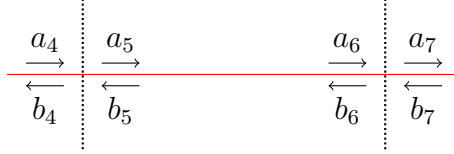


Figure 5.2: Schematic of the electric field components across two plane parallel interfaces.

Given a set of plane parallel interfaces a global transfer matrix may be calculated from a series of transfer and propagation matrices, relating the incident and excident field amplitudes at each side of the stack of layers. As such the field components $a_4^{(S)}$, $a_4^{(P)}$, $b_4^{(S)}$ and $b_4^{(P)}$ may be linked to $a_7^{(S)}$, $a_7^{(P)}$, $b_7^{(S)}$ and $b_7^{(P)}$ in figure 5.2 by a global transfer matrix $\mathbf{T}$ as

$$\begin{pmatrix} a_4^{(S)} \\ a_4^{(P)} \\ b_4^{(S)} \\ b_4^{(P)} \end{pmatrix} = \mathbf{T}_{4,7} \begin{pmatrix} a_7^{(S)} \\ a_7^{(P)} \\ b_7^{(S)} \\ b_7^{(P)} \end{pmatrix}, \quad (5.14)$$

$$= \mathbf{T}_{4,5} \mathbf{P}_{5,6} \mathbf{T}_{6,7} \begin{pmatrix} a_7^{(S)} \\ a_7^{(P)} \\ b_7^{(S)} \\ b_7^{(P)} \end{pmatrix}. \quad (5.15)$$

The forms of the matrices in equations 5.11 to 5.13 are described in the following section, and a more detailed treatment of this technique may be found in section 4.3.

A set of scattering and transfer matrix equations linking together a network of nodes, together with equations defining the inputs to the network, form a set of linear equations that may be written as a single matrix equation. By solving this system using standard numerical techniques the values of the electric fields at various points within the network may be found.

5.3 Matrices of common optical components

The transfer or scattering matrices for a range of optical components commonly employed in interferometry are presented here in order to demonstrate the utility of the modelling methodology. This selection of optical components is sufficient to model a

wide range of optical systems, however it is not an exhaustive list, and the technique is widely applicable to modelling linear optical components. All optical components described here are available in the accompanying software implementation [85].

5.3.1 Free space propagation

The simplest possible transfer matrix describes the propagation of light through free space. For propagation between node 0 and node 1 the transfer matrix equation is given by

$$\begin{pmatrix} a_1^{(P)} \\ b_1^{(P)} \\ a_1^{(S)} \\ b_1^{(S)} \end{pmatrix} = \begin{pmatrix} e^{ik_z d} & 0 & 0 & 0 \\ 0 & e^{-ik_z d} & 0 & 0 \\ 0 & 0 & e^{ik_z d} & 0 \\ 0 & 0 & 0 & e^{-ik_z d} \end{pmatrix} \begin{pmatrix} a_0^{(P)} \\ b_0^{(P)} \\ a_0^{(S)} \\ b_0^{(S)} \end{pmatrix}, \quad (5.16)$$

where k_z is the component of the wavevector normal to the optical interface planes and d is the distance between the interfaces. The above equation assumes an optically homogeneous and isotropic medium, propagation in birefringent media will introduce non-diagonal terms into the propagation matrix. For absorbing media the wavevector will be complex valued, and hence introduce an exponential amplitude decay with propagation distance as expected for negative imaginary components of the complex index when following the sign conventions in this work.

5.3.2 Material interface

The interface between two materials may also be represented as a transfer matrix. Extending the transfer matrix formalism commonly employed in thin film modelling [130] to handle both S and P polarisations [158, 161, 162], as previously described in section 4.3, the transfer matrix equation from the zeroth to the first node may be written in terms of the S and P polarised amplitude reflection and transmission coefficients $\mathbf{r}^{(S)}$, $\mathbf{r}^{(P)}$, $\mathbf{t}^{(S)}$ and $\mathbf{t}^{(P)}$ as

$$\begin{pmatrix} a_1^{(P)} \\ b_1^{(P)} \\ a_1^{(S)} \\ b_1^{(S)} \end{pmatrix} = \begin{pmatrix} \frac{1}{\mathbf{t}^{(P)}} & \frac{\mathbf{r}^{(P)}}{\mathbf{t}^{(P)}} & 0 & 0 \\ \frac{\mathbf{r}^{(P)}}{\mathbf{t}^{(P)}} & \frac{1}{\mathbf{t}^{(P)}} & 0 & 0 \\ 0 & 0 & \frac{1}{\mathbf{t}^{(S)}} & \frac{\mathbf{r}^{(S)}}{\mathbf{t}^{(S)}} \\ 0 & 0 & \frac{\mathbf{r}^{(S)}}{\mathbf{t}^{(S)}} & \frac{1}{\mathbf{t}^{(S)}} \end{pmatrix} \begin{pmatrix} a_0^{(P)} \\ b_0^{(P)} \\ a_0^{(S)} \\ b_0^{(S)} \end{pmatrix}. \quad (5.17)$$

This transfer matrix may be used in combination with the previously described transfer matrix for free space propagation to construct optical paths passing through several materials, and by taking the product of several such matrices propagation through multiple layers of materials may be modelled with a single transfer matrix [130]. For further details of this approach see section 4.3.

5.3.3 Linear polariser

A non-ideal linear polariser is described here by three properties: rotation angle (θ), extinction coefficient (ρ_{ex}) and loss (α). The rotation angle here is measured

clockwise from the S polarised axis looking from the zeroth node to the first although the physical meaning of this rotation will depend on the orientation in which the scattering matrix is coupled into the optical network. The extinction coefficient is defined as the intensity ratio of transmitted light polarised perpendicular to the transmission axis of the polariser, to the intensity of transmitted light polarised parallel to the transmission axis. Loss is similarly defined as the intensity ratio of transmitted light polarised parallel to the transmission axis to the intensity of the component of the incident light polarised along that axis. As such an ideal polariser has an extinction coefficient and loss of zero.

Whilst a linear polariser is a two port device, it is more readily represented as a scattering matrix. Using the above definitions the scattering matrix equation of a polariser aligned to the S polarisation axis is given by

$$\begin{pmatrix} b_0^{(P)} \\ b_0^{(S)} \\ b_1^{(P)} \\ b_1^{(S)} \end{pmatrix} = \begin{pmatrix} 0 & 0 & \sqrt{\rho_{\text{ex}}(1-\alpha)} & 0 \\ 0 & 0 & 0 & \sqrt{1-\alpha} \\ \sqrt{\rho_{\text{ex}}(1-\alpha)} & 0 & 0 & 0 \\ 0 & \sqrt{1-\alpha} & 0 & 0 \end{pmatrix} \begin{pmatrix} a_0^{(P)} \\ a_0^{(S)} \\ a_1^{(P)} \\ a_1^{(S)} \end{pmatrix},$$

$$= \mathbf{Q}_{\text{POL}}^{(S)} \begin{pmatrix} a_0^{(P)} \\ a_0^{(S)} \\ a_1^{(P)} \\ a_1^{(S)} \end{pmatrix}. \quad (5.18)$$

The scattering matrix of a rotated polariser may be found by applying the rotation matrix

$$\mathbf{R}(\theta) = \begin{pmatrix} \cos(\theta) & -\sin(\theta) & 0 & 0 \\ \sin(\theta) & \cos(\theta) & 0 & 0 \\ 0 & 0 & \cos(\theta) & -\sin(\theta) \\ 0 & 0 & \sin(\theta) & \cos(\theta) \end{pmatrix}, \quad (5.19)$$

to the vector of incident field components, effectively a rotation of the coordinate frame, then rotating back into the frame of the model. As such the scattering matrix of a rotated polariser is given by

$$\mathbf{Q}_{\text{POL}}^{(\theta)} = \mathbf{R}(-\theta) \mathbf{Q}_{\text{POL}}^{(S)} \mathbf{R}(\theta). \quad (5.20)$$

5.3.4 Waveplate

A waveplate introducing a phase difference η between the S and P polarised beams, orientated with the fast axis along the S polarised direction may be represented by

the scattering matrix equation

$$\begin{aligned} \begin{pmatrix} b_0^{(P)} \\ b_0^{(S)} \\ b_1^{(P)} \\ b_1^{(S)} \end{pmatrix} &= \begin{pmatrix} 0 & 0 & e^{-\frac{i\eta}{2}} & 0 \\ 0 & 0 & 0 & e^{\frac{i\eta}{2}} \\ e^{-\frac{i\eta}{2}} & 0 & 0 & 0 \\ 0 & e^{\frac{i\eta}{2}} & 0 & 0 \end{pmatrix} \begin{pmatrix} a_0^{(P)} \\ a_0^{(S)} \\ a_1^{(P)} \\ a_1^{(S)} \end{pmatrix}, \\ &= \mathbf{Q}_{\text{WP}}^{(S)} \begin{pmatrix} a_0^{(P)} \\ a_0^{(S)} \\ a_1^{(P)} \\ a_1^{(S)} \end{pmatrix}. \end{aligned} \quad (5.21)$$

This matrix can be seen to be closely related to the corresponding Jones matrix for an optical retarder [146], and Jones matrices can in general be easily rewritten as scattering matrices. In the same manner as previously discussed for a linear polariser, the scattering matrix of a rotated waveplate is given by

$$\mathbf{Q}_{\text{WP}}^{(\theta)} = \mathbf{R}(-\theta) \mathbf{Q}_{\text{WP}}^{(S)} \mathbf{R}(\theta), \quad (5.22)$$

where rotation is again measured clockwise from the S polarised axis looking from the zeroth node to the first.

5.3.5 Faraday rotator

Faraday rotators are commonly employed as optical isolators. Whilst the scattering matrix of an ideal optical isolator could be defined trivially, in some cases it may be of interest to model the effects of multiple reflections through a non-ideal isolator. Such a device may be modelled using the previously described equations for a linear polariser, in combination with a Faraday rotator.

The scattering matrix of a Faraday rotator with loss α as previously defined is given by

$$\mathbf{Q}_{\text{FR}}^{(\psi)} = \sqrt{1 - \alpha} \begin{pmatrix} 0 & 0 & \cos(\psi) & -\sin(\psi) \\ 0 & 0 & \sin(\psi) & \cos(\psi) \\ \cos(\psi) & -\sin(\psi) & 0 & 0 \\ \sin(\psi) & \cos(\psi) & 0 & 0 \end{pmatrix}, \quad (5.23)$$

where ψ is the angle of the polarisation rotation introduced by the polariser, with rotation measured clockwise looking from the zeroth node to the first.

5.3.6 Non-polarising beam splitter

The properties of an ideal lossless non-polarising beam splitter may be defined by the ratio of reflected intensity to transmitted intensity, ζ . For a set value of ζ , the

amplitude reflectivity $\mathfrak{r}$ and transmission $\mathfrak{t}$ coefficients are given by

$$\mathfrak{r} = \sqrt{\frac{\zeta}{\zeta + 1}} \quad (5.24)$$

and

$$\mathfrak{t} = \sqrt{\frac{1}{\zeta + 1}}. \quad (5.25)$$

Referring back to figure 5.1, the scattering matrix equation linking the four nodes is given by

$$\begin{pmatrix} b_0^{(P)} \\ b_0^{(S)} \\ b_1^{(P)} \\ b_1^{(S)} \\ b_2^{(P)} \\ b_2^{(S)} \\ b_3^{(P)} \\ b_3^{(S)} \end{pmatrix} = \begin{pmatrix} 0 & 0 & \mathfrak{r} & 0 & \mathfrak{t} & 0 & 0 & 0 \\ 0 & 0 & 0 & \mathfrak{r} & 0 & \mathfrak{t} & 0 & 0 \\ \mathfrak{r} & 0 & 0 & 0 & 0 & 0 & \mathfrak{t} & 0 \\ 0 & \mathfrak{r} & 0 & 0 & 0 & 0 & 0 & \mathfrak{t} \\ \mathfrak{t} & 0 & 0 & 0 & 0 & 0 & \mathfrak{r} & 0 \\ 0 & \mathfrak{t} & 0 & 0 & 0 & 0 & 0 & \mathfrak{r} \\ 0 & 0 & \mathfrak{t} & 0 & \mathfrak{r} & 0 & 0 & 0 \\ 0 & 0 & 0 & \mathfrak{t} & 0 & \mathfrak{r} & 0 & 0 \end{pmatrix} \begin{pmatrix} a_0^{(P)} \\ a_0^{(S)} \\ a_1^{(P)} \\ a_1^{(S)} \\ a_2^{(P)} \\ a_2^{(S)} \\ a_3^{(P)} \\ a_3^{(S)} \end{pmatrix}. \quad (5.26)$$

From this equation it is clear that the coupling between individual nodes and polarisations can be readily changed. Realistic beam splitter coatings can be modelled if desired by calculating the amplitude reflection and transmission coefficients for each polarisation (the Fresnel coefficients) using transfer matrix modelling techniques from the layer structure of the beam splitter coating.

5.3.7 Polarising beam splitter

A polarising beam splitter can be defined by a set of coefficients similar to those used previously to define a linear polariser. Extinction coefficients must be defined separately for the reflected and transmitted beams as $\rho_{\text{ex,R}}$, $\rho_{\text{ex,T}}$, and similarly loss coefficients for each polarisation as $\alpha^{(S)}$ and $\alpha^{(P)}$. This distinction is necessary as for commonly available thin film based polarising beam splitters, the extinction coefficient of the reflected beam is often significantly poorer than that of the transmitted beam, due to the reliance on Brewster's angle reflections. A polarising beam splitter aligned such that S polarised light is reflected and P polarised light transmitted may

be described with the scattering matrix equation

$$\begin{pmatrix} b_0^{(P)} \\ b_0^{(S)} \\ b_1^{(P)} \\ b_1^{(S)} \\ b_2^{(P)} \\ b_2^{(S)} \\ b_3^{(P)} \\ b_3^{(S)} \end{pmatrix} = \begin{pmatrix} 0 & 0 & \mathbf{r}^{(P)} & 0 & \mathbf{t}^{(P)} & 0 & 0 & 0 \\ 0 & 0 & 0 & \mathbf{r}^{(S)} & 0 & \mathbf{t}^{(S)} & 0 & 0 \\ \mathbf{r}^{(P)} & 0 & 0 & 0 & 0 & 0 & \mathbf{t}^{(P)} & 0 \\ 0 & \mathbf{r}^{(S)} & 0 & 0 & 0 & 0 & 0 & \mathbf{t}^{(S)} \\ \mathbf{t}^{(P)} & 0 & 0 & 0 & 0 & 0 & \mathbf{r}^{(P)} & 0 \\ 0 & \mathbf{t}^{(S)} & 0 & 0 & 0 & 0 & 0 & \mathbf{r}^{(S)} \\ 0 & 0 & \mathbf{t}^{(P)} & 0 & \mathbf{r}^{(P)} & 0 & 0 & 0 \\ 0 & 0 & 0 & \mathbf{t}^{(S)} & 0 & \mathbf{r}^{(S)} & 0 & 0 \end{pmatrix} \begin{pmatrix} a_0^{(P)} \\ a_0^{(S)} \\ a_1^{(P)} \\ a_1^{(S)} \\ a_2^{(P)} \\ a_2^{(S)} \\ a_3^{(P)} \\ a_3^{(S)} \end{pmatrix},$$

$$= \mathbf{Q}_{\text{PBS}} \begin{pmatrix} a_0^{(P)} \\ a_0^{(S)} \\ a_1^{(P)} \\ a_1^{(S)} \\ a_2^{(P)} \\ a_2^{(S)} \\ a_3^{(P)} \\ a_3^{(S)} \end{pmatrix}. \quad (5.27)$$

Defining the reflected and transmitted extinction coefficients as

$$\rho_{\text{ex},R} = \frac{|\mathbf{r}^{(P)}|^2}{|\mathbf{r}^{(S)}|^2} \quad (5.28)$$

and

$$\rho_{\text{ex},T} = \frac{|\mathbf{t}^{(S)}|^2}{|\mathbf{t}^{(P)}|^2} \quad (5.29)$$

and the S and P polarised loss coefficients as

$$\alpha^{(S)} = |\mathbf{r}^{(S)}|^2 + |\mathbf{t}^{(S)}|^2 \quad (5.30)$$

and

$$\alpha^{(P)} = |\mathbf{r}^{(P)}|^2 + |\mathbf{t}^{(P)}|^2 \quad (5.31)$$

leads to

$$|\mathbf{r}^{(S)}|^2 = \frac{\alpha^{(S)} - \alpha^{(P)} \rho_{\text{ex},T}}{1 - \rho_{\text{ex},T} \rho_{\text{ex},R}}, \quad (5.32)$$

$$|\mathbf{r}^{(P)}|^2 = \frac{\rho_{\text{ex},R} (\alpha^{(S)} - \alpha^{(P)} \rho_{\text{ex},T})}{1 - \rho_{\text{ex},T} \rho_{\text{ex},R}}, \quad (5.33)$$

$$|\mathbf{t}^{(S)}|^2 = \frac{\alpha^{(P)} \rho_{\text{ex},T} - \alpha^{(S)} \rho_{\text{ex},T} \rho_{\text{ex},R}}{1 - \rho_{\text{ex},T} \rho_{\text{ex},R}} \quad (5.34)$$

and

$$|\mathbf{t}^{(\text{P})}|^2 = \frac{\alpha^{(\text{P})} - \alpha^{(\text{S})}\rho_{\text{ex},R}}{1 - \rho_{\text{ex},T}\rho_{\text{ex},R}}. \quad (5.35)$$

Some care must be taken to ensure that the intensity reflectivity and transmission coefficients remain physical (in the interval $[0, 1]$) when applying these equations, however for real world values of

$$\alpha^{(\text{P})/(\text{S})} \approx 1 \quad (5.36)$$

and

$$\rho_{\text{ex},R/T} \approx 0 \quad (5.37)$$

equations 5.32 to 5.35 remain valid.

The polarising beam splitter may be rotated about two axes; the axis of transmission from node zero to node two (θ_0) and the axis of transmission from node one to node three (θ_1). As with previously defined rotation angles, these rotations are defined clockwise from the S polarised axis, looking from the zeroth node to the second, and from the first node to the third. This may be accomplished with the block diagonal 8×8 rotation matrix

$$\mathbf{R}(\theta_0, \theta_1) = \begin{pmatrix} \mathbf{U}(\theta_0) & 0 & \dots & 0 \\ 0 & \mathbf{U}(\theta_1) & \dots & 0 \\ \vdots & \vdots & \mathbf{U}(\theta_0) & \vdots \\ 0 & 0 & \dots & \mathbf{U}(\theta_1) \end{pmatrix}, \quad (5.38)$$

where the sub-matrices are given by

$$\mathbf{U}(\theta_i) = \begin{pmatrix} \cos(\theta_i) & -\sin(\theta_i) \\ \sin(\theta_i) & \cos(\theta_i) \end{pmatrix}, \quad (5.39)$$

such that the scattering matrix of the rotated polarising beam splitter is given by

$$\mathbf{Q}_{\text{PBS}}^{(\theta_0, \theta_1)} = \mathbf{R}(-\theta_0, -\theta_1) \mathbf{Q}_{\text{PBS}} \mathbf{R}(\theta_0, \theta_1). \quad (5.40)$$

5.3.8 Mirror

Mirrors may be modelled in two ways: firstly, by simply setting the incident and exident components to be equal for each polarisation state at a port of another optical component, for example a region of free space propagation. Alternatively, the mirror can be modelled as a material interface, accounting for the lossy properties of metallic mirrors. If this is done, the field amplitude incident on the back surface of the mirror must be fixed, in most practical applications at zero.

5.4 Modelling a polarisation splitting homodyne interferometer

To demonstrate the applications of the model, a homodyne interferometer with polarisation based quadrature fringe counting will be modelled. A schematic of this interferometer is shown in figure 5.3, with the field pairs at each node (0 through 21) shown.

Light is input into this interferometer polarised at 45° to the vertical axis, and is then split into two arms by a polarising beam splitter (PBS). A double pass through a quarter wave plate (QWP) in each arm then switches the polarisation allowing the previously reflected beam to transmit through the PBS and vice versa. In order to perform directional fringe counting two signals are required with a phase difference between them, ideally of 90° (quadrature). To achieve this, the output of the PBS is split with a non-polarising beam splitter (NPBS). For one output, the polarisations are mixed with a polariser at 45° degrees to the vertical, then detected at photodiode one (PD_1). For the second output (PD_2) a QWP with the fast axis vertical is used before the mixing polariser to introduce a 90° phase difference between the measurement and reference beams.

5.4.1 Model construction

The field components may be linked together with transfer and scattering matrices as previously described. For the interferometer depicted in figure 5.3 the first equation needed is the input, given by

$$\begin{pmatrix} a_0^{(P)} \\ a_0^{(S)} \end{pmatrix} = \begin{pmatrix} \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{pmatrix} \quad (5.41)$$

in the ideal case. The equation of the polarising beam splitter is then given by

$$\begin{pmatrix} b_1^{(P)} \\ b_1^{(S)} \\ b_2^{(P)} \\ b_2^{(S)} \\ b_3^{(P)} \\ b_3^{(S)} \\ b_4^{(P)} \\ b_4^{(S)} \end{pmatrix} = \mathbf{Q}_{\text{PBS}} \begin{pmatrix} a_1^{(P)} \\ a_1^{(S)} \\ a_2^{(P)} \\ a_2^{(S)} \\ a_3^{(P)} \\ a_3^{(S)} \\ a_4^{(P)} \\ a_4^{(S)} \end{pmatrix}, \quad (5.42)$$

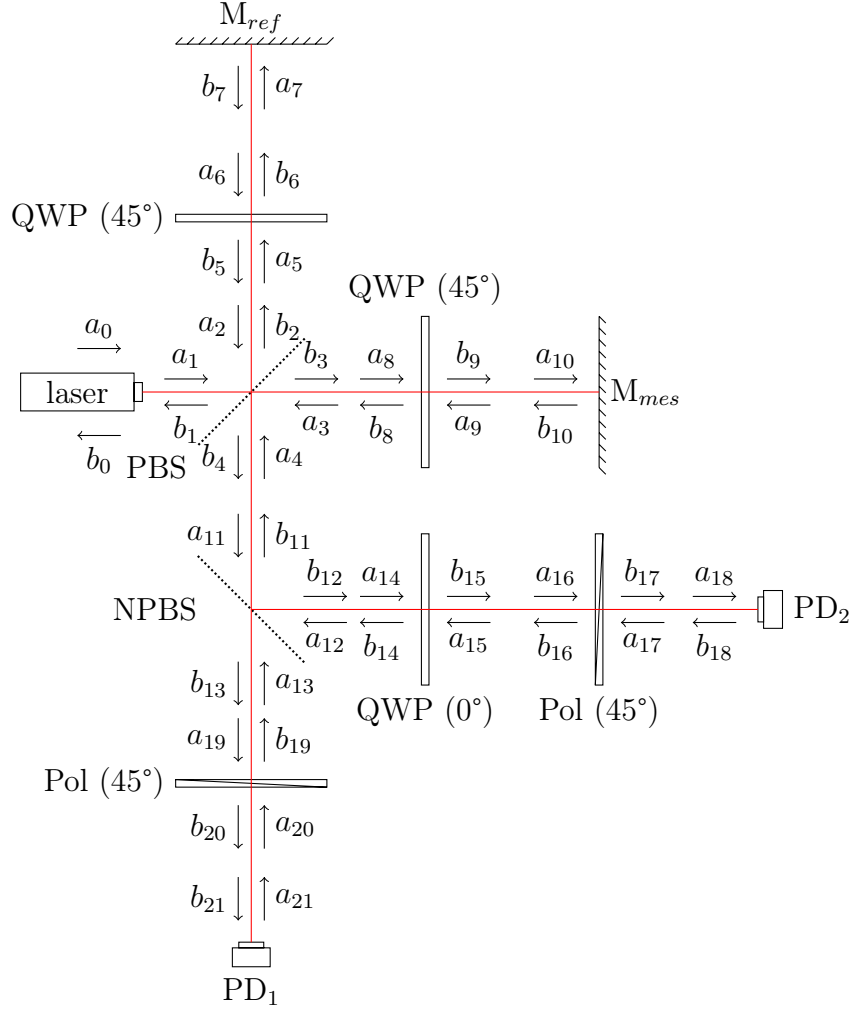


Figure 5.3: Schematic of the electric field components throughout a homodyne Michelson interferometer with polarising quadrature detection. Counter propagating field components are denoted by a_n and b_n , with the direction of propagation indicated by the corresponding arrows. QWP, quarter wave plate; PBS, polarising beam splitter; NPBS, non-polarising beam splitter; Pol, polariser; PD, photodiode; M_{ref} and M_{mes} , reference and measurement mirrors.

and the reference arm quarter wave plate by

$$\begin{pmatrix} b_5^{(P)} \\ b_5^{(S)} \\ b_6^{(P)} \\ b_6^{(S)} \end{pmatrix} = \mathbf{Q}_{WP}^{(45^\circ)} \begin{pmatrix} a_5^{(P)} \\ a_5^{(S)} \\ a_6^{(P)} \\ a_6^{(S)} \end{pmatrix}. \quad (5.43)$$

Similar equations follow for the other quarter wave plates, polarisers and the non-polarising beam splitter, leaving only the mirrors and photodiodes. Ideal mirrors may be modelled as previously discussed by setting the incident and exident beams

to be equal, for example

$$\begin{pmatrix} b_7^{(P)} \\ b_7^{(S)} \end{pmatrix} = \begin{pmatrix} a_7^{(P)} \\ a_7^{(S)} \end{pmatrix} \quad (5.44)$$

for the reference mirror. The field component incident from the back surface of the photodiodes must be set to zero (or another known value) for the model to be solvable. Back reflections from the photodiodes could be modelled as a material interface if desired, although if the photodiodes are to be angled, as is common to avoid back reflections in experimental work, fixing the incident field at zero may produce more applicable results.

Each component must then be linked together. Some care is needed to ensure that the outputs and inputs to each component are coupled correctly when mixing transfer and scattering matrices; for scattering matrices the ‘a’ component is always the input, for transfer matrices this is only the case for one side of the stack of materials described by the transfer matrix. In practise this problem can be solved by ensuring all components are linked by transfer matrices, with no scattering matrix to scattering matrix connections, and then checking the directions of the field components linked to each end of the transfer stack. In this case the laser may be connected to the beam splitter with the transfer matrix equation

$$\begin{pmatrix} a_0^{(P)} \\ a_0^{(S)} \\ b_0^{(P)} \\ b_0^{(S)} \end{pmatrix} = T_{0,1} \begin{pmatrix} a_1^{(P)} \\ a_1^{(S)} \\ b_1^{(P)} \\ b_1^{(S)} \end{pmatrix} \quad (5.45)$$

whereas the polarising beam splitter and reference arm quarter wave plate are linked by the equation

$$\begin{pmatrix} b_2^{(P)} \\ b_2^{(S)} \\ a_2^{(P)} \\ a_2^{(S)} \end{pmatrix} = T_{2,5} \begin{pmatrix} a_5^{(P)} \\ a_5^{(S)} \\ b_5^{(P)} \\ b_5^{(S)} \end{pmatrix}. \quad (5.46)$$

Note that the order of the ‘a’ and ‘b’ components has been switched in the above equation. Together, these scattering and matrix equations form a system of linear equations that may be written in the form of a single 88×88 network matrix equation,

$$\begin{pmatrix} 0 \\ \vdots \\ \vdots \\ \vdots \\ \vdots \\ 0 \\ a_0^{(P)} \\ a_0^{(S)} \end{pmatrix} = \mathbf{T}_{\text{NET}} \begin{pmatrix} a_0^{(P)} \\ a_0^{(S)} \\ b_0^{(P)} \\ b_0^{(S)} \\ \vdots \\ a_{21}^{(P)} \\ a_{21}^{(S)} \\ b_{21}^{(P)} \\ b_{21}^{(S)} \end{pmatrix}. \quad (5.47)$$

For small systems it may be possible to find an analytical solution to this network equation, however in the vast majority of practical applications it is preferable to find a numerical solution.

5.4.2 Model results

Applying the previously described modelling methodology to the interferometer depicted in figure 5.3 results, in the ideal case with the polarisation axes of all components perfectly aligned and no unwanted back reflections present, in the outputs (for nodes 18 and 21) shown in figure 5.4. As expected, the model produces sinusoidal signals with a phase difference of 90° , resulting in a circular Lissajous.

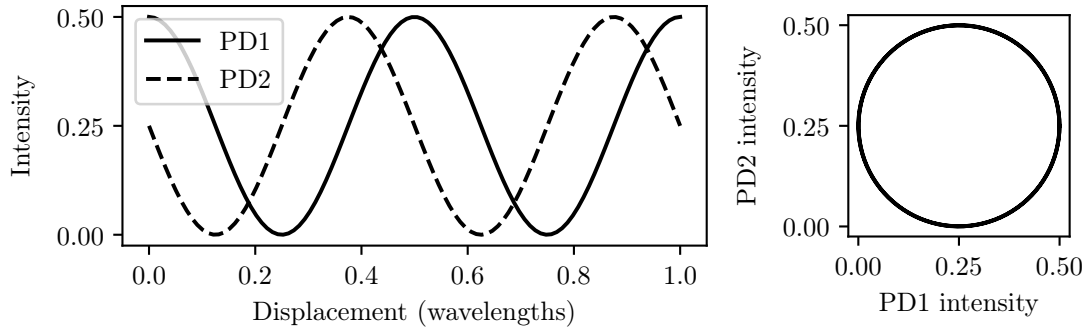


Figure 5.4: Ideal outputs of interferometer depicted in figure 5.3. Intensities are normalised to the input intensity.

Non-ideal behaviour may be introduced into the modelled signals by misaligning the waveplate in the PD2 output arm of the interferometer. Introducing a large rotational misalignment of 20° into the model results in the outputs shown in figure 5.5. This rotation of the output quarter wave plate introduces both an error in the phase difference between the output signals, and a reduction in the amplitude of the PD2 signal. Such effects would introduce errors if the interferometer outputs were to be used for displacement measurement [21].

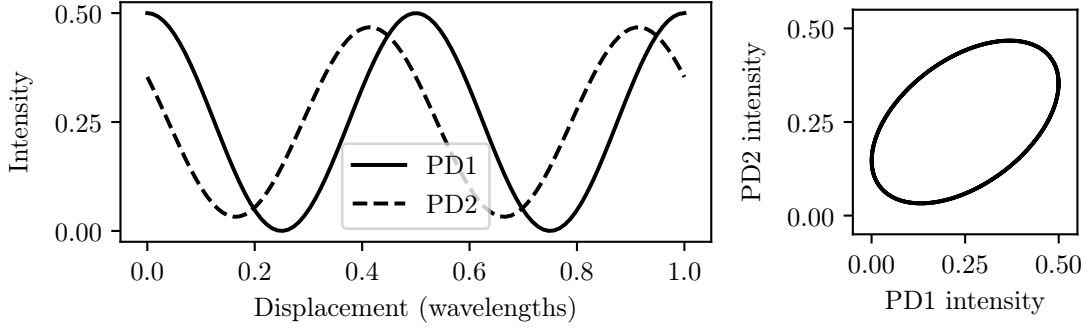


Figure 5.5: Outputs of interferometer depicted in figure 5.3 with the PD2 output quarter waveplate rotated by 20° . Intensities are relative to input intensity.

So far, the results shown are identical to those that could have been reached through Jones matrix calculation, and show only sinusoidal errors (that is to say, errors that result in a distortion of the output signals to a different sinusoid) that could be corrected for in a displacement measuring interferometer with a Heydemann style correction [87]. Of more interest with respect to the benefits of the proposed modelling approach are the effects of multiple reflections on the interferometer outputs, which in general result in the distortion of the output signals to a non-sinusoidal form. Finding an analytical form for the interferometer outputs with the presence of back reflections is challenging, indeed that is the problem that motivated the development of this model. In the simplified case of a single back reflection in the measurement arm of an otherwise ideal interferometer however, the analytical form is relatively simple and can be calculated for comparison with the modelled results. For an ideal polarising beam splitter and perfectly aligned waveplates, if a reflection at an air glass interface takes place between nodes 3 and 8 in figure 5.3 the amplitudes of the electric field components $b_4^{(P)}$ and $b_4^{(S)}$ are given by

$$b_4^{(P)} = \tilde{A}^{(S)} e^{ik2d_{ref}} \quad (5.48)$$

and

$$b_4^{(S)} = \tilde{A}^{(P)} t_{g \rightarrow a} t_{a \rightarrow g} \left(e^{ik2d_{mes}} + \sum_{n=0}^{\infty} r_{a \rightarrow g}^{2n+2} e^{ik(4n+6)d_{mes}} \right) \quad (5.49)$$

$$= \tilde{A}^{(P)} t_{g \rightarrow a} t_{a \rightarrow g} \left(e^{ik2d_{mes}} - \frac{r_{a \rightarrow g}^2 e^{ik6d_{mes}}}{r_{a \rightarrow g}^2 e^{ik4d_{mes}} - 1} \right) \quad (5.50)$$

where $\tilde{A}^{(S)}$ and $\tilde{A}^{(P)}$ are the S and P polarised amplitudes of the incident light, d_{ref} and d_{mes} are the lengths of the reference and measurement arms respectively, $r_{a \rightarrow g}$, $t_{g \rightarrow a}$ and $t_{a \rightarrow g}$ are the amplitude air to glass reflectivity, glass to air transmission and air to glass transmission coefficients, and k is the wavevector. From these amplitudes, the detected intensity at PD1 and PD2 may be calculated. If a refractive index of 1.515 is taken for the glass and 1 for the air, and the reference arm length is taken

to be an exact multiple of the wavelength, equations 5.48 and 5.50 along with the model result in the signals shown in figure 5.6. The model produces the expected result, with the combination of the quarter waveplate and polarising beam splitter preventing alternate multiple reflections from reaching the detector. If such signals were to be used for displacement measurement, even if a Heydemann style correction were to be applied, errors in the measured displacement would still be present.

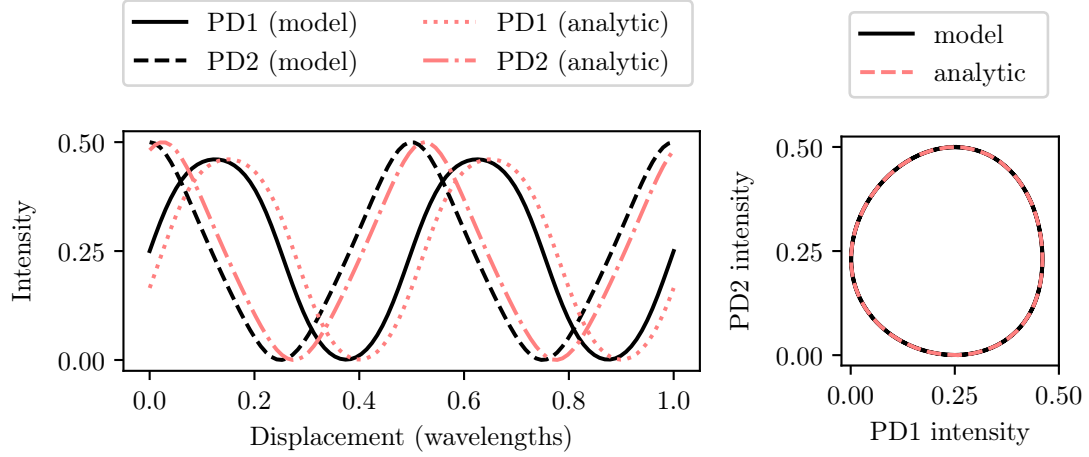


Figure 5.6: Outputs of interferometer depicted in figure 5.3 with a reflection for a glass-air interface between nodes 2 and 5, as calculated by both the described model and an exact analytical approach. A displacement offset has been added to the analytic signals in the left hand graph for clarity. Intensities are relative to input intensity.

The non-circular nature of the Lissajous figure, and the presence of additional frequencies, can be more clearly seen in figure 5.7, where the refractive index of the glass has been increased to 3 to produce a particularly extreme example of the effects of back reflections. The results of a first order application of equation 5.49 are also shown in figure 5.7, the same result that would arise from applying a second order (the first order would be eliminated by the polarising beam splitter) Jones matrix model.

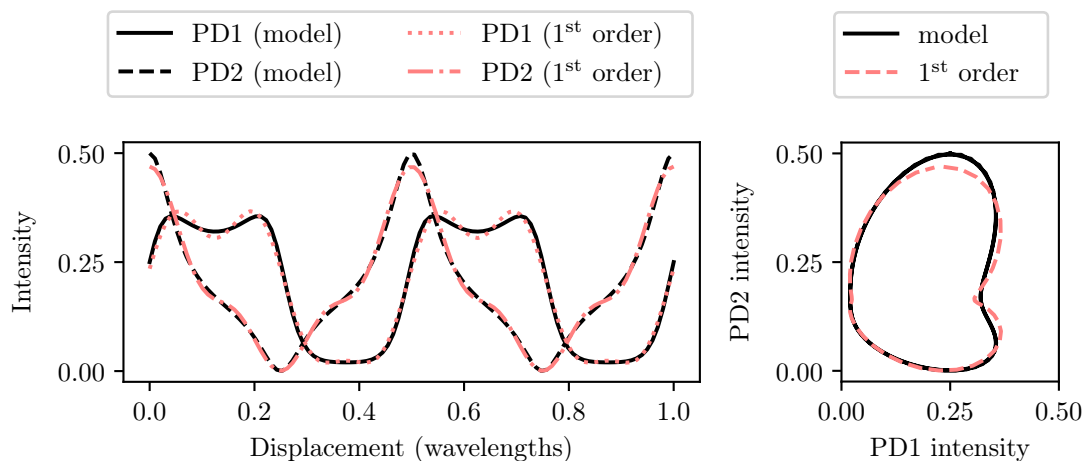


Figure 5.7: Outputs of interferometer depicted in figure 5.3 with a reflection for a glass-air interface between nodes 2 and 5, as calculated by both the described model and a first order analytical approximation. Intensities are relative to input intensity.

5.5 Discussion

The capability of the combined scattering and transfer matrix modelling approach to model the effects of cavity formation within the optics of homodyne interferometers has been demonstrated. In the case of the polarisation based quadrature interferometer cavity effects introduced by a single source of unwanted reflection have been compared for both the described modelling technique and an exact analytical solution, and agreement has been found.

In addition to the demonstrated ability to evaluate the outputs of an interferometer, the model also allows the electric field at other points within the interferometer optics to be investigated without having to re-derive the possible beam paths, for example back reflections into the laser source.

This modelling approach does however have some important caveats that must be considered when analysing the model results. Firstly, the outputs of the model place an upper limit on the effects of multiple passes within the interferometer optics on the output signals; in reality the optical surfaces will not be perfectly aligned, and this misalignment will result in an attenuation of higher order multiple passes for finite diameter beams. Additional sources of round trip loss are likely to be present in real world interferometer designs, some of which can be included in this model. For example, non-ideal mirrors may be included based on known properties of the optical components. More complex loss sources, in particular losses introduced by small misalignments that preferentially attenuate higher order reflections, cannot be fully modelled with the techniques described here.

Secondly, ideal lossless cavities cannot be modelled, as such cavities will result in a singular network matrix. Singular network matrices can also be introduced by cavities containing components with optical gain. Whilst this is an unlikely case to

be implemented intentionally, gain may be introduced by errors in the implementation of transmission through absorbing materials if sign conventions for complex refractive indices are not kept consistent.

Finally, when considering multiple coupled cavities the relative phases of multiple light paths can have a significant effect on the resulting output. This can present a problem when considering macroscopic optical elements, as the exact spacing between surfaces may not be known. In this case the model can be used to find the maximum and minimum distortions to the desired ideal signals. Similar techniques can be used to identify the combined effect of polarisation aberrations arising from multiple optical components based upon the specifications of the various components.

Taking these caveats into account, the modelling approach is most valuable when combined with experimental measurements. In particular, the model may be used to identify the spatial period of non-linearities arising from a range of sources, allowing the sources of non-linearities found in experimental measurements to be located, and if possible eliminated. The model may also be used to place upper limits on the non-linearities that may arise in a proposed interferometer design, and to identify which optical components are critical to the performance of the system.

5.6 Summary

A modelling approach has been developed that permits the effects of multiple reflections through arbitrary networks of planar optical interfaces, including polarising components, to be modelled, extending prior comparable approaches [153] that cannot describe the effects of polarisation. Scattering or transfer matrices for a range of commonly used optical components have been described. The application of this modelling technique to a polarisation splitting homodyne interferometer has been demonstrated, and agreement between the model and an exact analytical solution has been found under simplified conditions.

The modelling methodology presented here is well suited to computational implementation, and software has been developed [85, 159] to automate the generation and solving of a network matrix equation from a list of optical components coupled at nodes. This software has been published as an open source python module, backed by a transfer matrix modelling library implemented in C.

The work presented here focuses on the application of the described modelling technique to displacement measuring interferometers, however the method is widely applicable to any network of planar optical components. Future work may extend the modelling technique to include the treatment of birefringent optics from first principles, integrating existing transfer matrix methodologies for anisotropic layers [158, 161, 162]. Although fundamentally a homodyne technique, another avenue for future work is the application of the model to heterodyne interferometers by considering the sum of the amplitude outputs of two or more homodyne interferometers operating at differing optical frequencies. The combined scattering and transfer matrix modelling technique is applied to a more complex optical system in chapter 7, where the model is applied to identifying complex non-linearity sources in the NPL plane mirror differential interferometer design.

Chapter 6

Correction of non-linearities in homodyne interferometry

6.1 Introduction

Non-linearities, defined in this work as errors in the measured interferometric phase that are periodic with some harmonic of the illuminating wavelength, affect almost all interferometric measurements of displacement. As discussed in chapter 3, non-linearity may be the dominant uncertainty source for short range measurements, before refractive index uncertainties begin to dominate, for measurements made in air. As such significant effort has been invested in the correction of non-linearities in homodyne interferometry, starting with the ellipse fitting Heydemann correction [87], a number of variations of which have been proposed in the literature [21, 83, 111, 163–166]. In addition to ellipse fitting corrections, other non-linearity correction algorithms have been proposed in the literature aiming specifically at making real time corrections in cases where non-linearities may vary on a fringe by fringe basis [167–170]. Some of these correction methodologies are aimed specifically at the correction of displacements covering less than one full optical fringe [163, 169]. A final class of ellipse fitting corrections aim to correct for higher order non-linearities in homodyne interferometers, with one such method previously proposed in the literature [68].

This chapter will first review existing non-linearity correction algorithms, with a focus on techniques that are suitable for precision interferometry applications. A detailed comparison of existing ellipse fitting correction algorithms will be made, considering the performance of the algorithms, both for displacements covering multiple optical fringes, and for sub-fringe displacements. Although in most cases experimental measurements of non-linearity are preferable, as discussed in chapter 3, the ability to propagate uncertainties through the correction algorithms will also be considered. A brief review of algorithms suitable for real time application will also be presented, although these algorithms will not be compared in detail as existing approaches are designed for higher speed, lower precision applications, and are therefore not within the direct scope of this thesis.

Following the review of existing techniques, three new methodologies are proposed, each aiming to fill a specific gap in the performance of existing corrections:

1. Two-wavelength non-linearity correction.
2. Multiple intensity reference non-linearity correction.
3. Measurement arm normalisation non-linearity correction.

The first employs two-wavelength interferometry to correct for harmonic distortion in the quadrature signals, therefore correcting for higher order non-linearities, which cannot be corrected by the Heydemann correction. The second makes use of intensity reference signals taken from the return beams in each arm of the interferometer separately, allowing the offsets, and in some cases the gain ratio, of the quadrature signals to be corrected for before a constrained ellipse fit is made, and is aimed specifically at reducing non-linearities in cases where the displacement covers a small portion of an optical fringe, an area where traditional ellipse fitting corrections leave significant residual non-linearities, as will be demonstrated in section 6.2.1. The final methodology also makes use of intensity references taken from the return beams in each arm of the interferometer, but instead focusses upon using the partial information about unwanted higher order effects contained within these intensity references to correct for higher order non-linearities. This final methodology is also employed by the interferometer described in chapter 7, and therefore although the methodology is introduced here, detailed discussion of the performance and limitations in practical application, along with the presentation of experimental data, is reserved for chapter 7. The novel non-linearity correction methodologies presented in this chapter have previously been published [69], or submitted to a journal [171, 172], in the course of this PhD, and such much of the text presented in the corresponding sections of this chapter is reproduced from those publications.

Finally, the various non-linearity correction methodologies proposed both in the literature, and in this work, are summarised, and the most appropriate methodology for different measurement challenges identified.

6.2 Review of existing techniques

6.2.1 Ellipse-fitting algorithm comparison

Ellipse fitting for the purposes of non-linearity correction was first proposed by Heydemann [87], and was later applied to precision optical interferometry by Birch [21] and later authors [83]. Whilst the basic principle of the ellipse fitting correction (as introduced in section 3.3) is relatively simple, with either the quadrature signals being corrected using equation 3.26 and 3.27, or the phase measurement being corrected directly with equation 3.29, the quality of the ellipse fit, and therefore the choice of fitting algorithm, may have a significant effect on residual non-linearities.

Numerous ellipse fitting algorithms have been proposed in the literature [173–177] that go beyond the least squares conic fit outlined in section 3.3, with many developed for applications within the field of computer vision. Several such algorithms have been applied to optical interferometry [163–165]. Additionally, algorithms have been developed specifically for optical interferometry [111, 166, 178]. This section

Table 6.1: Ellipse fitting algorithms included in the comparison.

Algorithm	Year	Abbreviation	Reference
Least-squares	1981	LS	[87]
Bias-corrected direct ellipse fit	2008	DEF	[175]
Best linear unbiased estimator minimisation	2014	BLUE	[111]
Orthogonal distance regression	1989	ODR	[179]

aims to review a number of ellipse algorithms that form a representative sample of those proposed in the literature.

Algorithm selection

For the purposes of this work the algorithms given in table 6.1 will be considered. Due to the large number of algorithms proposed in the literature not all possible algorithms are compared, with the included algorithms selected based on promising performance in the field of precision interferometry. In particular, algorithms designed for application to particularly noisy data, for example [176], were not considered due to the low noise requirements in the quadrature signals where sub-nanometre performance is required. The dynamic algorithm proposed in [166] was also excluded for the comparison, despite being developed specifically for optical interferometry, due to the assumption that the interferometric phase varies smoothly, which, depending on the spatial sampling rate and displacement profile, may not be the case.

The Least Squares (LS), Direct Ellipse Fit (DEF) and Best Linear Unbiased Estimator (BLUE) algorithms were selected as these algorithms have previously been applied to optical interferometry in [21], [163] and [111]. The DEF algorithm does not provide a method for uncertainty propagation, however, this algorithm was still of interest as it guarantees the fitted conic section is an ellipse or circle. If an ellipse specific fit was found to have benefits, it is possible that additional work could yield a method for uncertainty propagation.

Considering algorithms that have not yet been applied to interferometry in the literature, only algorithms that account for non-homogeneous and non-isotropic noise in the fitted data were considered. In homodyne interferometry, as was demonstrated in chapter 3, the noise in the quadrature signals will be strongly correlated, either due to optical intensity noise, or due to noise in the intensity noise normalisation process being coupled into both quadrature signals, leading to non-isotropic noise, and the noise will also vary with position on the Lissajous figure, resulting in non-homogeneous noise. As a result of the strong noise correlation, accounting for the covariance in the fitting process is important both for uncertainty propagation, and for the fitting process, where the non-isotropic nature of the covariant noise may be used by some algorithms to improve the fit accuracy. For this reason the ODRPACK implementation of the Orthogonal Distance Regression (ODR) algorithm [179] was included in the comparison. The Approximate Maximum Likelihood (AML) implementation of Szpak [177] was also considered, however the algorithm

was found to occasionally diverge, and preliminary testing showed no improvements over the ODR method, so the AML algorithm was excluded from the comparison.

Algorithm implementation

Of the selected algorithms, the LS [87] fit has the simplest implementation, and is the most widely employed in homodyne interferometry. This algorithm was implemented as previously described in section 3.3, fitting a conic section of the form

$$Ax^2 + By^2 + Cxy + Dx + Ey = 1, \quad (6.1)$$

with covariances estimated using equation 3.48. The LS algorithm does not account for uncertainties in the data points.

The DEF [175] fit aims to improve upon the LS fit by guaranteeing that the fitted conic section is an ellipse or circle, rather than a parabola or hyperbola, with the implementation employed here also correcting for bias in the fit present in the original implementation [174]. This algorithm was implemented as described in [175] for an ellipse specific fit. This implementation of the DEF algorithm fits a conic section of the form

$$A'x^2 + B'xy + C'y^2 + D'x + E'y + F' = 0, \quad (6.2)$$

and for the purposes of this work the coefficients used in the uncertainty propagation equations derived in chapter 3 were calculated from those used in equation 6.2 as

$$A = -\frac{A'}{F'}, \quad (6.3)$$

$$B = -\frac{C'}{F'}, \quad (6.4)$$

$$C = -\frac{B'}{F'}, \quad (6.5)$$

$$D = -\frac{D'}{F'}, \quad (6.6)$$

$$E = -\frac{E'}{F'}. \quad (6.7)$$

No covariance estimation is provided by the DEF algorithm, and the algorithm does not account for uncertainties in the data points.

Both the LS and DEF algorithms attempt to minimise the sum of the squared algebraic distances between the data points to be fitted and the fitted curve, that is for a conic section given by equation 6.1 the algorithms minimise the quantity given by

$$\sum_i (Ax_i^2 + By_i^2 + Cx_iy_i + Dx_i + Ey_i - 1)^2, \quad (6.8)$$

where the sum is over all data points (x_i, y_i) included in the fit.

The BLUE [111] and ODR [179] algorithms instead aim to minimise the geometric distance from each data point to the fitted curve, although in the case of the BLUE and AML algorithms this process is approximate. The geometric distance

is the distance between a data point and the closest point on the fitted curve, as opposed to the point on the curve defined by equation 6.1 corresponding to the same x and y coordinates.

The BLUE algorithm was modified from the form given in [111] to fit a conic section of the form given by equation 6.1, as opposed to the form

$$x^2 + B^\dagger y^2 + C^\dagger xy + D^\dagger x + F^\dagger y + G^\dagger = 0 \quad (6.9)$$

used in [111]. The modification to the implementation of [111] required only straight forward changes to the derivative matrices used in the BLUE algorithm, and is not anticipated to significantly effect the performance of the algorithm.

The ODRPACK [179] implementation of the ODR algorithm was used to fit of equation 6.1 directly, using the implicit fit capability of ODRPACK, with the LS algorithm used to generate the initial fit coefficients. It should be noted that using the ODR algorithm in this way does not guarantee an elliptical fit, however this was not found to be a severe limitation, as in cases where a parabola or hyperbola resulted from the fitting process, the ellipses fitted by algorithms that guarantee an elliptical result were such poor fits as to be unusable for precision interferometry. The ODR algorithm can also make use of covariance matrices in each data point, and again these were provided. An estimated covariance matrix for the fit coefficients is produced by the algorithm.

Monte-Carlo simulation procedure

In order to compare the algorithms presented in table 6.1 Monte-Carlo simulations were carried out following the measurement model described in section 3.3. Optical power signals were generated for the quadrature and reference outputs using equations 3.100 to 3.102 at N phase points for an interferometer with an operating wavelength of 632.8 nm, an intensity pick off ratio of 10 %, $\mathfrak{R} = \mathfrak{T} = 0.3$, and initially a 90° quadrature phase shift and an input power of 1 mW.

These optical powers were then used to generate shot noise limited voltage signals according to equations 3.76 to 3.78, 3.79 and 3.80 for photodiode detectors as described in table 3.2, with a bandwidth of 100 kHz. For the purposes of the Monte-Carlo simulation, random noise was generated with standard deviations as described by equations 3.79 to 3.80 for M Monte-Carlo repeats. Normalised quadrature signals were then calculated from the voltage signals according to equations 3.33 and 3.34 for each repeat, with the signal gains set to give a peak voltage of unity.

In all cases signals were generated for interferometric phases in the range $[0, 2\pi)$ (representing the half open interval $0 \leq \varphi < 2\pi$) with $N = 1000$ phase steps (equivalent to a 158 pm displacement spacing for a double pass interferometer operating at 632.8 nm) and $M = 10000$ Monte-Carlo repeats. For each repeat, the desired number of fit points uniformly distributed throughout the desired fit range were selected from the N phase points, with a randomly generated offset, different for each repeat, in the range $[0, N)$ applied such that any effects introduced by the location of the selected phase points within the fringe were averaged out. Where this process resulted in phase point indexes greater than N , the indexes were rolled over to begin again from zero. This approach was chosen as for simulations considering partial fringes

the performance of the algorithms may be dependent on the location of the data within a fringe, something that cannot generally be freely chosen in experimental work. Additionally, the uncertainty in the change in the interferometric phase for a given displacement length is dependent on the initial value of the instantaneous phase, which again cannot be freely chosen in experimental work.

Each ellipse fitting algorithm was applied to the selected data points separately for each of the M repeats, and the interferometric phase φ_{calc} and the change in the interferometric phase $\Delta\varphi_{\text{calc}}$ were calculated from the resulting fit coefficients using equations 3.29 and 3.31 for the full range of phase points. Uncertainties in the resulting changes in the interferometric phase were estimated as

$$\sigma_{\Delta\varphi} = \sqrt{\frac{1}{M-1} \sum_{m=0}^M (\Delta\varphi_{\text{calc},m} - \Delta\varphi_m)^2}, \quad (6.10)$$

where the known value of the change in the interferometric phase is used rather than the average of the calculated phases so as to account for both noise and residual non-linearities in the calculated interferometric phase. A 95 % confidence interval was calculated for each estimate of $\Delta\sigma_{\varphi}$ using the chi squared distribution, with the confidence interval given by [180]

$$\left(\sqrt{\frac{M-1}{\chi_{0.025}^2}} \sigma_{\Delta\varphi}, \sqrt{\frac{M-1}{\chi_{0.975}^2}} \sigma_{\Delta\varphi} \right), \quad (6.11)$$

where χ_{α}^2 is the inverse of the cumulative distribution function for the chi squared distribution evaluated for $M-1$ degrees of freedom at a probability α . Such confidence intervals are shown by the error bars throughout this section.

For ease of interpretation and comparison, the interferometric phase uncertainty estimates and corresponding confidence intervals were converted to displacement uncertainties using equation 3.103 with zero dead path, $N_{\text{pass}} = 4$ and a constant wavelength of 632.8 nm assumed, neglecting all other uncertainty sources. For the purposes of residual non-linearity comparison, the maximum estimated displacement uncertainty within the fringe was taken for comparison.

To assess the accuracy of the covariance estimates provided by the algorithms, equation B.2 was used to calculate uncertainty estimates in the interferometric phase, with the estimates provided by the algorithms used for the fit coefficient variances and covariances. Variances and covariances in the normalised quadrature signals were estimated directly from the Monte-Carlo repeats. In addition to the fit coefficient variance and covariance estimates provided by the algorithms, covariance matrices in the fit coefficients were also calculated directly from the Monte-Carlo process to allow to validity of equation B.2 to be tested.

For comparison, uncertainties were also calculated using the analytic equations derived in section 3.3, excluding the effects of uncertainties in the fit parameters. Analytic variances were also calculated with a randomly generated offset in the range $[0, N)$ applied for $M = 10000$ repeats, again in order to average the effect of varying the starting and ending instantaneous phase. The calculated analytic variances were

then averaged across the M repeats, and the standard uncertainty calculated from the square root of the variance. Any increase above the uncertainty limit indicated by the analytic solution shows the increase in measurement uncertainty introduced by the uncertainty in the ellipse fit coefficients.

Residual non-linearity comparison

First, the performance of the algorithms when fitting a full ellipse was examined. The described Monte-Carlo procedure was used to generate quadrature signals with a varying number of phase points. The results of the Monte-Carlo simulation are shown in figure 6.1. The performance of the LS, DEF and BLUE algorithms can be seen to be approximately equal, with any differences within the uncertainty introduced by the relatively low number of repeats included in the Monte-Carlo simulation as a result of the computational complexity of the geometric fit algorithms. A small improvement in performance can be seen for the ODR algorithm, with the benefits increasing where a smaller number of points are included in the fit. This process was repeated for a higher noise level by reducing the illuminating power used in the Monte-Carlo simulation to 0.1 mW, with a similar result, as shown in figure 6.2. The performance differences between the algorithms can be seen to decrease as the number of points included in the fit increases, and as such the simulation was not continued past 100 fit points.

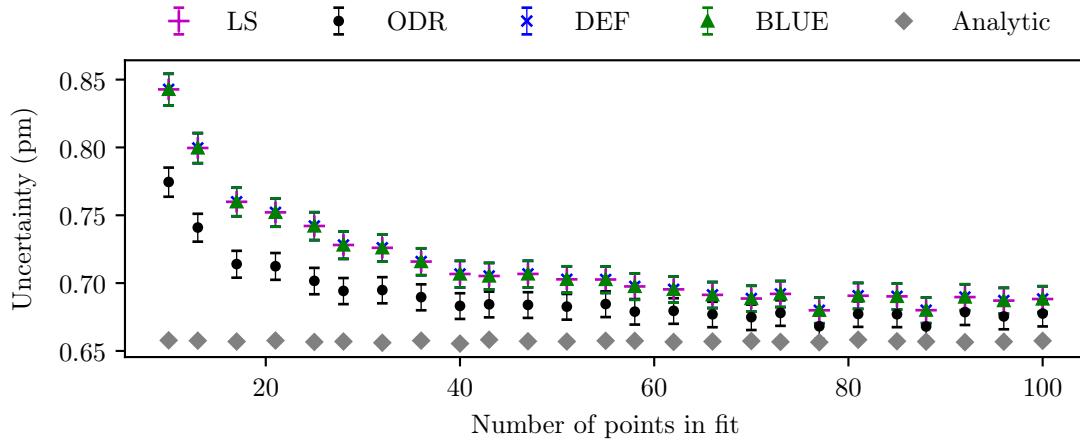


Figure 6.1: Estimated displacement uncertainty for varying numbers of fit points after applying the algorithms given in table 6.1 evaluated at the maximum error point within a fringe for an illuminating power of 1 mW, with noise in the quadrature signals simulated as described in the text. Error bars show a 95 % confidence interval.

The sensitivity of the algorithms to varying quadrature phase shift and phase quadrature coating reflectivity to transmission ratio was also investigated for the full fringe case with 100 phase points included in the fit. As shown in figures 6.3 and 6.4, the results are in agreement with figures 3.8 and 3.10 with a small increase in uncertainty due to the uncertainty in the fit coefficients. No significant differences between the algorithms are visible, and the locations of the optimum coating

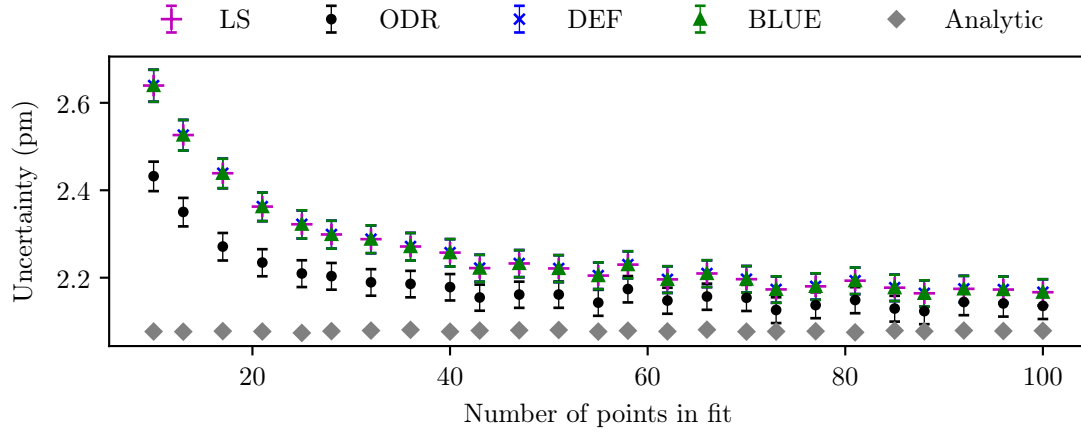


Figure 6.2: Estimated displacement uncertainty for varying numbers of fit points after applying the algorithms given in table 6.1 evaluated at the maximum error point within a fringe for an illuminating power of 0.1 mW, with noise in the quadrature signals simulated as described in the text. Error bars show a 95 % confidence interval.

properties are not changed.

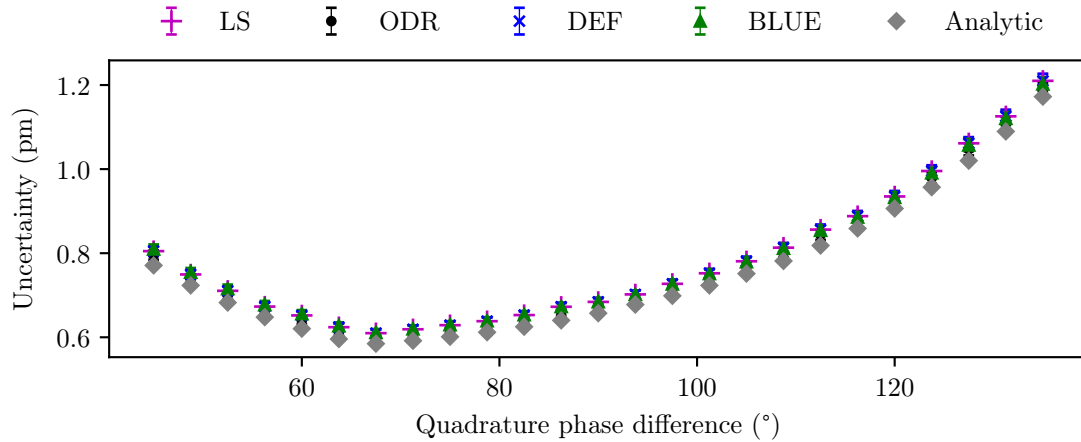


Figure 6.3: Estimated displacement uncertainty for varying quadrature phase shifts after applying the algorithms given in table 6.1 evaluated at the maximum error point within a fringe for an illuminating power of 1 mW, with noise in the quadrature signals simulated as described in the text. Error bars show a 95 % confidence interval.

The algorithms were then tested on data covering less than a full ellipse, for a constant phase spacing of 250 steps per fringe, equivalent to a 0.633 nm displacement spacing for a double pass interferometer at 632.8 nm. This test was again repeated for two illuminating power levels (1 mW, shown in figure 6.5 and 0.1 mW, shown in figure 6.6), corresponding to two noise levels. In both cases the residual

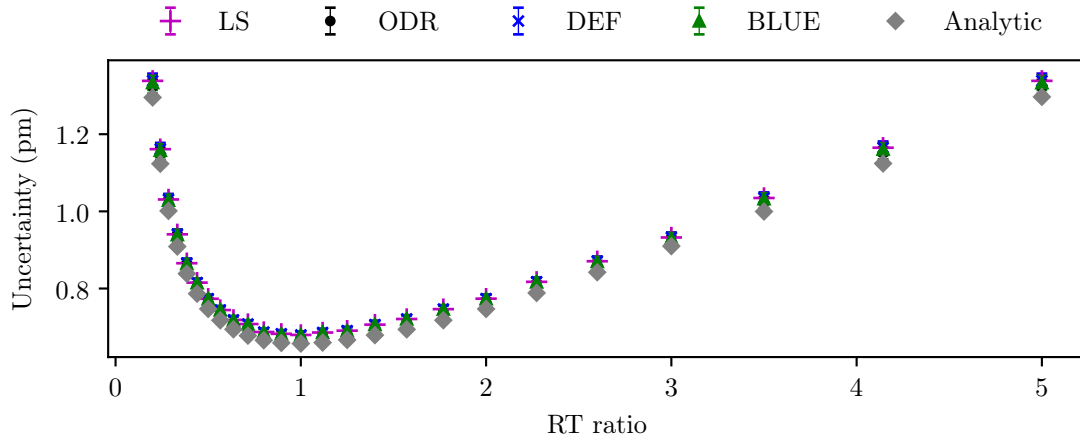


Figure 6.4: Estimated displacement uncertainty for varying phase quadrature coating reflectivity to transmission ratio after applying the algorithms given in table 6.1 evaluated at the maximum error point within a fringe for an illuminating power of 1 mW, with noise in the quadrature signals simulated as described in the text. Reflectivity to transmission ratios are shown for a fixed coating absorption of 40 %. Error bars show a 95 % confidence interval.

displacement errors can be seen to increase as the fit range is reduced for all algorithms. Although the performance of the algorithms is broadly comparable, some differences are apparent. For small fringe fractions, corresponding to displacement ranges of less than approximately 50 nm for a double pass interferometer operating at 632.8 nm, the benefits of the guaranteed ellipse fit provided by the DEF algorithm can be seen, with this being the only algorithm to return a valid fit in some cases. The elliptical nature of the fitted conic sections was not checked directly, however, in cases where any of the repeats returned fit coefficients resulting in a negative square root value in equation 3.29 no data is plotted, as the fit was considered invalid. The BLUE algorithm can also be seen to outperform the LS and ODR algorithms for small fringe fractions, although this algorithm also failed to return elliptical fits for extremely small fringe fractions. For displacement ranges greater than 100 nm the ODR algorithm was found to give a small performance improvement over the other algorithms under the conditions tested here, although the improvement in performance was negligible for displacement ranges approaching a full optical fringe, as would be expected from figures 6.1 and 6.2. This may be more clearly seen in figure 6.7, where the simulation has been restricted to fringe fractions corresponding to displacements greater than 50 nm, and a linear ordinate axis scale has been employed.

The previous tests of varying quadrature phase and phase quadrature coating reflectivity to transmission ratio shown in figures 6.3 and 6.4 were also repeated for one half of an optical fringe with 100 data points included in the fit, the results of which are shown in figures 6.8 and 6.9, where the benefits of the ODR algorithm can be clearly seen, as would be expected from figure 6.7. As in the full fringe case, there

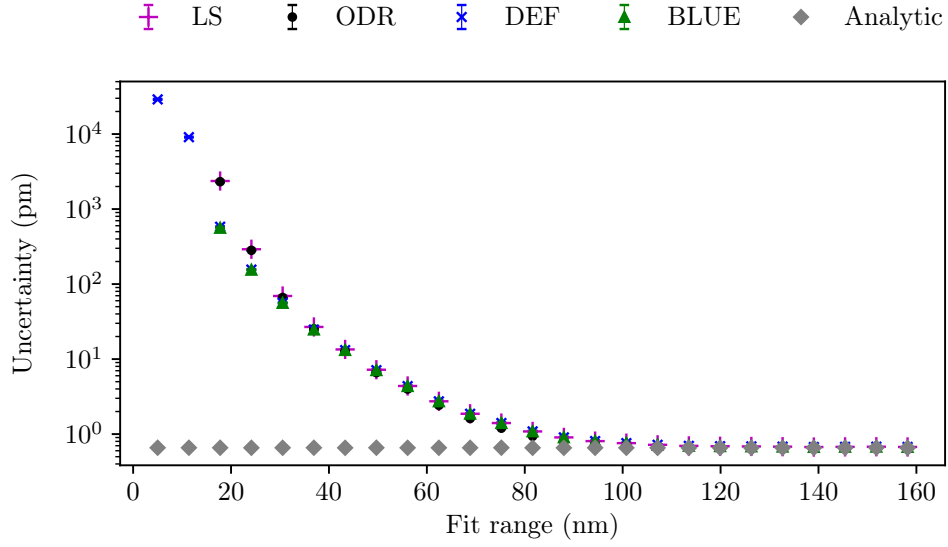


Figure 6.5: Estimated displacement uncertainty after applying the algorithms given in table 6.1 evaluated at the maximum error point within a fringe for an illuminating power of 1 mW for sub-fringe displacements, with noise in the quadrature signals simulated as described in the text. Error bars show a 95 % confidence interval.

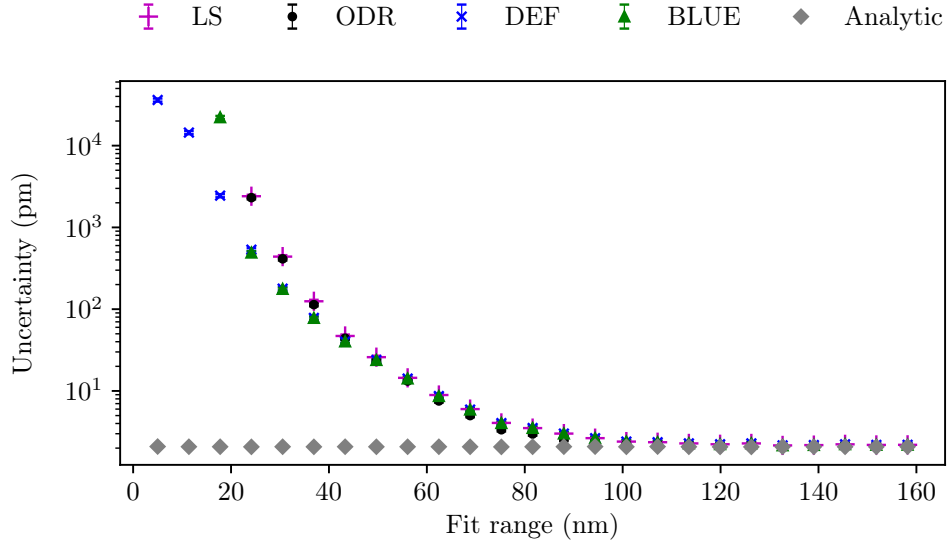


Figure 6.6: Estimated displacement uncertainty after applying the algorithms given in table 6.1 evaluated at the maximum error point within a fringe for an illuminating power of 0.1 mW for sub-fringe displacements, with noise in the quadrature signals simulated as described in the text. Error bars show a 95 % confidence interval.

is no change in the optimal value of the quadrature phase shift, or the reflectivity to transmission ratio, relative to the idealised analytic solution.

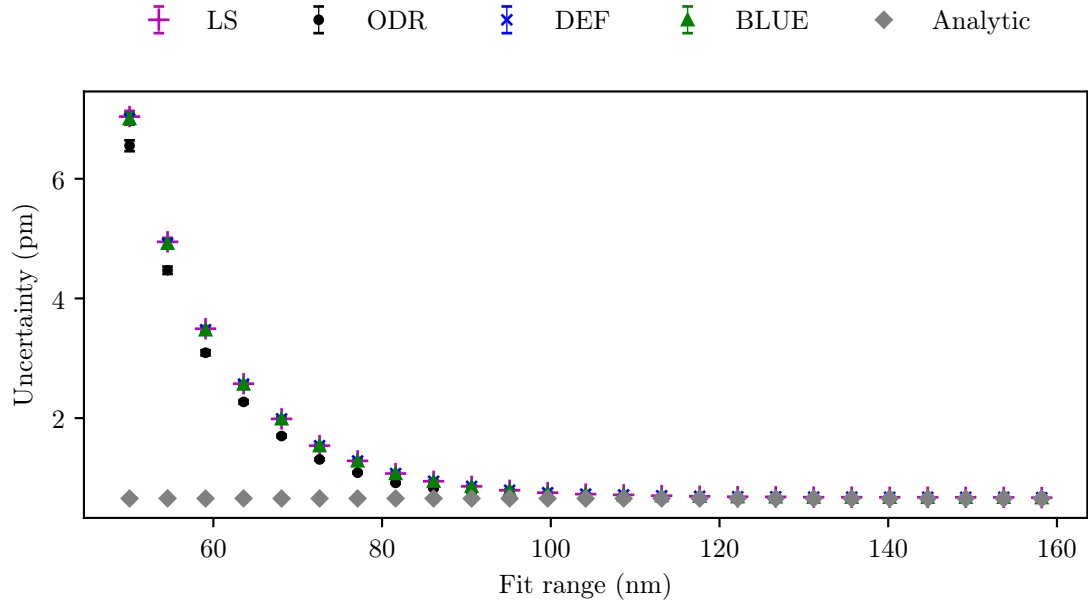


Figure 6.7: Repeat of the simulation shown in figure 6.5, restricted to larger fit ranges, and with a non-logarithmic scale. Error bars show a 95 % confidence interval.

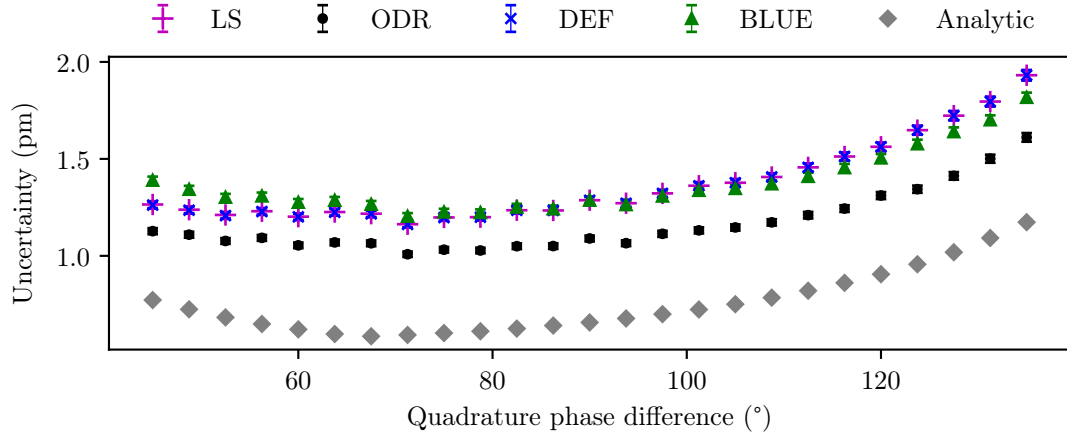


Figure 6.8: Average magnitude of displacement errors for varying quadrature phase shifts when fitting one half of an optical fringe after applying the algorithms given in table 6.1 evaluated at the maximum error point within a fringe for an illuminating power of 1 mW, with noise in the quadrature signals simulated as described in the text. Error bars show the standard error in the mean.

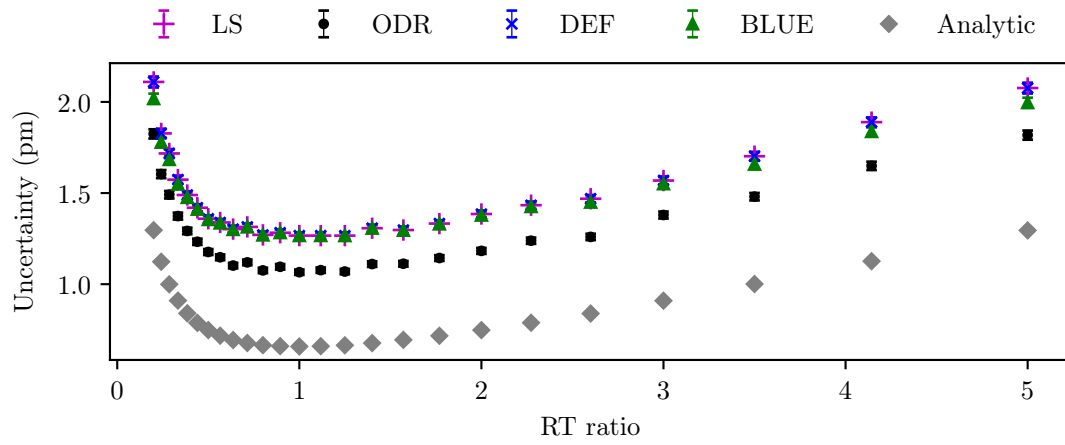


Figure 6.9: Average magnitude of displacement errors for varying phase quadrature coating reflectivity to transmission ratio when fitting one half of an optical fringe after applying the algorithms given in table 6.1 evaluated at the maximum error point within a fringe for an illuminating power of 1 mW, with noise in the quadrature signals simulated as described in the text. Error bars show the standard error in the mean.

Uncertainty propagation comparison

In addition to the capabilities of an ellipse fitting algorithm to minimise residual non-linearities, the ability of an algorithm to estimate a covariance matrix may be of interest, if fit coefficient uncertainties are to be used to estimate displacement uncertainties. For the algorithms compared in this work, established methods exist to estimate the fit coefficient covariance matrix for the LS, ODR and BLUE algorithms. No covariance estimate is provided for the DEF algorithm, and as such this algorithm will not be included in this part of the comparison.

In order to compare the accuracy of the covariance estimates provided by the algorithms, Monte-Carlo simulations were again employed using the previously described methodology, comparing uncertainties calculated directly from the Monte-Carlo results ('Monte-Carlo'), uncertainties calculated from the covariances produced by the algorithms ('Algorithm covariance'), and uncertainties calculated from equation B.2 using covariances in the fit coefficients estimated from the Monte-Carlo repeat results ('Monte-Carlo covariance'). Uncertainties were again also calculated with the analytic expression derived in equations 3.3 to indicate the baseline uncertainty that would arise if no additional uncertainty was introduced by the correction process.

First, the uncertainty propagation properties of the algorithms were tested on a full fringe. Results for fits including 100 data points are shown in figure 6.10, where all algorithms can be seen to be in reasonable agreement with the Monte-Carlo result, although some deviations around the minima and maxima can be seen. Reducing the number of points included in the fit to 20 demonstrates the greater accuracy of the covariance estimates produced by the ODR algorithm as compared to the LS and BLUE algorithms, as can be seen in figure 6.11, where the algorithm covariance and Monte-Carlo covariance lines can be seen to be in agreement for the ODR algorithm. All algorithms, however, fail to match the Monte-Carlo result, even when the fit coefficient covariance estimate is accurate, as in the ODR case.

This finding suggests that the assumption that the uncertainties in individual instantaneous phase measurements are uncorrelated, made in section 3.3, is not accurate when the ellipse fit coefficients contribute significantly to the measurement uncertainty. Errors in the fit coefficients, however, contribute only to the non-linearity error, not to the statistical uncertainty, as no additional 'noise' is added by the use of erroneous fit coefficients. As a result, the uncertainty budget presented in section 3.3 remains accurate if, as suggested in section 3.3, experimental non-linearity measurements are made under the same conditions, including the number of points in the ellipse fit, as the displacement measurement.

The results of applying the algorithms to one half of a full fringe with 100 data points included in the fits are shown in figure 6.12. In all cases the algorithms underestimate the uncertainty. Even larger underestimates are seen in figure 6.13, where the results of repeating the test with one quarter of a full optical fringe are shown. This further validates the statement that it is preferable to measure residual non-linearities experimentally, rather than to rely upon uncertainties estimated based upon fit coefficient uncertainties. Making use of the Monte-Carlo estimates of the fit coefficient uncertainties provides a better approximation, as least of the maximum uncertainty value, however, significant deviations from the full Monte-Carlo

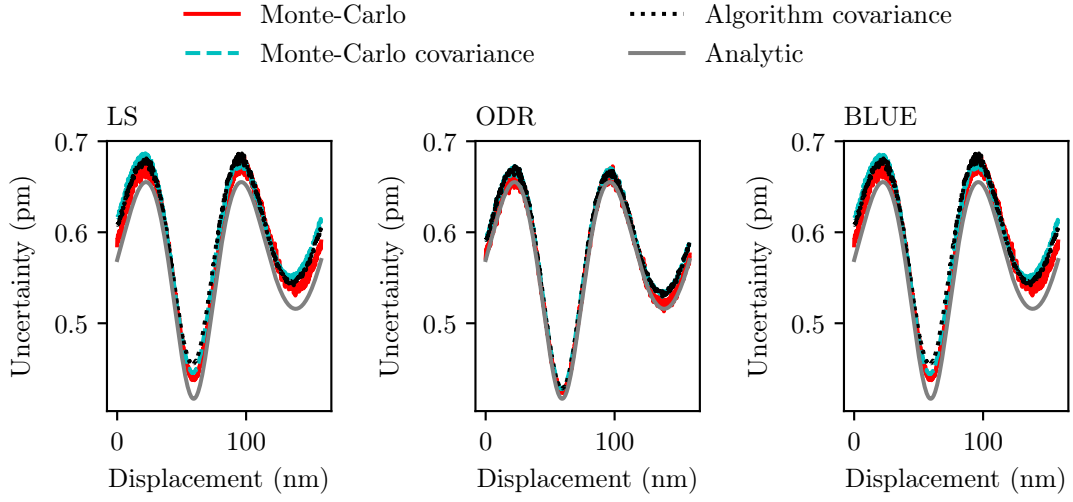


Figure 6.10: Displacement uncertainties calculated with the described Monte-Carlo methodology for the LS, ODR and BLUE algorithms for full fringe fits including 100 data points.

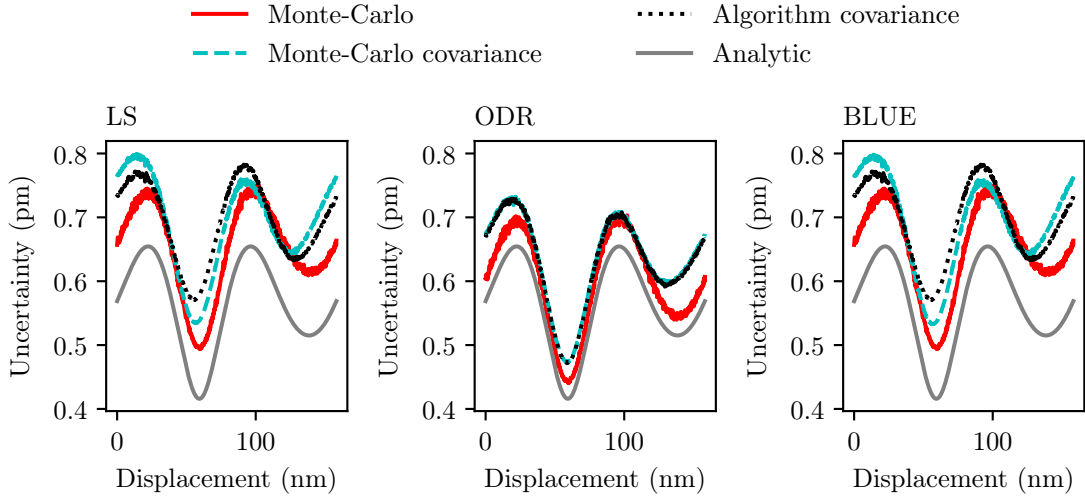


Figure 6.11: Displacement uncertainties calculated with the described Monte-Carlo methodology for the LS, ODR and BLUE algorithms for full fringe fits including 20 data points.

case are still present. These deviations are again attributed to the assumption that the uncertainties in individual instantaneous phase measurements are uncorrelated.

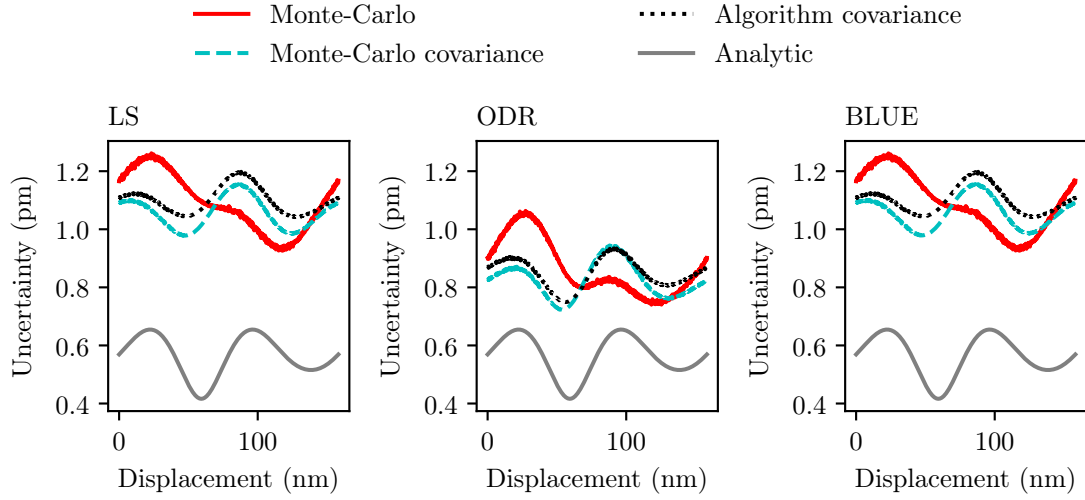


Figure 6.12: Displacement uncertainties calculated with the described Monte-Carlo methodology for the LS, ODR and BLUE algorithms for half fringe fits including 100 data points.

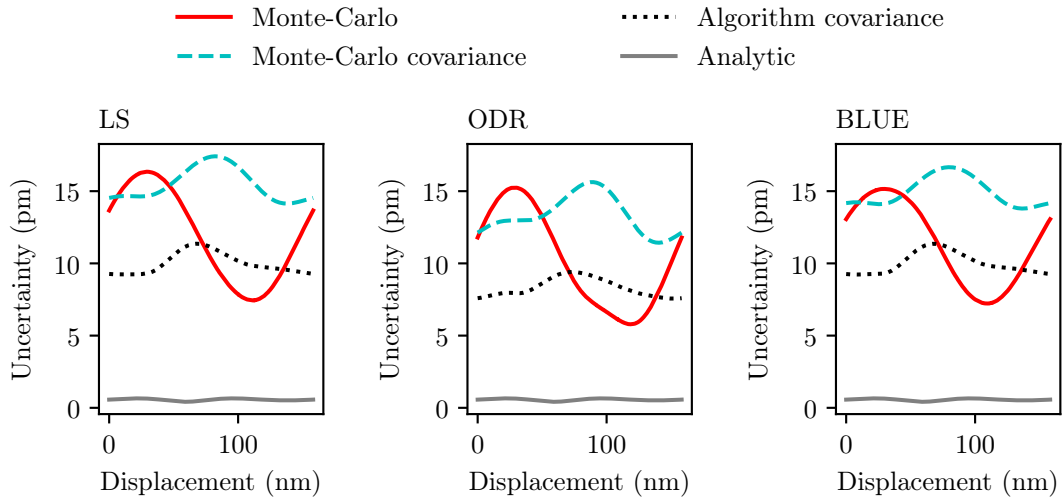


Figure 6.13: Displacement uncertainties calculated with the described Monte-Carlo methodology for the LS, ODR and BLUE algorithms for quarter fringe fits including 100 data points.

Summary of ellipse fitting algorithms

The results of this comparison suggest that, for the algorithms given in table 6.1, where a full ellipse is to be fitted, with a large number of data points included in the fit relative to the noise level, any differences between the performance of the ellipse

fitting algorithms are negligible, as indicated by the trends shown in figures 6.1 and 6.2. This is in agreement with the findings of [110]. Additional uncertainty in the displacement measurement is introduced by application of ellipse fitting non-linearity corrections due to the uncertainties in the fit coefficients, with the additional uncertainty introduced being reduced by including a greater number of data points in the fit, an effect that can be observed as the deviation from the analytic solution shown by all algorithms in figures 6.1 and 6.2. Beyond the uncertainty scaling effects introduced by the stretching effects of the ellipse fitting non-linearity correction, figures 6.3 and 6.4 demonstrate that no significant additional uncertainties are introduced when correcting data with non-ideal quadrature phase shifts or phase quadrature coating reflectivity to transmission ratios.

Differences between the algorithms begin to appear as the number of points included in the fit is reduced, with figures 6.1 and 6.2 showing that the ODR algorithm becomes preferable in this case. Differences also arise when displacements covering less than a full optical fringe (corresponding to a full ellipse) are to be corrected. For displacement ranges of approximately 50 nm and upwards the ODR algorithm again has small benefits over the other algorithms considered, shown most clearly in figure 6.7. As shown in figures 6.5 and 6.6, for smaller displacement ranges, the DEF and BLUE algorithms out perform the LS and ODR algorithms. For very small fringe fractions the DEF algorithm was the only algorithm tested to consistently return valid ellipse fits, however, in this regime the ellipse fits returned result in non-linearities of hundreds of picometres, which in many cases may be impractically large for precision interferometry applications.

Of the algorithms that provide variance and covariance estimates in the fit coefficients (the LS, BLUE and ODR algorithms), the ODR algorithm was found to produce the most accurate variance and covariance estimates, as shown in particular by figure 6.11, although with large numbers of points included in the fit the differences between the algorithms again become minor. For displacements covering less than a full optical fringe all algorithms were found to significantly underestimate the fit coefficient variances and covariances, resulting in the significant underestimates of the maximum displacement uncertainty shown in figures 6.12 and 6.13. Furthermore, the previously stated assumption that uncertainties in individual measurements of the instantaneous phase was found to be inaccurate where significant uncertainty is contributed by the fit coefficients. When estimating measurement uncertainties following the procedure described in section 3.3 this is not a problem, as this additional uncertainty is accounted for in the non-linearity measurement. If uncertainties in the fit coefficients were to be used for uncertainty estimation, however, an expression for the uncertainty in the measured phase difference, rather than the instantaneous phase value as presented in appendix B, accounting for the covariance between the fit coefficients would need to be derived.

The ODR algorithm was found to be preferable in all cases involving displacements greater than 50 nm in this comparison, with the small fringe fraction performance likely to be limited by the non-ellipse specific nature of the fit in the implementation applied in this work. Further work may develop an ellipse specific fit using the ODR algorithm, potentially extending the performance benefits of this algorithm to the small fringe fraction case. The computational complexity of

the algorithms has not been considered in this comparison, as for post processing corrections none of the algorithms presented in table 6.1 present a computational challenge on modern hardware. Where the algorithms are to be employed as a pseudo-real time correction (discussed briefly in the following section) this may give an advantage to the simpler LS and DEF algorithms, with previous work in the field indicating the DEF algorithm is preferable [163] with a minimal performance penalty.

6.2.2 Non-ellipse fitting algorithms

In addition to the ellipse fitting algorithms discussed in the previous section, non-linearity correction algorithms have been proposed that do not rely upon ellipse fitting. Such algorithms are in general more diverse than the ellipse fitting techniques previously considered, and are in many cases developed to perform in a specific measurement regime, for example at high speed, or for very small sub-fringe displacements. A meaningful direct comparison is therefore more challenging, and this section will aim to give a general overview of the non-ellipse fitting algorithms proposed in the literature, without making a detailed comparison of the performance limits of the algorithms.

For applications where real time updates to the non-linearity correction are required, for example where the interferometer displacement is to be used as a control signal, algorithms have been proposed that are suitable for implementation on a Field Programmable Gate Array (FPGA). The simplest such algorithm is a Peak Finding (PF) algorithm developed in [167] that detects maxima and minima of both quadrature signals and uses this information to normalise the quadrature signals to the range ± 1 . The sum and difference of the normalised signals are then taken to correct for any quadrature phase error, and the peak finding and normalisation process is repeated to renormalise the sum and difference signals. The interferometric phase is then calculated with equation 3.28. In the published implementation, this correction was applied with update rates of up to 10 MHz.

Another algorithm [168] designed for real time applications, which will be termed the Low Error Point (LEP) algorithm in this work, exploits the fact that at interferometric phases of

$$\varphi = \frac{m\pi}{4}, \quad (6.12)$$

for $m = 0, 1, \dots, 7$ the errors due to elliptical distortions to the Lissajous figure are minimised, with the error due to the offset in the x quadrature signal minimised at $m = 0$ and 4, the error due to the offset in the y quadrature signal minimised at $m = 2$ and 6, the error due to a gain ratio error minimised at $m = 0, 2, 4$ and 6, and the error due to quadrature phase shift error minimised at $m = 1, 3, 5$ and 7. The phase at the $\pi/4$ points therefore approximates the error free phase, and the quadrature signal values at the $\pi/4$ points can be used to recover the offsets, gain ratio and quadrature phase error required to correct the quadrature signals. Where multiple error sources occur, for example both an offset error and a gain ratio error, the phase error minimisation at the $\pi/4$ points is only approximate, however, if the algorithm is applied iteratively, with the quadrature signal values at the $\pi/4$ points updated each time a $\pi/4$ point is reached, the algorithm will approach the correct

values.

A final real time algorithm, the Kalman Filter (KF) algorithm, aims to estimate the conic section coefficients produced by ellipse fitting algorithms with an extended Kalman filter algorithm [170], a technique that is described as more suitable for real time implementation than the matrix based ellipse fitting techniques described previously. This approach has been shown to be capable of updating the correction coefficients at 1 MHz whilst applying the non-linearity correction at 50 MHz.

The most important aspect of the algorithms to be considered is the residual non-linearity after correction. In all cases values are provided by the authors, however, as the performance of the algorithms is likely to be dependent on the optical interferometer they are paired with, a direct comparison should be treated with caution. Nonetheless, for reference, the PF algorithm was shown to reduce peak-to-peak errors from 40 nm to 0.6 nm [167], the LEP algorithm from 4 nm to 1 nm [168] and the KF algorithm from 40 nm to 4 nm [170].

An additional aspect that must be considered is the fact that all three real time algorithms considered here will have some convergence displacement; the displacement range that must be covered for the ultimate performance of the algorithm to be reached. For the PF algorithm the authors indicate approximately 1.8 fringes are required for convergence [167], the LEP algorithm requires approximately 2.6 fringes [168] and the KF algorithm requires approximately one fringe [170]. In all cases the required convergence displacement may be dependent on the noise in the quadrature signals, although this has only been demonstrated explicitly for the KF approach [170].

Considering the possible performance limits of the algorithms, the PF [167] and LEP [168] algorithms both make use of the value of the quadrature signals at a relatively small number of points to make the correction. As such, they will be more sensitive to noise than the ellipse fitting approach, as the ellipse fitting approach effectively averages the noise around the ellipse when making the fit. The KF approach [170] does not suffer this limitation, and may therefore be more suitable in cases where high speed correction is required for relatively noisy quadrature signals. A related limitation of the PF [167] and LEP [168] algorithms is the required sampling rate. As data from fixed points of the ellipse is used to apply the correction, in cases where few data points are recorded for each fringe, which may be the case for example where high velocities must be measured, the accuracy of the algorithms may be severely impacted. The KF approach [170] is again anticipated to be less sensitive to this effect.

Comparing the three algorithms, the PF algorithm has significant advantages in ease of implementation, and due to the relatively simplicity of the algorithm it may be possible to implement the algorithm on lower cost hardware, or alternatively at higher speeds on equivalent hardware. The experimental data presented by the authors also indicates the algorithm may be capable of greater non-linearity error reductions than the LEP algorithm, although without comparing both algorithms on the same optical interferometer system this cannot be stated for certain. The KF algorithm is significantly more complex, and as such may require higher cost hardware. The experimental data presented in [170] demonstrates a lower performance than

the experimental implementation of the PF algorithm, however the performance appears to be limited primarily by noise, and as such may be improved for a lower noise system. In particular, it seems likely the KF algorithm may have significant advantages where relatively few data points are available for each optical fringe. Further work on the KF algorithm is required to fully explore the performance limits.

Another possible pseudo-real time correction methodology is to apply one of the previously considered ellipse fitting algorithms at a lower frequency, with the ellipse fit being performed for example on a microprocessor, whilst applying non-linearity correction through equation 3.29 at the full measurement frequency, for example on a FPGA. In cases where the ellipse fit coefficients are relatively stable, this approach may be sufficient for real time measurement.

Two non-ellipse fitting algorithms have not been considered in this section. The first relies upon linearisation of the measured displacement to a second displacement measurement made with a capacitive sensor [68] in a manor analogous to that employed by the two-wavelength non-linearity correction proposed in section 6.3. The second uses shutters in the measurement and reference arms to normalise the Lissajous signal to a unit circle [169]. A further development of this approach is presented in section 6.4. The benefits and limitations of these corrections are discussed in sections 6.3 and 6.4.

6.3 Two-wavelength non-linearity correction

6.3.1 Introduction

Unwanted cavity formation within interferometer optics, either due to polarisation leakage or unwanted back reflections (ghost reflections) has been identified as a source of non-linearities in a number of homodyne and heterodyne interferometer designs [60,61,63,65,84,114,115]. Whilst analyses of such non-linearity sources have been made [59,60,84,85,114,115], and in some cases the effects of these non-linearities can be reduced through improved optical design or alignment [28,115,181], complete elimination is often impossible. As a result, residual non-linearity errors that are highly dependant on the exact alignment of the interferometer optics, and that range from tens to hundreds of picometres in magnitude, may remain in interferometer systems

As previously discussed in this chapter, a wide range of non-linearity correction algorithms have been proposed in the literature, many following the ellipse fitting technique developed by Heydemann [87] for ultrasonic instruments, and first applied to optical interferometry by Birch [21]. A number of improvements to the Heydemann correction have been proposed, applying differing ellipse fitting algorithms to the problem, as discussed in section 6.2.1, however all are ultimately limited to the correction of elliptical non-linearities, and as such are unable to correct for higher order non-linearities that result in a non-elliptical Lissajous figure [84,85].

Alternative non-linearity correction methodologies aiming to compensate only for quadrature phase differences, relative gains and offsets have also been proposed based on peak detection [167] and identification of low error points on the ellipse [168]. These approaches allow certain non-linearities to be corrected for in real time,

but are also limited to the correction of elliptical non-linearities.

One correction approach has been proposed in the literature [68] that is capable of the correction of non-elliptical non-linearities, based upon fitting a sum of sinusoidal non-linearity harmonics with variable phase and amplitude to the difference between a measured interferometric displacement and the same displacement measured with a capacitance sensor. This approach is limited to applications in which attaching a capacitance sensor to the target to be measured is possible, and to displacement ranges suitable for capacitive measurement, although the authors do suggest alternative secondary displacement measures. In addition to this, fixed amplitude and phase non-linearity terms are fitted in [68]. Higher order non-linearity terms are highly sensitive to the optical alignment, and therefore may vary with both displacement and time, limiting the performance of the capacitive correction where non-linearities vary on a fringe to fringe basis.

Two-wavelength interferometer designs have previously been developed for refractive index compensation [88, 182, 183], absolute length measurement [184], and non-linearity evaluation [11, 48]. The possibility of exploiting the periodic nature of non-linearity errors to evaluate residual non-linearities, by observing the magnitude of the difference between displacement measurements made at two wavelengths has previously been recognised [11, 48], however, the correction of non-linearities through two-wavelength interferometry has not previously been demonstrated.

In this section a non-linearity correction algorithm is proposed, implemented as a post-processing technique, that makes use of two coaxial optical interferometers operating along the same beam path but at differing wavelengths. The difference between the measured displacements is calculated, and, through Fourier analysis, non-linearity terms at spatial harmonics of each wavelength are separated. The non-linearity errors in each interferometer are then subtracted from the measured displacements, resulting in a corrected displacement measurement. In contrast to the commonly employed Heydemann style ellipse fitting correction, this technique allows for the correction of higher order non-linearity terms that produce non-sinusoidal quadrature signals and non-elliptical Lissajous figures. In this section the basic principle of the proposed correction algorithm is outlined, followed by a discussion of the data processing techniques required to minimise edge effects and spectral leakage arising from the Fourier approach. The correction algorithm is then demonstrated in a proof of concept experiment for a variety of displacement profiles. As this is a proof of concept work, potential improvements to the technique and required future work are discussed.

The work presented in this section has been previously published in the course of this PhD [69].

6.3.2 Theory and data processing methods

In order to describe the two-wavelength non-linearity correction methodology, a basic introduction to the principles of the technique will first be given, followed by a discussion of the limits of the Fourier approach and the data processing techniques required to mitigate these limits.

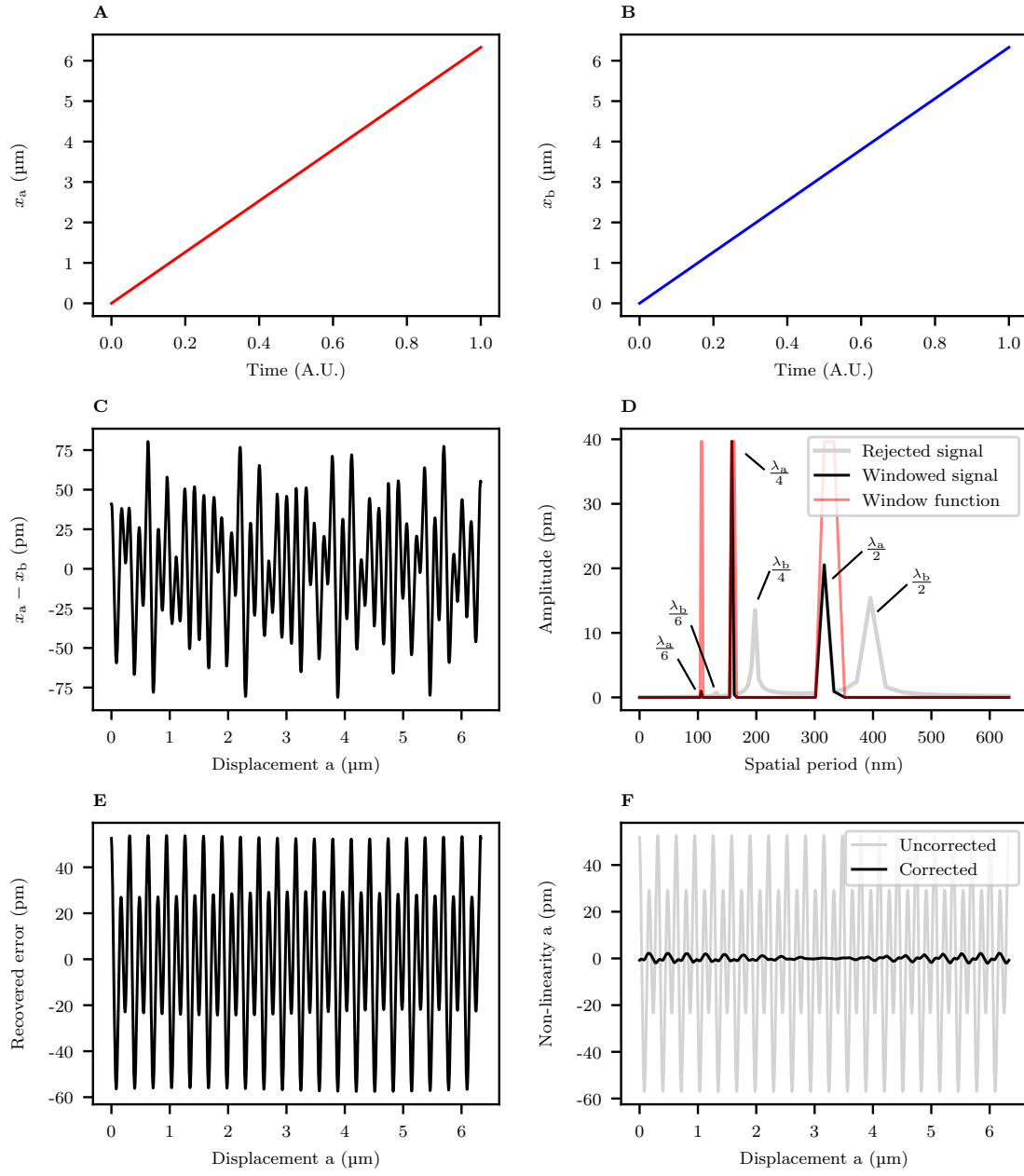


Figure 6.14: Basic principles of the two-wavelength non-linearity correction methodology, illustrated here for the simple case of a linear movement, as applied to modelled data for $\lambda_a = 633\text{nm}$ and $\lambda_b = 785\text{nm}$. **A** - displacement is measured with interferometer ‘a’, operating at λ_a . **B** - displacement is measured simultaneously with interferometer ‘b’, operating at λ_b . **C** - the difference between the ‘a’ and ‘b’ displacements is calculated. **D** - a Fourier transform is applied to the displacement difference and windowed around spatial harmonics of λ_a . **E** - estimated non-linearity errors in the λ_a displacement are recovered by inverting the windowed Fourier transform of the displacement difference. **F** - estimated non-linearity errors are subtracted from displacement a , resulting in a displacement measurement with reduced non-linearity. Edge effects are visible, resulting from a non-integer number of periods being included in the Fourier transform.

Principles of the two-wavelength non-linearity correction

Defining non-linearities as errors in the measured interferometric displacement that are periodic with harmonics of the illuminating wavelength, and neglecting all other error sources, the measured interferometric displacement at a wavelength λ_a may be written as

$$x_a = x + \sum_{j=1}^{j_{\max}} A_{a,j} \cos \left(4\pi \frac{j}{\lambda_a} x + \phi_{a,j} \right), \quad (6.13)$$

where x is the true displacement and $A_{a,j}$ and $\phi_{a,j}$ are amplitude and phase coefficients at the j^{th} harmonic respectively. An optical configuration resulting in only even harmonics being present in the measured displacement is assumed in this work. This is the case for the majority of displacement measuring interferometer designs, which operate almost universally in a two or four pass configuration, with one optical fringe corresponding to one half or one quarter of the illuminating wavelength in displacement.

The true displacement x cannot in general be recovered from x_a , as, although the error is known to be periodic, it is periodic with x , which is unknown. x could be recovered from x_a if the sample spacing was known to be uniform in x , however uniform sample spacing would require a constant velocity, and therefore severely limit the applicability of the technique.

In this two-wavelength interferometry technique displacement is measured simultaneously with two interferometers operating at different wavelengths along a common beam path. A second displacement measurement is therefore introduced, given by

$$x_b = x + \sum_{j=1}^{j_{\max}} A_{b,j} \cos \left(4\pi \frac{j}{\lambda_b} x + \phi_{b,j} \right). \quad (6.14)$$

The shortest spatial period harmonic that can be resolved will be determined in the case of uniform sample spacing by the Nyquist frequency, and will therefore be given, for a sample spacing Δx , by

$$j_{\max} = \left\lfloor \frac{\min(\lambda_a, \lambda_b)}{4\Delta x} \right\rfloor. \quad (6.15)$$

Higher order harmonics will tend to be strongly attenuated by losses and minor misalignments within the interferometer optics, and as a result of this in many cases $j_{\max}$ will be limited by the optical design of the interferometer, rather than the spatial sampling rate. Typically harmonics of up to four times the fringe period of the interferometer are observed.

The basic principles of the non-linearity correction methodology are shown in figure 6.14, with figure 6.14 **A** and **B** showing x_a and x_b as defined in equations 6.13 and 6.14. Comparing equations 6.13 and 6.14, it can be seen that if the harmonic coefficients $A_{a/b,j}$ or $\phi_{a/b,j}$ can be determined, the true displacement can be recovered from the displacement measurement at either wavelength. Taking the difference between the two displacement measurements yields a combined sum over both sets

of harmonics as

$$d = x_a - x_b, \quad (6.16)$$

$$= \sum_{j=1}^{j_{\max}} \left(A_{a,j} \cos \left(4\pi \frac{j}{\lambda_a} x + \phi_{a,j} \right) - A_{b,j} \cos \left(4\pi \frac{j}{\lambda_b} x + \phi_{b,j} \right) \right), \quad (6.17)$$

shown in figure 6.14 **C**.

The problem then becomes one of separating harmonics of differing spatial periods, pointing naturally towards a Fourier approach. The amplitudes of the Discrete Fourier Transform (DFT) of the displacement difference are presented in figure 6.14 **D**, where in order to separate the non-linearity harmonics a rectangular window function has been applied around the first three even harmonics of λ_a , in this case corresponding to spatial periods of 316.5 nm, 158.25 nm and 105.5 nm. The width of the window around each non-linearity harmonic is set in this work to twice the frequency resolution of the DFT. Having applied this window function, the DFT is then inverted, resulting in an estimate of the non-linearity error in the measured displacement x_a , shown in figure 6.14 **E**, that approximates the non-linearity error as

$$e_a \approx \sum_{j=1}^{j_{\max}} A_{a,j} \cos \left(2\pi \frac{j}{\lambda_a} x + \phi_{a,j} \right). \quad (6.18)$$

This error estimate is then subtracted from the measured displacement, resulting in a displacement measurement with reduced non-linearities, given by

$$x_{a,c} = x_a - e_a, \quad (6.19)$$

as shown in 6.14 **F**. Whilst in this example the correction has been applied to x_a only, the same approach may also be applied to x_b .

The spatial period windowing approach adopted here restricts this technique to cases where the non-linearity harmonics do not overlap, a condition that is dependant on the choices of λ_a and λ_b , the displacement range covered, and the stability of the non-linearities. Assuming uniform spatial sampling, ideally stable wavelengths, and non-linearities that remain stable in both amplitude and phase throughout the measurement such that no peak broadening takes place, the frequency spacing between non-linearity terms must remain greater than the frequency resolution of the DFT, given by the inverse of the displacement range covered. In this work, wavelengths of 632.8 nm and 785 nm were adopted, with the first closely spaced spatial harmonics occurring at 79.1 nm (the 8th harmonic of 632.8 nm) and 78.5 nm (the 10th harmonic of 785 nm). Separating these harmonics under ideal conditions would require a displacement range of 10.3 μ m, with a sample spacing of less than 39.25 nm. As previously discussed, higher order harmonics are likely to be strongly attenuated by the interferometer optics, and as such this restriction is not anticipated to limit the performance of the correction. If the harmonics present in the interferometer

outputs are known, λ_a and λ_b may be chosen to avoid overlaps.

Spatial resampling

For a (spatially) uniformly sampled displacement measurement, the described approach may prove sufficient, however many real world measurements will not be uniformly spatially sampled due to non-linear displacement paths. Further problems arise with the methodology proposed in figure 6.14 due to the finite sample length and non-integer number of periods contained in the sample. The results of this are apparent in 6.14 **D** where the amplitude peaks in the DFT are broadened, and in figure 6.14 **E** and **F** as a low frequency modulation in the estimated non-linearity error amplitude where none exists in the modelled data. In addition to these problems, the measured displacement, either x_a or x_b , chosen to form the ‘timebase’ of the DFT contains the non-linearity errors, and the effects of this timebase modulation must be considered.

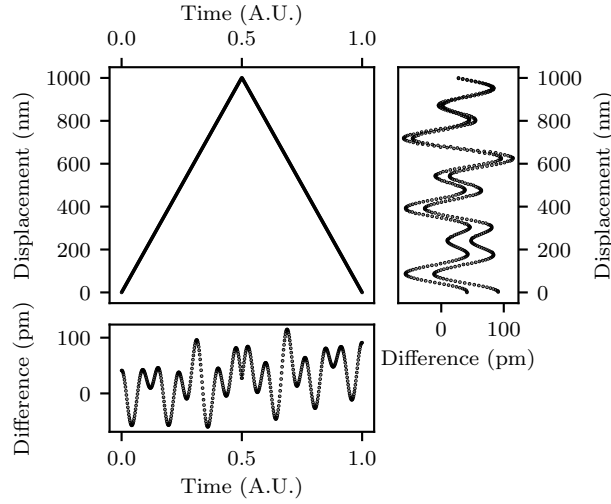


Figure 6.15: Demonstration of the effect of multiple displacement values at each displacement point on the interpolation process for a measurement including a linear 50 pm drift over the full measurement period between the λ_a and λ_b displacement measurements.

To counter first the problem of non-uniform sampling, the calculated displacement differences d must be interpolated onto a linear displacement basis, resulting in

$$d_{a,L} = \sum_{j=1}^{j_{\max}} \left(A'_{a,j} \cos \left(4\pi \frac{j}{\lambda_a} x_{a,L} + \phi'_{a,j} \right) - A'_{b,j} \cos \left(4\pi \frac{j}{\lambda_b} x_{a,L} + \phi'_{b,j} \right) \right) \quad (6.20)$$

and

$$d_{b,L} = \sum_{j=1}^{j_{\max}} \left(A'_{a,j} \cos \left(4\pi \frac{j}{\lambda_a} x_{b,L} + \phi'_{a,j} \right) - A'_{b,j} \cos \left(4\pi \frac{j}{\lambda_b} x_{b,L} + \phi'_{b,j} \right) \right) \quad (6.21)$$

for interpolations carried out with respect to the x_a and x_b measurements respectively, where $x_{a,L}$ and $x_{b,L}$ are the linearly spaced displacements and the primed amplitude and phase coefficients indicate that these coefficients are not identical to those in equation 6.16 due to the aforementioned timebase modulation. A nearest neighbour interpolation algorithm is applied here, chosen to avoid biasing the results with an interpolation algorithm more closely matched to the physical displacements generated in the experimental work. In the case of displacement profiles that are self-overlapping, this interpolation process will, in the case of nearest neighbour interpolation, introduce some high frequency fluctuations between the multiple displacement values if drift between the two interferometers occurs, as demonstrated in figure 6.15. These high frequency fluctuations will be filtered out by the Fourier transform and windowing process, and as such will not affect the corrected displacement.

The effects of a non-integer number of periods within the finite sample length are mitigated by first clipping $d_{a,L}$ and $d_{b,L}$ to an integer number of fringes for each wavelength respectively, then tiling the clipped data onto the start and end of the data set. Separate interpolations of the displacement difference are carried out for the x_a and x_b displacements as due to the periodicity of the non-linearities, the errors due to the non-linearities at harmonics of each wavelength are zero at multiples of that wavelength. The use of separate interpolations therefore allows the data to be clipped to an integer number of fringes as precisely as possible for each wavelength. For the experimental results gathered here this approach was found to be successful, however, it should be noted that by tiling the data the implicit assumption is made that non-linearity errors at the start and end of the displacement range are similar. Displacement differences falling outside of an integer number of fringes are therefore, as a consequence of the tiling, corrected based on the errors at the opposite end of the integer fringe range. This extrapolation is only appropriate where the previously stated assumption holds.

Following the frequency space windowing and DFT inversion the recovered non-linearity estimates $e_{a,L}$ and $e_{b,L}$ must then be interpolated back onto the measured displacements x_a and x_b before subtracting to give the corrected measurements. In this work a nearest neighbour interpolation approach was once again used for the back interpolation step.

6.3.3 Experimental methods

In order to demonstrate the two-wavelength non-linearity methodology experimentally, a homodyne Michelson interferometer was constructed making use of a phase-quadrature beam splitter (PQBS) [25] cube designed originally for use at 633 nm.

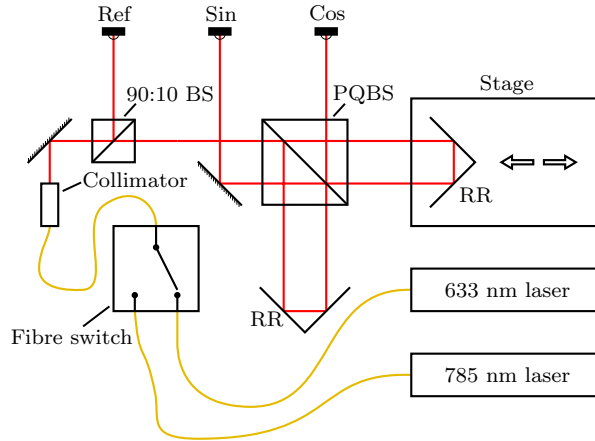


Figure 6.16: Schematic of the two-wavelength interferometer. RR, retro-reflector; PQBS, phase-quadrature beam splitter; 90:10 BS, reference pick off beam splitter; Ref, intensity reference photodiode; Sin, sine quadrature photodiode; Cos, cosine quadrature photodiode.

A schematic of the interferometer is shown in figure 6.16. The external faces of the PQBS and the front surfaces of the solid glass corner cube retroreflectors were deliberately aligned to introduce higher order non-linearities. λ_a was provided by a thermally stabilised Helium-Neon (HeNe) laser operating at approximately 632.8 nm (REO 32734) and λ_b was provided by a 785 nm diode laser (Schäfter + Kirchhoff 51nano-S-785-12-Q06-P-5-2-18-0-150). The choice of wavelengths in this case was primarily determined by availability of equipment, however, a number of factors may influence the suitability of pairs of wavelengths; this is discussed further in section 6.3.5. The two wavelengths were multiplexed together with a MEMS fibre optic switch (Sercalo SN-1x2-4-N) using a switching frequency of approximately 30 Hz. By fibre coupling both laser sources into the same fibre, coaxial propagation of the two interferometers is ensured. A mirror collimator was employed to ensure beams at both wavelengths were collimated. ThorLabs SM600 fibre was used from the fibre switch to the collimator, with a cut off wavelength between 500 nm and 600 nm. At a wavelength of 785 nm this fibre may exhibit some undesirable bend losses; the effect of any varying losses were minimised by normalisation to an intensity reference output. The reference, sine and cosine photodiode outputs shown in figure 6.16 were amplified and digitised simultaneously with electronics developed in house at the National Physical Laboratory (NPL). Demultiplexing of the two wavelengths was carried out in software based on rising edges in the reference channel signal, with the sine and cosine channels normalised relative to the demultiplexed reference for each wavelength in order to minimise the effects of longer term drift in the laser powers. A further normalisation based upon a Gaussian envelope fitted to maxima and minima of the sine and cosine fringes was carried out for the λ_b fringes. This normalisation was required to compensate for the low coherence length (approximately 300 μm) of the 785 nm laser.

A Heydemann correction [21, 87] was applied to both sets of fringes prior to applying the two-wavelength correction. The Heydemann correction fits an ellipse to

the Lissajous figure formed by plotting the two quadrature outputs of a homodyne interferometer against each other, and uses the resulting ellipse fit coefficients to transform the Lissajous figure to the ideal origin centred circular form. As a result, any quadrature phase errors, signal offsets and gain ratio errors are minimised.

Whilst applying the Heydemann correction in this case was not strictly necessary, as the two-wavelength correction may also be employed to correct for the effects of elliptical distortions to the quadrature signals, some offset must be applied to the measured optical powers in order to ensure the origin falls within the approximately elliptical Lissajous figure, in order to allow fringe counting to take place. As the Heydemann correction has no disadvantages in this case, with the restrictions to displacements of greater than one optical fringe and non-real time correction shared with the two-wavelength approach, the Heydemann correction was employed as a pre-processing step.

The interferometric phase was then calculated through a four-quadrant arctangent function, and unwrapped at discontinuities of greater than π . For this proof of concept experiment no environmental parameters were recorded, and therefore no Edlén style wavelength corrections [92] were applied and the laser wavelengths were assumed constant. The small scaling errors introduced by this assumption do not affect the ability of the system to distinguish the presence of non-linearity harmonics, and any error in the estimated amplitude of the non-linearity harmonics introduced by this fixed wavelength assumption is estimated to be less than 1 pm. As no wavelength changes were measured, and the interferometer was operated with a path length difference of the order of micrometres, displacements for the λ_a interferometer were calculated without a dead path correction term [49] as

$$x_a = \frac{\lambda_a}{4\pi} \Delta\varphi_a, \quad (6.22)$$

where φ_a is the unwrapped interferometric phase for the λ_a interferometer, based on a wavelength of 632.8 nm. λ_b displacements were calculated from the unwrapped interferometric phase at λ_b , φ_b , as

$$x_b = \frac{\max(x_a) - \min(x_a)}{\max(\varphi_b) - \min(\varphi_b)} \Delta\varphi_b, \quad (6.23)$$

negating any requirement for precise knowledge of the wavelength λ_b .

Test displacements were generated with a Queensgate Instruments NPS-X15A nanopositioning stage, and non-linearities before and after correction were estimated from a DFT of the displacement error relative to the stage position, permitting non-linearities to be separated from higher frequency noise and lower frequency drift terms.

This system, whilst non-optimal for precision interferometry, proved sufficient to serve as a proof of concept of the two-wavelength methodology. The limitations of this system are considered in the discussion section, along with suggested improvements for a system suitable for practical application.

Three displacement scan profiles were used to demonstrate the two-wavelength correction: a unidirectional linear scan, a mirrored linear scan and a sinusoidal

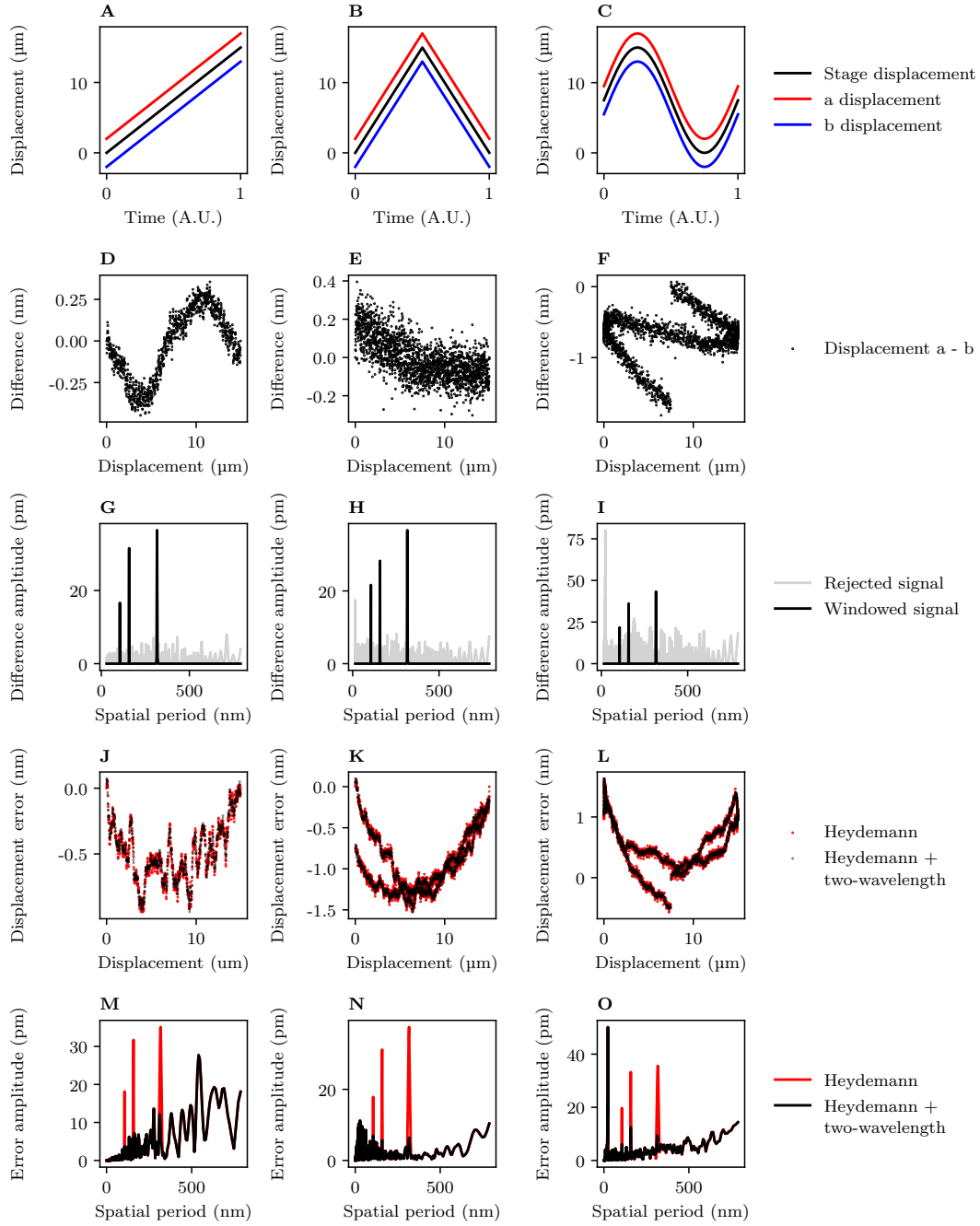


Figure 6.17: Stages of the two-wavelength non-linearity correction process for three displacement profiles (A, B and C) as applied to correction of the λ_a displacement measurement. Offsets have been applied to the displacement profiles in A, B and C for clarity. D, E and F show the difference between the λ_a and λ_b interferometer displacement measurements. G, H and I show the DFT of the displacement difference following the interpolation, clipping and tiling process described in the text, with interpolation performed relative to the λ_a displacements. J, K and L show the errors in the λ_a displacement measurements, relative to the stage position, for the three displacement profiles, and for both Heydemann corrected and Heydemann and two-wavelength corrected data. M, N and O show DFTs of the data presented in J, K and L, the error harmonics relative to the stage position.

scan, as shown in figure 6.17 **A**, **B** and **C**. In all cases the stage was moved in 1000 steps through a 15 μm range, allowing both the effects of non-uniform spacing (in the case of the sinusoidal scan) and multiple displacement difference values at each displacement point (in the case of the mirrored linear scan) to be tested.

6.3.4 Results

Considering first the linear displacement profile (figure 6.17 **A**), the calculated displacement difference between the two interferometers at each stage position is shown in figure 6.17 **D**. Due to the relatively large amount of drift between the interferometers, introduced by a combination of drift in λ_b , refractive index fluctuations and thermal fluctuations over the course of the measurement, non-linearities are not clearly visible. The combined effects of refractive index fluctuations (including dispersion between the two wavelengths) and thermal drifts can also be seen in figure 6.17 **E**, **F**, **J**, **K** and **L**. These effects are non-repeatable. The DFT of the interpolated displacement differences is shown for the linear displacement profile in figure 6.17 **G**, with peaks apparent at spatial periods of $\lambda_a/2$, $\lambda_a/4$ and $\lambda_a/6$. Figure 6.17 **J** shows the results of the correction made based upon the windowed DFT shown in figure 6.17 **G**, with displacement errors shown relative to the stage position for interferometric displacement measurements made at λ_a following the application of both a Heydemann correction, and a combined Heydemann and two-wavelength correction. DFTs of the λ_a displacement errors relative to the stage positions are shown for Heydemann corrected and Heydemann and two-wavelength corrected displacements in figure 6.17 **M**, with the estimated non-linearity amplitude, calculated as the sum of the $\lambda_a/2$, $\lambda_a/4$ and $\lambda_a/6$ DFT peaks, reduced from 84 pm to 11 pm.

Corresponding results are also shown in figure 6.17 for the mirrored linear displacement profile (**B**, **E**, **H**, **K** and **N**) and sinusoidal displacement profile (**C**, **F**, **I**, **L** and **O**). Non-linearities, again estimated as the sum of the first three even harmonics of λ_a , were reduced from 86 pm to 18 pm for the mirrored linear displacement profile. For the sinusoidal displacement profile estimated non-linearities were reduced from 89 pm to 26 pm. It should be noted that as non-linearities were estimated with a Fourier technique, the estimates may be less accurate for the non-uniform mirrored linear and sinusoidal displacement profiles, as interpolation onto a linear displacement basis was required prior to applying the DFT. The increased magnitude of errors with spatial periods above 300 nm present in figure 6.17 **J** are responsible for the increased noise floor in figure 6.17 **M** as compared to figure 6.17 **N** and **O**. These fluctuations are again attributed to air turbulence effects.

In order to investigate the effects of reduced displacement scan ranges on the correction, data from the linear displacement scan shown in figure 6.17 was taken in segments covering 100 displacement ranges of 1 μm to 7.5 μm with randomly generated offsets from the starting displacement in the range 0 μm to 7.5 μm applied. Non-linearity amplitudes in the λ_a interferometer were then estimated from the sum of the displacement error DFT peaks for the first three even harmonics. This process was repeated with 1000 randomly generated offsets, and means and standard errors in the means calculated, as plotted in figure 6.18. As may be expected from a Fourier technique the non-linearity correction can be seen to be more effective where a

larger displacement range is available, however some non-linearity reduction remains possible down to a displacement range of 1 μm . Oscillations in the two-wavelength corrected non-linearity are introduced by the varying number of fringes contained in the displacement range. It should be noted that for smaller displacement ranges the Fourier estimation of non-linearity amplitudes becomes less reliable.

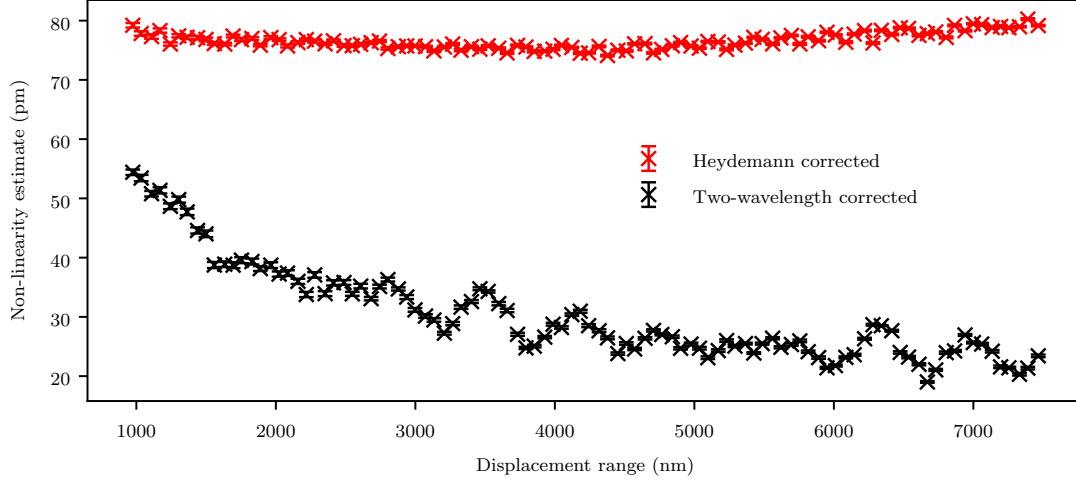


Figure 6.18: Variation of the non-linearity amplitude, estimated from the sum of the displacement error DFT peaks for the first three even harmonics, with displacement range.

6.3.5 Discussion

The results presented here demonstrate the capability of the two-wavelength non-linearity correction methodology to improve upon the non-linearities achieved by application of the commonly employed Heydemann style [21,87] ellipse fitting correction. In particular, non-linearities introduced by unwanted cavity formation within the optics, present in many interferometer designs due to back reflections (as is the case in this work) or polarisation leakage, may be corrected for with the two-wavelength methodology, a correction that to the knowledge of the authors has not previously been possible with a purely optical instrument. Additionally, the required precision of the ellipse fitting correction is reduced, as residual non-linearities after the ellipse fitting correction are corrected by the two-wavelength correction. Whilst a homodyne interferometer design has been demonstrated in this case, the technique is anticipated to be equally applicable to heterodyne systems.

Residual non-linearities have been estimated in this work using a Fourier technique. As such, there is some uncertainty associated with the residual non-linearity estimates. Non-linearities that vary with displacement may be underestimated by Fourier techniques, as the amplitudes of peaks in the Fourier transform are based on the assumption of ideally periodic signal components. In this proof of concept work, a clear reduction in the observed peaks in the Fourier transform of the displacement

error has been demonstrated, however, further work measuring the non-linearity errors directly in a more stable system, for example against an X-ray interferometer as in [11], is needed before the performance of the two-wavelength technique can be stated conclusively. For this reason, the term ‘estimated residual non-linearity’ is used throughout this work when discussing the achieved results.

The remainder of this discussion will be broken into sub-sections focussing on several key aspects of the two-wavelength non-linearity correction. First, experimental considerations that are not explored in this proof of concept work will be discussed. Following this, the existing limitations of, and potential improvements to, the proposed processing algorithm will be considered. A qualitative discussion of uncertainty contributions will then be made. Finally, potential avenues for further work will be proposed.

Experimental considerations

Several aspects of the two-wavelength correction methodology are sensitive to experimental conditions that are not explored in this work. Firstly the choice of λ_a and λ_b must be considered. In this case a HeNe laser operating at 633 nm was selected for λ_a due to the widespread use of this wavelength for metrological applications. The choice of λ_b was primarily determined by equipment availability, however it is anticipated that in practice a lower cost unstabilised laser is likely to be employed as the second wavelength, for example a diode laser, or a HeNe lasing at a different wavelength. The use of unstabilised lasers is possible as exact knowledge of λ_b is not required as a result of the rescaling of the λ_b displacements based upon the λ_a displacements. Several restrictions are placed on λ_b and the associated laser source. Any non-linearity harmonics present in the interferometer must not overlap between the two wavelengths, with the acceptable frequency spacing between harmonics determined by the frequency resolution of the displacement difference DFT (and therefore the displacement range covered and spatial sampling rate), the wavelength stability of the source laser and the stability of the non-linearities themselves. As alluded to by the previous restriction, λ_b must remain sufficiently stable for the duration of the measurement such that non-linearity harmonics can be separated in the DFT, with the required stability determined by the duration of the measurement, the non-linearity harmonics present in the interferometer, and the spacing of the wavelength harmonics. The λ_b source laser must also have a sufficient coherence length to cover the desired measurement range. In addition to these more fundamental restrictions, further restrictions may be imposed on λ_b if λ_a and λ_b are to be fibre coupled, due to the finite single mode wavelength range of commonly available optical fibres, necessitating the use of more expensive photonic-crystal fibres if the spacing between λ_a and λ_b is large. If both wavelengths are to be detected by the same detector, as is the case in this work, the availability detectors suitable for operation at both λ_a and λ_b must also be considered.

Along with the stability of the laser wavelengths, the stability of the non-linearities themselves may affect the correction. Residual non-linearities after applying a Heydemann style correction are likely to be higher order effects, existing due to unwanted cavity formation within the interferometer optics, with the cavity necessarily including a reflection from the moving mirror or retro-reflector for

a non-linearity to be produced. As such the non-linearities targeted with the two-wavelength correction may be highly sensitive to the alignment of the interferometer optics with respect to the moving mirror or retro-reflector, an alignment that is prone to vary with both displacement and time. Variation in the amplitudes of the non-linearity harmonics manifests as broadened or split DFT peaks, and as such a wider frequency domain window will be required, again placing limits on the frequency spacing of harmonics of λ_a and λ_b in systems where non-linearities are prone to drift.

In this work combination and separation of the two wavelengths was achieved by multiplexing the fibre coupled laser sources together with a fibre switch at a relatively low switching frequency of 30 Hz. This approach, whilst adequate as a proof of concept, severely limits the measurement speed of the system and increases the effect of mechanical vibrations at higher frequencies than the switching speed on the displacement difference signal due to aliasing effects. As such alternatives would be preferable in practical application. Several approaches are possible, including spatially separating the two interferometers, modulation and lock-in amplification at different frequencies for each wavelength, separation with dichroic beam splitters or diffraction gratings, and separation via orthogonal polarisation. Of these techniques, modulation and lock-in amplification would be particularly well suited to heterodyne interferometer designs as such a modulation is already present. Most heterodyne interferometer designs are however reliant upon polarising optics, and finding polarising optics with suitable performance at both wavelengths may prove restrictive. Spatially separating the two interferometers would allow use of often lower cost and higher performing monochromatic optical components in both heterodyne and homodyne interferometer designs, and Abbe and cosine errors introduced by small misalignments between the optical axes of the two interferometers would, over sufficiently small measurement ranges, be compensated for by the rescaling effect of equation 6.23. Dichroic beam splitters offer a relatively straightforward implementation of the two-wavelength technique, however imperfect dichroic separation may result in the presence of λ_b harmonics in the λ_a displacement and visa versa. Leakage between the wavelengths of this sort would result in non-linearities common to both interferometers, which would not appear in the displacement difference and would therefore remain uncorrected. Leakage between the two wavelengths is also likely to prove problematic with diffraction grating or polarisation based separation.

A final experimental consideration is the required spatial sampling rate and displacement range. The minimum sampling rate is determined by the highest order non-linearity harmonic present in the interferometer, and as such will be highly dependant on the optical design of the interferometer. For the interferometer presented in this work higher order non-linearities are introduced by ghost reflections from the external surfaces of the phase-quadrature beam splitter cube and the stage mounted retro-reflector, and as such are attenuated by both the anti-reflection coating optics and by small misalignments between the optical surfaces. As such the highest order non-linearity observed occurs at a spatial period of approximately 105.5 nm ($\lambda_a/6$), requiring a spatial sampling period of 52.75 nm to resolve. Shorter wavelengths and the presence of higher order non-linearities will both serve to increase

the required spatial sampling rate. The required displacement range for the correction to prove effective is related to the choice of wavelengths and spacing of the extant non-linearity harmonics as previously discussed, however as can be seen from figure 6.18 under the experimental conditions in this work, increased displacement range led to reduced residual non-linearities, with small increases in performance where the displacement range covered an integer number cycles of a non-linearity harmonic. This relationship may not hold true where non-linearities are highly variable with displacement, and in such cases over larger displacement ranges a moving overlapped DFT approach may prove preferable.

Data processing considerations

In addition to the experimental considerations, the performance of the correction may be sensitive to the data processing methods applied, in particular the displacement difference interpolation and tiling, and the frequency spacing windowing steps. Where the measured displacement profile is self overlapping, as is the case for both the mirrored linear and sinusoidal displacement profiles examined in this work, multiple displacement difference values exist at each measured displacement point. The effect of this on the interpolation process can be observed in figure 6.17 **H** and **I** where peaks are present at short spatial periods due to the nearest neighbour interpolation algorithm switching between displacement differences measured at similar displacement values at differing time points. These short spatial period artefacts did not negatively affect the results as they occurred at spatial periods shorter than the highest order non-linearity harmonic present, and due to the highest resolved non-linearity harmonic being limited by the spatial sampling rate, this should remain the case in all systems. In this proof of concept experiment non-linearities are not the dominant uncertainty source, however, it should be noted that in practical application the two-wavelength correction is only beneficial in cases where the repeatability of the measurement is better than the Heydemann corrected non-linearities.

Both the choice of frequency space windowing function and frequency space window width have the potential to affect the correction process. In this work, possibly due to the relatively large amounts of both noise and drift present in the displacement measurements, the window function employed was not found to affect the residual non-linearities after correction, as such a simple rectangular window function was applied. In a more sensitive instrument, the ringing effects introduced by a rectangular frequency space window may become apparent, and alternative window functions may be required. The widths of the windows were found to strongly affect the results, with a window width determined in this work from the frequency resolution of the DFT. Increased window widths may be beneficial under some circumstances, for example where large variations in the non-linearities take place during the measurement process, leading to broadened DFT peaks.

Uncertainty

Uncertainties in the corrected displacements will be determined by the uncertainties in the displacement measurements at both wavelengths, and will in general be increased relative to the uncorrected uncertainty to a degree determined by the width

of the frequency space window function. Non-linearities that vary with displacement, or wavelengths that vary over the course of the measurement will therefore increase the uncertainty in the corrected result due to the requirement for wider window functions. The frequency space windowing process will also introduce a degree of statistical correlation between individual corrected displacement measurements, and this may need to be considered if the corrected displacement values are to be further processed.

An additional uncertainty contribution arises from the fact that the displacement difference is sampled on a displacement basis that is itself non-linear, and this effect limits the ultimate performance of the correction. For the typically sub-nanometre non-linearity values encountered in precision displacement measuring interferometry, the effects of this ‘timebase’ modulation will be small. Higher order non-linearities will have a stronger timebase modulation effect as the ratio of the amplitude to the period of the distortion is increased, all else being equal.

Uncertainties in the values of the two wavelengths will also have an effect on the measurement uncertainty, with contributions arising from both uncertainty in the laser frequency, and uncertainty in the refractive index in the beam path. The λ_b wavelength is not used in the correction as implemented here, with the λ_b displacement scaled relative to the λ_a displacements, with the effective uncertainty in the average value of λ_b over the course of the measurement therefore determined by the uncertainty in the interferometric phase measurements made at λ_b (which are not affected by wavelength), and the uncertainty in λ_a . λ_a uncertainties in this case were dominated by the uncertainty in the refractive index as no Edlén correction was performed.

In this proof of concept work uncertainties in the measured displacements were dominated by longer period drift terms, rather than statistical uncertainties in individual measurement points or non-linearities, and as such a detailed assessment of the measurement uncertainty has not been made here.

Further work

As previously discussed, the Fourier technique employed in this work to estimate residual non-linearities has the potential to underestimate the residual non-linearity. As such, before the two-wavelength correction can be employed in rigorous metrological applications, a more accurate assessment of the residual non-linearities should be made. Direct traceable measurement of picometre scale non-linearities is possible through X-ray interferometry [8, 11].

In addition to verification of the performance of the correction, an uncertainty budget would be required for the displacement measurement in metrological applications. Uncertainties have not been assessed quantitatively in this work as it is clear from the results presented in figure 6.17 **J**, **K** and **L** that non-linearities are not the dominant uncertainty source, and therefore any assessment of the uncertainty resulting from the two-wavelength correction could not be verified experimentally. Future work may develop an uncertainty budget, and this could again be verified with X-ray interferometry [11].

In addition to the reduction in residual non-linearities, two-wavelength interferometry also has the potential to allow for improved refractive index compensation [88], and this may be an avenue for future work. Further work may also investigate the application of the Fourier algorithm developed here to other differently non-linear b displacement measures, for example capacitive sensors as suggested in [68]. Conversely, the harmonic fitting approach proposed in [68] could be applied to two-wavelength interferometry, and may prove preferable to the Fourier approach in cases where noise levels are low and non-linearities are stable with respect to both displacement and time over the ranges covered.

6.3.6 Conclusions

The two-wavelength non-linearity correction methodology has been shown to be capable of correcting for residual non-linearity errors that remain after application of Heydemann style [21, 87] ellipse fitting corrections. Residual non-linearities after Heydemann correction may be caused by polarisation leakage or unwanted back reflections (ghost reflections) within the interferometer optics, and affect both homodyne [84] and heterodyne [60, 61, 63, 65] interferometers. The two-wavelength correction is not dependant on a particular interferometer detection regime, and as such is anticipated to be applicable to both homodyne and heterodyne designs. In the best case a reduction in estimated non-linearity from 84 pm to 11 pm as compared to Heydemann corrected displacements was demonstrated for a linear displacement profile. In the worst case this reduction was from 89 pm to 26 pm for a sinusoidal displacement profile. This work is intended as a proof of concept, and as such the results presented here are not anticipated to represent the performance limits of the two-wavelength technique. Equally, as this is a proof of concept, further work is needed to quantify the uncertainties associated with the two-wavelength technique.

The proposed correction is implemented in post processing, and future work may implement the two-wavelength correction in a psuedo-real time manor, with rolling overlapping Fourier transforms. Additional future work could combine the two-wavelength non-linearity correction with two-wavelength refractive index compensation [88].

6.4 Multiple intensity reference non-linearity correction

6.4.1 Introduction

As previously discussed within this chapter, a wide variety of methods have been described in the literature to correct for non-linearities in homodyne interferometry [21, 83, 87, 111, 163, 169], including methods designed for sub-fringe displacements [163, 169]. Existing sub-fringe correction methods proposed in the literature are however limited to residual non-linearities of approximately 1 nm where displacements of less than one half of an optical fringe are to be corrected [163, 169], far in excess of the tens of picometre non-linearities that can be achieved by optical

interferometry in the multi-fringe case [11], although the performance limits may be expected to be dependent on the exact interferometer system to which the correction is applied.

A non-linearity correction methodology is presented in this work that is capable of correction of non-linearities in homodyne optical interferometers for sub-fringe displacements, outperforming the Heydemann correction for displacements of less than approximately 120 nm. Two closely related techniques are proposed, both based upon measurement of the optical power returned from each arm of the interferometer, following a related methodology to that of [169] with the goal of reducing residual non-linearities to tens of picometres. First, the theory behind the methodology will be presented, along with the theory behind the calibration procedure developed in this work. The methodologies used to investigate the performance of the correction will then be described, followed by experimental results validating the correction. The limitations and benefits of the correction will then be considered and compared with the Heydemann [87] and beam shuttering [169] techniques.

6.4.2 Theory

Quadrature fringe counting in homodyne displacement measuring interferometry relies upon the interferometer producing two signals, separated by a 90° phase difference in the ideal case [21]. A phase quadrature coating [25] based interferometer is depicted in figure 6.19 with quadrature outputs $I_{\sin}$ and $I_{\cos}$, and intensity reference outputs I_{mes} and I_{ref} . Two closely related non-linearity correction methodologies are presented in this section, an overview of which is given in figure 6.20.

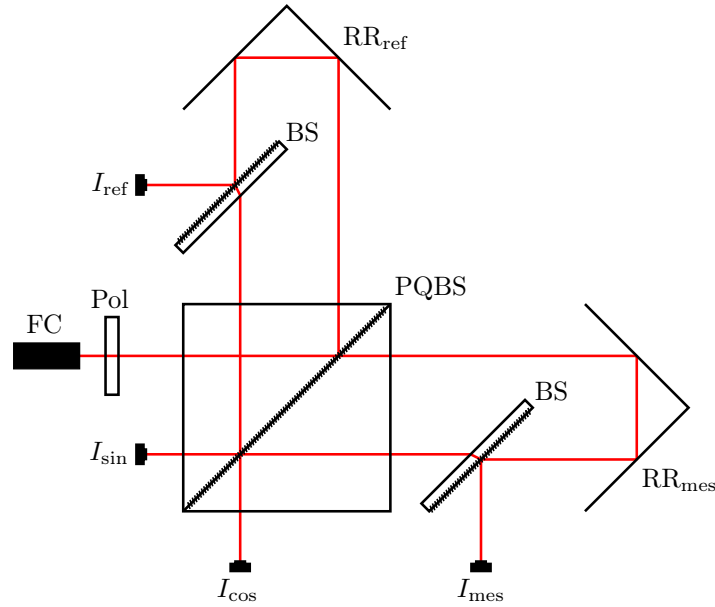


Figure 6.19: Schematic of the phase quadrature coating based homodyne interferometer. FC, fibre collimator; PQBS, phase quadrature beam splitter; Pol, polariser; BS, 50:50 beam splitter; RR_{mes} and RR_{ref} , measurement and reference retroreflectors; I_{mes} , measurement beam reference intensity; I_{ref} , reference beam reference intensity; $I_{\sin}$, sine quadrature output; $I_{\cos}$, cosine quadrature output.

Assuming an optical intensity after the polariser of I_0 , the optical intensity signals at the four outputs will be given by

$$I_{\cos} = \Re_Q \Im_Q \Im_P \left(\Im_{\text{mes}} + \Im_{\text{ref}} + 2\sqrt{\Im_{\text{mes}} \Im_{\text{ref}}} \cos(\varphi) \right) I_0, \quad (6.24)$$

$$I_{\sin} = \Im_P \left(\Im_{\text{ref}} \Re_Q^2 + \Im_{\text{mes}} \Im_Q^2 + 2\Re_Q \Im_Q \sqrt{\Im_{\text{mes}} \Im_{\text{ref}}} \sin(\varphi + \delta') \right) I_0, \quad (6.25)$$

$$I_{\text{ref}} = \Im_{\text{ref}} \Re_Q \Re_P I_0, \quad (6.26)$$

$$I_{\text{mes}} = \Im_{\text{mes}} \Im_Q \Re_P I_0, \quad (6.27)$$

where $\Re_Q$ and $\Im_Q$ are the power reflectivity and transmissivity coefficients of the phase-quadrature beam splitter (PQBS) coating, $\Re_P$ and $\Im_P$ are the power reflectivity and transmissivity coefficients of the non-polarising beam splitter (BS) coatings, $\Im_{\text{ref}}$ and $\Im_{\text{mes}}$ are the power transmissivity coefficients of the reference and measurement arm beam paths, φ is the interferometric phase, determined by the illuminating wavelength, refractive index, and relative path lengths of the reference and measurement arms of the interferometer, and δ' is the quadrature phase error.

Equations 6.24 and 6.25 assume perfect interference between the beams returned from the reference and measurement arms, however in reality wavefront or polarisation aberrations may result in a reduction in fringe contrast at the sine and cosine outputs, without affecting the intensity measured at the reference and measurement intensity outputs. Such effects will alter the amplitude of the measured quadrature signals, but will not affect the offsets of the quadrature signals, a fact that will be exploited later in the correction process.

The optical intensity signals are not measured directly, and are instead commonly detected and amplified to give a voltage signal with a generally unknown gain and offset for each signal, resulting in

$$V_{\cos} = G_{\cos} I_{\cos} + D_{\cos}, \quad (6.28)$$

$$V_{\sin} = G_{\sin} I_{\sin} + D_{\sin}, \quad (6.29)$$

$$V_{\text{ref}} = G_{\text{ref}} I_{\text{ref}} + D_{\text{ref}}, \quad (6.30)$$

$$V_{\text{mes}} = G_{\text{mes}} I_{\text{mes}} + D_{\text{mes}}, \quad (6.31)$$

where the coefficients G represent (positive) gain, and the coefficients D represent the voltage offset that is present in the absence of laser illumination, introduced by both background illumination and electronic effects. For simplicity of notation in the following equations, offset free voltage signals will be introduced as

$$V'_{\cos} = V_{\cos} - D_{\cos}, \quad (6.32)$$

$$V'_{\sin} = V_{\sin} - D_{\sin}, \quad (6.33)$$

$$V'_{\text{ref}} = V_{\text{ref}} - D_{\text{ref}}, \quad (6.34)$$

$$V'_{\text{mes}} = V_{\text{mes}} - D_{\text{mes}}. \quad (6.35)$$

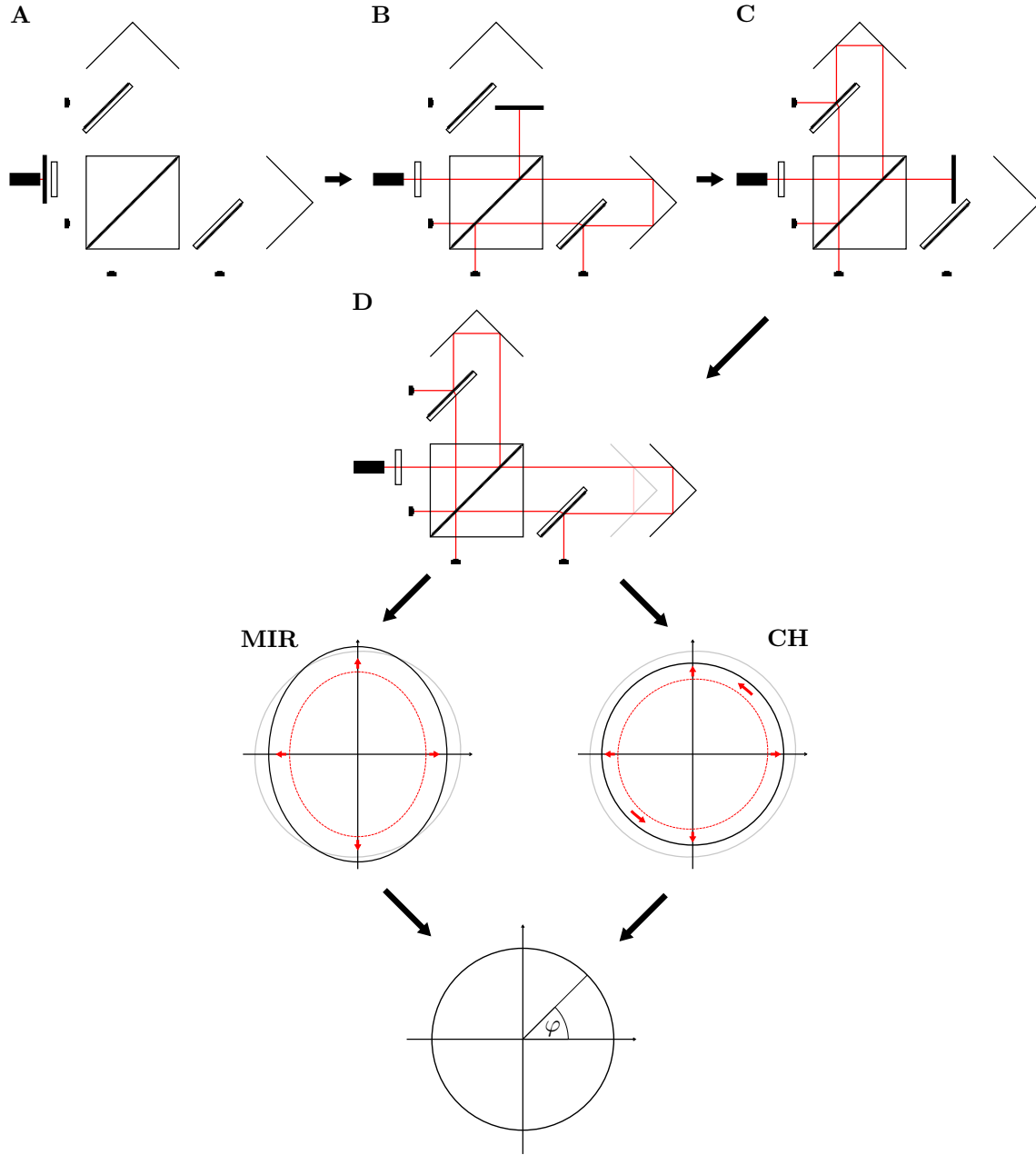


Figure 6.20: Flowchart demonstrating the MIR and constrained Heydemann (CH) non-linearity correction methodologies. First, the input beams are blocked and dark voltages measured (**A**). The reference (**B**) and measurement (**C**) arms are then blocked in turn, and calibration measurements taken. Following calibration, a displacement measurement is taken (**D**), and equations 6.38 and 6.39 are applied to centre the quadrature signals, and to attempt to correct the signal amplitudes. For the MIR approach, the sum and difference of the centred signals are taken, and an ellipse fit is made, centred on the origin, and with the ellipse axes aligned to the data axes (**MIR**). For the CH approach, the centred signals are fitted with an ellipse again constrained to be centred on the origin, but free to rotate about the origin (**CH**). Finally for both approaches the fit coefficients are used to correct the quadrature signals, and the phase is recovered from the now circular Lissajous figure.

This system of equations may be solved for the zero centred and equal amplitude quadrature signals $S_{\cos}$ and $S_{\sin}$, independently of illumination intensity and losses in the interferometer arms, as

$$S_{\cos} = \cos(\varphi),$$

$$= -\frac{G_{\text{ref}}G_{\cos}\mathfrak{I}_P\mathfrak{R}_QV'_{\text{mes}} + G_{\text{mes}}G_{\cos}\mathfrak{I}_P\mathfrak{I}_QV'_{\text{ref}} - G_{\text{mes}}G_{\text{ref}}\mathfrak{R}_P V'_{\cos}}{2G_{\cos}\mathfrak{I}_P\sqrt{G_{\text{mes}}G_{\text{ref}}\mathfrak{R}_Q\mathfrak{I}_QV'_{\text{mes}}V'_{\text{ref}}}}, \quad (6.36)$$

$$S_{\sin} = \sin(\varphi + \delta'),$$

$$= \frac{G_{\text{ref}}G_{\sin}\mathfrak{I}_P\mathfrak{I}_QV'_{\text{mes}} + G_{\text{mes}}G_{\sin}\mathfrak{I}_P\mathfrak{R}_QV'_{\text{ref}} - G_{\text{mes}}G_{\text{ref}}\mathfrak{R}_P V'_{\sin}}{2G_{\sin}\mathfrak{I}_P\sqrt{G_{\text{mes}}G_{\text{ref}}\mathfrak{R}_Q\mathfrak{I}_QV'_{\text{mes}}V'_{\text{ref}}}}, \quad (6.37)$$

however this form requires the optical and electronic properties of the interferometer system to be measured directly.

Considering equations 6.36 and 6.37, some simplifications can be made. Hodydyne interferometers recover the interferometric phase using the four quadrant arctangent function [21], which is insensitive to scaling factors common to both quadrature signals. As a consequence, the signals can be arbitrary scaled by a common factor, allowing equations 6.36 and 6.37 to be written in terms of gain ratios, rather than absolute gains, and reflectivity to transmission ratios, rather than absolute reflectivity and transmission coefficients. The cosine signal may also be inverted without affecting the relative phase measurement, allowing both signals to be rewritten with common signs for each component. Combining these changes results in

$$S'_{\cos} = \Gamma_Q V'_{\text{mes}} + G_{\text{m,r}} V'_{\text{ref}} - G_{\text{m,c}} \Gamma'_P V'_{\cos}, \quad (6.38)$$

$$S'_{\sin} = V'_{\text{mes}} + G_{\text{m,r}} \Gamma_Q V'_{\text{ref}} - G_{\text{m,s}} \Gamma'_P V'_{\sin}, \quad (6.39)$$

where

$$G_{\text{m,r}} = \frac{G_{\text{mes}}}{G_{\text{ref}}}, \quad (6.40)$$

$$G_{\text{m,c}} = \frac{G_{\text{mes}}}{G_{\cos}}, \quad (6.41)$$

$$G_{\text{m,s}} = \frac{G_{\text{mes}}}{G_{\sin}}, \quad (6.42)$$

$$\Gamma_Q = \frac{\mathfrak{R}_Q}{\mathfrak{I}_Q}, \quad (6.43)$$

$$\Gamma'_P = \frac{\mathfrak{R}_P}{\mathfrak{I}_P \mathfrak{I}_Q}. \quad (6.44)$$

Calibration procedure

In order to recover $S'_{\cos}$ and $S'_{\sin}$ the gain and reflectivity to transmission ratios must be measured. Whilst direct measurement would be possible, it is desirable to perform this calculation in-situ, as a calibration step. This may be achieved through beam blocking techniques, in a method related to that proposed in [169]. First, the dark voltage coefficients required to calculate the offset free voltage signals V' from the measured voltage signals V may be measured directly by blocking the input beam to the interferometer. By blocking first the reference, and then the measurement beam, separately, eight voltages may be recorded, some of which can be used to calibrate the required ratios. Blocking the reference beam results in voltages

$$V_{\cos}^{(R)} = G_{\cos} \Re_Q \Im_Q \Im_P \Im_{\text{mes}} I_0, \quad (6.45)$$

$$V_{\sin}^{(R)} = G_{\sin} \Im_Q^2 \Im_P \Im_{\text{mes}} I_0, \quad (6.46)$$

$$V_{\text{ref}}^{(R)} = 0, \quad (6.47)$$

$$V_{\text{mes}}^{(R)} = G_{\text{mes}} \Im_Q \Re_P \Im_{\text{mes}} I_0. \quad (6.48)$$

Similarly, blocking the measurement beam results in voltages

$$V_{\cos}^{(M)} = G_{\cos} \Re_Q \Im_Q \Im_P \Im_{\text{ref}} I_0, \quad (6.49)$$

$$V_{\sin}^{(M)} = G_{\sin} \Re_Q^2 \Im_P \Im_{\text{ref}} I_0, \quad (6.50)$$

$$V_{\text{ref}}^{(M)} = G_{\text{ref}} \Re_Q \Re_P \Im_{\text{ref}} I_0, \quad (6.51)$$

$$V_{\text{mes}}^{(M)} = 0. \quad (6.52)$$

From equations 6.45 to 6.52 the required ratios and products of ratios may then be determined as

$$G_{\text{m,r}} = \frac{V_{\text{mes}}^{(R)}}{V_{\text{ref}}^{(M)}} \sqrt{\frac{V_{\sin}^{(M)} V_{\cos}^{(M)}}{V_{\sin}^{(R)} V_{\cos}^{(R)}}}, \quad (6.53)$$

$$\Gamma_Q = \sqrt{\frac{V_{\sin}^{(M)} V_{\cos}^{(R)}}{V_{\cos}^{(M)} V_{\sin}^{(R)}}}, \quad (6.54)$$

$$G_{\text{m,c}} \Gamma'_P = \frac{V_{\text{mes}}^{(R)}}{V_{\cos}^{(R)}} \Gamma_Q, \quad (6.55)$$

$$G_{\text{m,s}} \Gamma'_P = \frac{V_{\text{mes}}^{(R)}}{V_{\sin}^{(R)}}. \quad (6.56)$$

From equations 6.38 and 6.39 and 6.53 to 6.56 equal amplitude zero centred quadrature signals can be calculated, however the quadrature phase error, δ' is still present. Two approaches are possible to correct for the quadrature error, which will be outlined in the following subsections.

Multiple intensity reference technique

The first approach, termed in this work the multiple intensity reference (MIR) technique, assumes the signals have ideally equal amplitudes, and takes the sum and difference of the quadrature signals, yielding

$$S_{c+s} = S'_{\cos} + S'_{\sin}, \quad (6.57)$$

$$= 2 \cos\left(\frac{1}{2}\delta'\right) \cos\left(\varphi + \frac{1}{2}\delta'\right), \quad (6.58)$$

$$S_{c-s} = S'_{\cos} - S'_{\sin}, \quad (6.59)$$

$$= 2 \sin\left(\frac{1}{2}\delta'\right) \sin\left(\varphi + \frac{1}{2}\delta'\right), \quad (6.60)$$

where the quadrature error has been eliminated at the cost of reintroducing a difference between the amplitudes of the signals. A constrained ellipse fit may then be employed to fit an ellipse centred on the origin, with the major and minor axes aligned with the data axes, in order to correct for the amplitude difference introduced by the sum and difference process. This approach is however vulnerable to differences in the signal amplitudes before taking the sum and difference.

An ellipse fit constrained as described may be made by applying least squares fit of the form

$$1 = \cos^2\left(\varphi + \frac{1}{2}\delta'\right) + \sin^2\left(\varphi + \frac{1}{2}\delta'\right), \quad (6.61)$$

$$= \left(\frac{S_{c+s}}{2 \cos\left(\frac{1}{2}\delta'\right)}\right)^2 + \left(\frac{S_{c-s}}{2 \sin\left(\frac{1}{2}\delta'\right)}\right)^2, \quad (6.62)$$

$$= AS_{c+s}^2 + BS_{c-s}^2. \quad (6.63)$$

Equation 6.63 may be written in matrix form as

$$\mathbf{M}_{\pm} \begin{pmatrix} A \\ B \end{pmatrix} = \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix}, \quad (6.64)$$

where $\mathbf{M}_{\pm}$ is the $2 \times N$ data matrix

$$\mathbf{M}_{\pm} = \begin{pmatrix} S_{c+s,0}^2 & S_{c+s,1}^2 & \cdots & S_{c+s,N}^2 \\ S_{c-s,0}^2 & S_{c-s,1}^2 & \cdots & S_{c-s,N}^2 \end{pmatrix} \quad (6.65)$$

and the right hand side is a $N \times 1$ vector of ones for a fit made to N data points. Solving this matrix equation following the ordinary least squares method results in

closed form solutions for the fit coefficients A and B given by

$$A = \frac{(\sum S_{s-c}^2)(\sum S_{s-c}^2 S_{s+c}^2) - (\sum S_{s+c}^2)(\sum S_{s-c}^4)}{(\sum S_{s-c}^2 S_{s+c}^2)^2 - (\sum S_{s+c}^4)(\sum S_{s-c}^4)}, \quad (6.66)$$

$$B = \frac{(\sum S_{s+c}^2)(\sum S_{s-c}^2 S_{s+c}^2) - (\sum S_{s-c}^2)(\sum S_{s+c}^4)}{(\sum S_{s-c}^2 S_{s+c}^2)^2 - (\sum S_{s+c}^4)(\sum S_{s-c}^4)}, \quad (6.67)$$

where the index of sums, which is not shown for clarity of notation, runs over the full range of the data points to be included in the fit. Equivalently, the sums may be replaced with averages, for example to avoid fixed point overflows, or with moving averages covering a fixed displacement length where the technique is to be applied to displacements covering multiple fringes. Both fit coefficients have a common divisor, which may be neglected as the absolute radius of the ellipse does not effect the phase measurement process, and as a result the fit coefficients may be expressed as

$$A' = \left(\sum S_{s+c}^2 \right) \left(\sum S_{s-c}^4 \right) \quad (6.68)$$

$$- \left(\sum S_{s-c}^2 \right) \left(\sum S_{s-c}^2 S_{s+c}^2 \right), \quad (6.69)$$

$$B' = \left(\sum S_{s-c}^2 \right) \left(\sum S_{s+c}^4 \right) \quad (6.70)$$

$$- \left(\sum S_{s+c}^2 \right) \left(\sum S_{s-c}^2 S_{s+c}^2 \right). \quad (6.71)$$

The sum and difference signals may then be corrected as

$$C_{s+c} = \sqrt{A'} S_{s+c}, \quad (6.72)$$

$$C_{s-c} = \sqrt{B'} S_{s-c}. \quad (6.73)$$

Constrained Heydemann technique

The second possible quadrature phase correction approach, termed here the CH technique, deals with the unequal amplitude problem by again making a constrained ellipse fit, fitting an ellipse with variable signal amplitudes and quadrature phase errors, but constrained to be centred on the origin.

The CH approach makes a least squares fit of the form

$$1 = C S_{\cos}^2 + E S_{\sin}^2 + F S_{\cos} S_{\sin}. \quad (6.74)$$

The form of equation 6.74 may be derived following the same approach as for equation 6.63, solving

$$S_{\cos} = \alpha \cos(\varphi), \quad (6.75)$$

$$S_{\sin} = \beta \sin(\varphi + \delta'), \quad (6.76)$$

for the sine and cosine of φ , and taking the sum of the square of sine and cosine to be equal to one. Coefficients α and β have been introduced to account for the non-ideal signal amplitudes.

Closed form solutions may also be found for the fit coefficients in this case, however the form is more complex and will not be described here. Instead, the fit coefficients C , E and F will be calculated following the ordinary least squares method as

$$\begin{pmatrix} C \\ E \\ F \end{pmatrix} = (\mathbf{M}^T \mathbf{M})^{-1} \mathbf{M}^T \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix}, \quad (6.77)$$

where

$$\mathbf{M} = \begin{pmatrix} S_{\cos,0}^2 & S_{\cos,1}^2 & \cdots & S_{\cos,N}^2 \\ S_{\sin,0}^2 & S_{\sin,1}^2 & \cdots & S_{\sin,N}^2 \\ S_{\cos,0} S_{\sin,0} & S_{\cos,1} S_{\sin,1} & \cdots & S_{\cos,N} S_{\sin,N} \end{pmatrix}. \quad (6.78)$$

Following the methodology of Heydemann [87] whilst accounting for the fact that the signals are now constrained to be centred on the origin, the amplitude ratio of the quadrature signals and the sine and tangent of the quadrature phase error may be calculated from the fit coefficients as

$$G_{xy} = \sqrt{\frac{E}{C}}, \quad (6.79)$$

$$\tan(\delta') = \frac{-F}{\sqrt{4CE - F^2}}, \quad (6.80)$$

$$\sin(\delta') = \frac{-2\sqrt{CE}}{\sqrt{4CE - F^2}}, \quad (6.81)$$

and the corrected quadrature signals may be recovered as

$$C_{\cos} = S_{\cos}, \quad (6.82)$$

$$C_{\sin} = G_{xy} \sin(\delta') S_{\sin} + \tan(\delta') S_{\cos}. \quad (6.83)$$

6.4.3 Experimental methods

In order to test both the MIR and CH techniques, an interferometer was constructed following the schematic representation shown in figure 6.19. Laser illumination was provided by a fibre coupled frequency stabilised helium neon (HeNe) laser (REO Model 32734) operating at approximately 632.8 nm. The fibre coupled laser was then collimated and polarised with a Glan-Thompson polariser (extinction ratio 100000:1), ensuring a stable polarisation state at the input to the interferometer. A custom phase-quadrature beam splitter (PQBS) [25] cube formed the core of the interferometer, splitting the beam into the measurement and reference arms of the interferometer, and introducing an approximately 90° quadrature phase shift between the interferometer outputs. The quadrature phase shift introduced by the PQBS is polarisation dependent, and as sub-fringe fluctuations in the quadrature phase shift are not corrected for with the techniques described here, the high extinction ratio input polariser is required to minimise non-linearities in the interferometer. Solid glass cube corner retroreflectors were employed in both arms of the interferometer, with the arm lengths approximately balanced to within millimetres. The measurement arm retroreflector, RR_{mes} , was mounted on a Queensgate Instruments

NPS-X-15A nanopositioning stage, with a manufacturer quoted position noise of 50 pm RMS, allowing precise path length changes to be generated. Plate beam splitters were used in both the reference and measurement arms, with a 50:50 reflectivity to transmission ratio, and beam splitters marked as being from the same coating batch were selected as both the MIR and CH techniques assume identical beam splitter properties in each arm of the interferometer. All optical components with planar optical surfaces were tilted slightly to minimise any multiple reflection effects, whilst maintaining parallelism between the measurement arm beams and stage displacement axis. In order to ensure consistent power collection at each of the four detectors, the output beams of the interferometer were focused onto the photodiode detectors (Hamamatsu S2386-18K/L) with 40 mm focal length lenses (not shown in figure 6.19). Apertures were placed before each lens set to a diameter that did not visible clip the beams in order to minimise the detection of stray light. The photodiodes were operated in photovoltaic mode to minimise dark currents, with the photocurrent amplified by custom electronics. The amplified signals were digitised with 16 bit analogue to digital converters (National Instruments NI7833R) operating at a sampling frequency of 200 kHz, and, for the purposes of this work, were logged for post processing.

For both the MIR and CH corrections identical calibration steps are required. These were performed in three steps. First, a beam block was placed between the input polariser and PQBS, and the dark voltage signals ($D_{\cos}$, $D_{\sin}$, D_{ref} and D_{mes}) recorded. The beam block was then moved to between the reference retroreflector (RR_{ref}) and the reference arm beam splitter, and the reference blocked voltage signals ($V_{\cos}^{(\text{R})}$, $V_{\sin}^{(\text{R})}$, $V_{\text{ref}}^{(\text{R})}$ and $V_{\text{mes}}^{(\text{R})}$) were recorded. Finally, the beam block was moved to between the measurement arm retroreflector and the measurement arm beam splitter, and the measurement blocked voltage signals ($V_{\cos}^{(\text{M})}$, $V_{\sin}^{(\text{M})}$, $V_{\text{ref}}^{(\text{M})}$ and $V_{\text{mes}}^{(\text{M})}$) were recorded. For each calibration step 100000 data points were recorded synchronously for all four signals. Ratios and products of ratios were then calculated according to equations 6.53 to 6.56 using the full 100000 data points. Averages were then taken for each ratio, and the averaged ratios used to correct the measured signals following equations 6.38 and 6.39. By taking averages only after the ratios were calculated, the calibration process was made insensitive to laser power fluctuations between each calibration step, and any bias introduced by taking ratios of averaged random variables was eliminated.

At this point, the two proposed techniques diverge. Considering first the MIR technique, sum and difference signals were calculated from equations 6.57 and 6.60. Fit coefficients were then calculated according to equations 6.68 and 6.71, and the sum and difference signals were corrected using equations 6.72 and 6.73. In the case of the CH technique, fit coefficients C , E and F were calculated from equation 6.77, and the correction was applied following equations 6.79 to 6.83.

Following signal correction, the interferometric phase was, for both techniques, calculated using the four quadrant arctangent function as

$$\varphi_{\text{CH}} = \text{atan2}(C_{\sin}, C_{\cos}) \quad (6.84)$$

and

$$\varphi_{\text{MIR}} = \text{atan2}(C_{\text{s+c}}, C_{\text{s-c}}) \quad (6.85)$$

for the CH and MIR corrections respectively. The corresponding displacement was then calculated for both techniques from a change in the interferometric phase as

$$x = \frac{\lambda}{4\pi} \Delta\varphi. \quad (6.86)$$

For the purposes of this work any dead path effects [49] introduced by path length imbalances at the start of the measurement were not considered, as path length imbalances were minimal, and any dead path effects present do not affect the non-linearity of the interferometer, which is the focus of this work. Additionally, the effects of refractive index fluctuations [92], were not considered, as such effects do not affect the ability of the interferometer system described here to resolve non-linearities, instead having only a negligible scaling effect on the observed non-linearity.

In order to test the correction techniques, the stage was stepped through the full 15 μm closed loop range in uniform steps. At each stage position V_{cos} , V_{sin} , V_{ref} and V_{mes} were recorded, with a 100 point moving average applied to reduce the effects of noise and higher frequency mechanical vibrations. Residual non-linearities were evaluated by comparison to the stage set position, with the displacement calculated from the stage set positions scaled to cover the same range as the interferometric displacement measurements, minimising the effects of angular misalignments between the stage and interferometer measurement axes on the non-linearity measurement. The differences between the interferometer displacements and the stage displacements were then calculated. In order to separate residual non-linearities of tens of picometres from the combined effects of air turbulence, mechanical noise, and electrical noise, a Fourier transform of the calculated displacement differences was employed. The calculated displacement differences were windowed with a Hann window function [185] to reduce spectral leakage, and zero padded, with a number of zeros equal to the number of data points adjoined to each end of the data. To compensate for the amplitude scaling effects of the window function, the amplitudes resulting from the Fourier transform were multiplied by a scaling factor calculated as the number of displacement difference data points divided by the sum over all values in the window function. For comparison, displacements corrected with the Heydemann correction [87] as implemented by Birch [21] were also calculated, along with displacements corrected with the optical shuttering technique of [169] in some cases. In order to minimise the impact of laser intensity fluctuations the quadrature signals were normalised to reference arm intensity reference signal before applying the Heydemann correction.

6.4.4 Results

Full-fringe case

First, the full-fringe case (the case for which a full ellipse of data is available for correction) was considered, for a displacement covering a range of 15 μm in 15 nm steps. A comparison of the Heydemann [87], beam shuttering [169], MIR and CH

algorithms is shown in figure 6.21. Residual twice per fringe non-linearities (having a spatial period of 158 nm) can be seen for the MIR algorithm, with an amplitude of approximately 9 pm. For both the MIR and CH algorithms a small improvement in performance can be seen over the Heydemann algorithm at a spatial period of 316.4 nm. This effect appears to be a result of the correction for variable in-arm losses made for both corrections, whereas the Heydemann correction can only compensate for laser intensity fluctuations affecting both arms. The beam shuttering technique [169] yields residual twice per fringe non-linearities with an amplitude of 65 pm for the interferometer used in this work. These residual non-linearities are in this case a result of the shuttering technique not correcting for quadrature phase shifts. A small non-linearity component with an amplitude of 8 pm can be seen at a spatial period of 105.5 nm for all algorithms, corresponding to $\lambda/6$, which in this interferometer system appears to be introduced by internal reflections within the solid glass corner cube retroreflector. Smaller higher order harmonics are also likely to be present due to this multiple reflection effect, however they cannot be separated from the noise floor in this system.

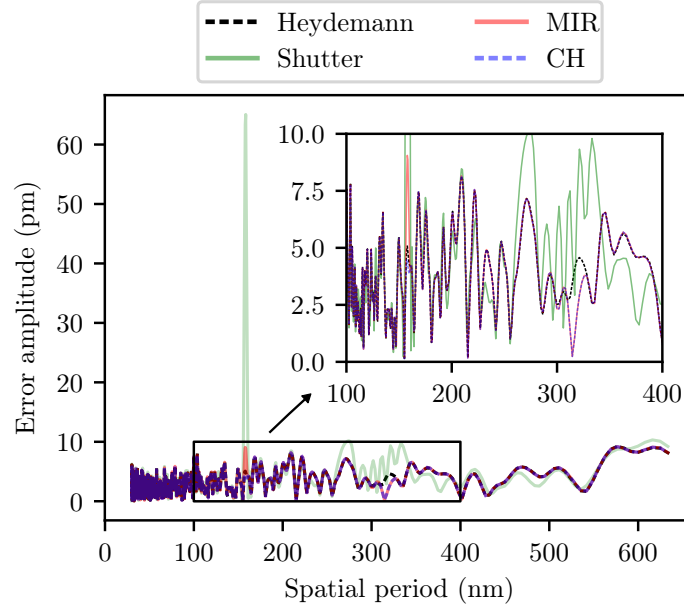


Figure 6.21: Comparison of the full-fringe performance of the Heydemann [87], beam shuttering [169], multiple intensity reference (MIR) and constrained Heydemann (CH) non-linearity correction algorithms when applied to a displacement covering 15 μm with a sample spacing of 15 nm. Inset shows region of interest. The Heydemann, MIR, and CH techniques overlap almost exactly away from harmonics of 632.8 nm.

The effects of wavefront distortions on the MIR algorithm, motivating the development of the CH algorithm, are demonstrated in figures 6.22 and 6.23. In this case some clipping was introduced into the reference arm of the interferometer, with the reference beam falling on a joint between the reflecting surfaces in the reference arm corner cube retroreflector. The high sensitivity of the beam shuttering technique to wavefront distortions can be seen. The large non-linearities seen in this case

may also be a result of the sensitivity of the beam shuttering technique to changes in laser intensity between shuttering steps, and between the shuttering steps and measurement taking place. For this reason, the beam shuttering technique is not included in further comparisons. A non-linearity term can also be seen to arise with a spatial period of 158 nm and an amplitude of approximately 50 pm for the MIR algorithm, resulting from imperfect amplitude correction by equations 6.38 and 6.39. Similar increases in non-linearity were regularly observed whilst testing the MIR algorithm in cases where wavefront distortions were introduced in either arm of the interferometer before recombination of the beams.

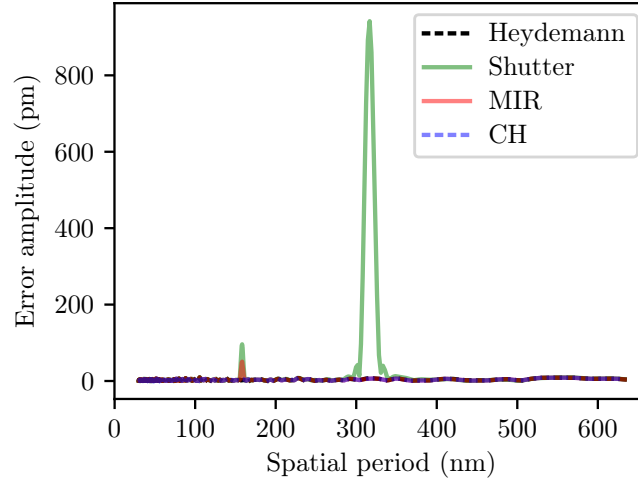


Figure 6.22: Comparison of the full-fringe performance of the Heydemann [87], beam shuttering [169], multiple intensity reference (MIR) and constrained Heydemann (CH) non-linearity correction algorithms in the presence of wavefront distortion, when applied to a displacement covering $15\text{ }\mu\text{m}$ with a sample spacing of 15 nm. All techniques plotted overlap almost exactly away from harmonics of 632.8 nm.

The full-fringe case was also utilised to investigate the stability of the calibration process. A single set of calibration measurements were made followed by a series of $15\text{ }\mu\text{m}$ displacements, again made in 15 nm steps, with the displacement measurement repeated every 2.5 min. The previously described Fourier method was then applied, and the values of the Fourier amplitudes at spatial periods of 316.4 nm and 158.2 nm (corresponding to $\lambda/2$ and $\lambda/4$) added for the Heydemann, MIR and CH techniques at each time point. The amplitude sums were then doubled to give an estimate of the peak-to-peak non-linearity, assuming perfect constructive interference between the amplitude terms, with the results shown in figure 6.24, along with a 1 h moving average of the resulting non-linearity estimates. Moving averages were calculated based upon a 1 h window centred on the plot time point, with any portion of the window falling outside of the data set neglected.

Sub-fringe case

The development of the MIR and CH algorithms was driven by the desire for improved non-linearity corrections in the case where displacements cover less than a full

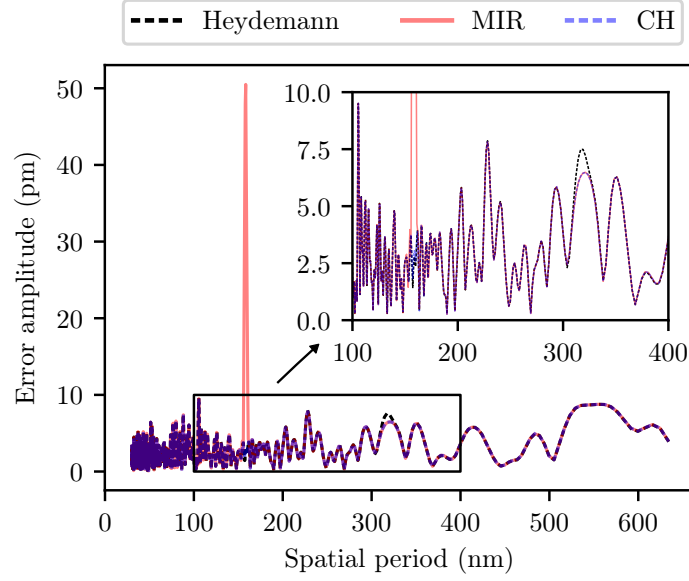


Figure 6.23: Comparison of the full-fringe performance of the Heydemann [87], multiple intensity reference (MIR) and constrained Heydemann (CH) non-linearity correction algorithms in the presence of wavefront distortion, when applied to a displacement covering $15\mu\text{m}$ with a sample spacing of 15 nm . Inset shows region of interest. All techniques plotted overlap almost exactly away from harmonics of 632.8 nm .

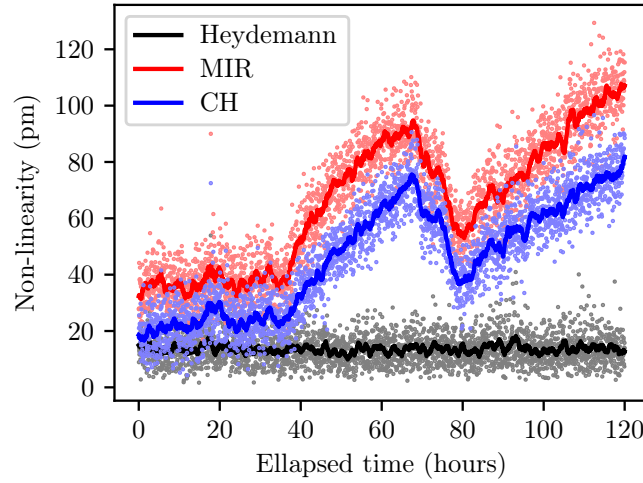


Figure 6.24: Time stability of MIR and CH calibration process, compared to the Heydemann correction in the full-fringe case. Both individual measurements taken every 2.5 min and 1 h moving averages are shown. Estimated peak to peak non-linearities are plotted, considering only non-linearity components with spatial periods of 316.4 nm and 158.2 nm . The laboratory remained within $(20 \pm 1)^\circ\text{C}$ throughout the measurements.

optical fringe. As such, the sub-fringe performance of the MIR and CH algorithms

was compared to that of the Heydemann correction. To test the sub-fringe performance, a displacement through a $15\text{ }\mu\text{m}$ range was performed in steps of 1.5 nm . A number of data points corresponding to the displacement range indicated on the x axis of figure 6.25 were taken from the start of the full $15\text{ }\mu\text{m}$ range, and the correction coefficients (A and B for the MIR algorithm; C , E and F for the CH algorithm and the full ellipse fit for the Heydemann algorithm) calculated based on this restricted data set. The full $15\text{ }\mu\text{m}$ displacement was then corrected based upon the calculated coefficients, and the previously described Fourier technique used to estimate non-linearities, as for figure 6.24. This process was then repeated 100 times, with randomly generated offsets applied to the restricted data sets for each repeat such that the restricted data sets were taken from different parts of the full displacement range. Averages and standard errors were then calculated for each fit range, with the results plotted in figure 6.25. This approach was required as the Fourier technique used to measure non-linearities in this work requires multiple fringes to give reliable results, and cannot be applied to sub-fringe measurements directly. An important consequence of this approach is that the stability of the non-linearities is also implicitly measured, along with the performance of the algorithms.

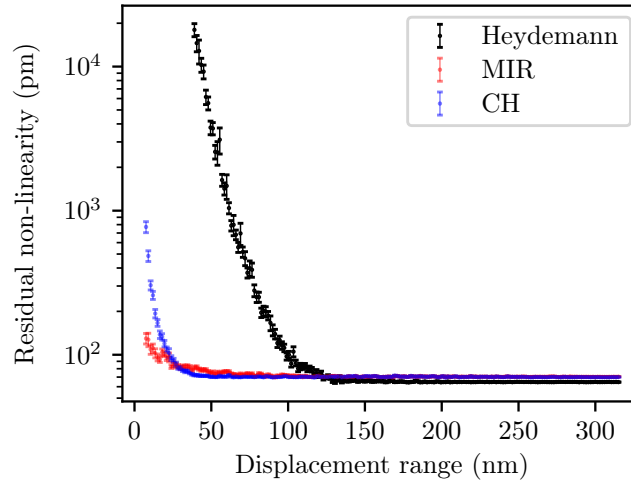


Figure 6.25: Sub-fringe performance of the Heydemann, MIR and CH algorithms. Estimated peak to peak non-linearities are shown, considering only non-linearity components with spatial periods of 316.4 nm and 158.2 nm calculated as described in the text. Error bars indicate standard uncertainties.

It should be noted that the data presented in figure 6.25 shows a higher large displacement non-linearity limit for all three algorithms than may be expected from figure 6.21. This is a consequence of two effects. Firstly, the decreased displacement step size required, which, in combination with the poor mechanical stability of this proof of concept interferometer, and the Fourier non-linearity measurement technique, resulted in a higher ‘noise floor’, introduced primarily by mechanical drifts. Secondly, the stability of the ellipse fit parameters across the full displacement range also influences the measured residual non-linearity.

6.4.5 Discussion

In this work, the MIR and CH algorithms have been compared to the commonly employed ellipse fitting Heydemann correction for both full-fringe and sub-fringe displacements. Considering first the full fringe results, figure 6.21 demonstrates that the CH algorithm is capable of reaching the same non-linearity performance as the Heydemann correction, with a small increase in performance for once per fringe non-linearities possible. The MIR algorithm has been shown to have residual non-linearities with a twice per fringe spatial period, in this case having an amplitude of 9 pm. As demonstrated in figure 6.23, the MIR algorithm is highly sensitive to any effect that results in incomplete interference between the beams returned from each arm, with twice per fringe non-linearities rising to 50 pm in the presence of wavefront distortions from a solid glass corner cube seam. This sensitivity to incomplete interference is a result of the assumption made by the MIR technique that the amplitudes of the quadrature signals may be determined from the intensity reference outputs, an assumption that is only valid in the ideal case of identical wavefronts being returned by each interferometer arm, with identical polarisation states, and with both beams being fully coherent. As a result, some residual non-linearities are likely to be present after applying the MIR correction in most practical interferometer designs, where incomplete interference may be introduced by wavefront distortions in the optics, misalignment between the interfering beams, or polarisation aberrations. In the interferometer described in this work the residual twice per fringe non-linearities after applying the MIR algorithm seen in figure 6.21 are likely to be a result of the cumulative effect of small angular alignment errors between the interfering beams introduced by imperfect cube corner retroreflectors, and polarisation aberrations introduced by reflections from the silver coated surfaces of the retroreflectors. Whilst in this case residual non-linearities introduced by imperfect interference were reduced to 9 pm in amplitude through careful alignment, this may not be possible in a system where non-linearities cannot be evaluated in-situ, and as a result, these highly sensitive residual non-linearity sources motivated the development of the CH approach.

Whilst figure 6.21 demonstrates the best case performance achieved with this system, with the displacement measurement made immediately after calibration, the stability of the calibration process is also important if the MIR or CH techniques are to be used in practical applications. From figure 6.24 it can be seen that the calibration remained stable over time scales of a few hours, with the residual non-linearities beginning to increase for both the MIR and CH algorithms after 35 h. Some caution should be employed when drawing general conclusions from the data presented in figure 6.24, as the stability of the calibration is dependent on a range of factors that may vary between interferometer systems, and between laboratories. The first effect to be considered is the mechanical stability of the optical system. Whilst variable losses in each arm of the interferometer are accounted for by both the MIR and CH algorithms, variable losses between the pick off beam splitters and intensity reference detectors, or between the phase-quadrature beam splitter and the quadrature detectors are not corrected. In part, this may be compensated by ensuring the optical path lengths from the fibre collimator to the four detectors are equalised, however changes in optical alignment that affect only one detector

will still result in an increase in residual non-linearities. The interferometer used in this work was assembled as a proof of concept, and as such greater mechanical stability may be expected from an interferometer designed for practical measurement tasks. Mechanical stability will also be influenced by the thermal stability of the environment. In this case the temperature remained within $\pm 0.25^\circ\text{C}$ over the course of the measurements presented in figure 6.24. In addition to the mechanical stability of the system, the stability of the amplification electronics may affect the residual non-linearities. In particular, both the MIR and CH algorithms assume the voltage offsets in the absence of laser illumination are constant, and changes to these offsets, or changes in the amplifier gains, will result in increased residual non-linearities. The intensity stability of the laser source is not critical for either the MIR or CH algorithms, as both algorithms correct for any change in source intensity.

In comparison to the beam shuttering technique previously proposed in the literature [169] both the MIR and CH techniques have been demonstrated in figure 6.22 to be less sensitive to incomplete interference effects. Additionally, the beam shuttering technique does not correct for quadrature phase shift errors, limiting the ultimate performance for the interferometer system described in this work, as shown in figure 6.21. The beam shuttering technique as described in [169] also does not include an intensity reference, and if implemented in this manner the correction is sensitive to changes in the laser intensity, both with regards to high frequency fluctuations and low frequency drift.

The main advantage of the MIR and CH algorithms over the Heydemann correction is demonstrated in figure 6.25; for displacements ranges of less than approximately 120 nm (for a single pass interferometer operating at 632.8 nm) residual non-linearities are reduced. In this work, an ordinary least squares fit following the methodology described in [21] was employed by the Heydemann correction, which may not be optimal for fitting small portions of an ellipse [163], however, even where enhanced ellipse fitting algorithms are employed, residual non-linearities in excess of 100 pm have been shown to arise for displacements covering less than one half of an optical fringe [163]. Similarly, ordinary least squares fits were used by both the MIR and CH algorithms when making constrained ellipse fits, and performance improvements may be possible with more advanced fitting algorithms.

Whilst in this work the CH algorithm has been demonstrated to be more robust than the MIR algorithm, in applications where near complete interference can be guaranteed, it can be seen from figure 6.25 that the MIR algorithm outperforms the CH algorithm for displacements of less than approximately 25 nm. This cross over point is likely to be highly dependent on noise levels in the interferometer system, however in general, for applications where displacements cover only a few nanometres, the MIR algorithm is likely to outperform both the CH and Heydemann algorithms. The relative computation simplicity of the fit coefficient calculations in the case of the MIR algorithm is also a significant advantage, and future work may investigate the possibility of implementing this algorithm in real time on a field programmable gate array (FPGA) device.

6.4.6 Conclusion

Two novel non-linearity correction algorithms have been proposed, both aiming to compensate for non-linearities where homodyne interferometers are applied to displacements covering less than one full optical fringe. The first, the MIR technique, is relatively mathematically simple, and may be suitable for real time implementation, however the performance has been shown to be limited in cases where incomplete interference occurs. The second approach, the CH technique, was developed for applications where incomplete interference does occur, and has been demonstrated to be capable of achieving residual non-linearities comparable to that of the Heydemann correction, whilst remaining effective down to small fractions of an optical fringe, in this work outperforming the Heydemann correction for displacements of less than approximately 120 nm. For the interferometer described here, non-linearity peaks at all harmonics of the illuminating wavelength (632.8 nm) were reduced to below 10 pm in amplitude by both techniques under ideal conditions.

Both the MIR and CH techniques have been compared to a beam shuttering technique previously described in the literature [169], and have been found to be less sensitive to incomplete interference and quadrature phase shift errors. The MIR and CH techniques do however require additional optical components and detection electronics as compared to the beam shuttering approach, and this may be a disadvantage in cost sensitive applications.

Future work may implement the MIR technique in real time on an FPGA device, and explore the potential for the CH algorithm to be applied to atomic force microscope cantilever measurements. Further work may also compare the performance of the algorithms to the NPL X-ray interferometer [54], permitting sub-fringe non-linearities to be evaluated directly, without recourse to a Fourier technique.

6.5 Measurement arm intensity reference normalisation

A simple non-linearity compensation approach is presented in this section, based upon normalisation of the quadrature outputs to an intensity reference taken from the return path of the measurement beam of the interferometer. This approach allows for the partial correction of non-linearities introduced by higher order harmonics, with complete correction possible under certain conditions. This section describes the implementation of the measurement arm normalisation technique for the modified NPL plane mirror differential interferometer design shown in figure 6.26. This interferometer design is analysed in chapter 7, and as such more detailed discussion of the performance of the interferometer design, including the performance of the measurement arm normalisation technique, is reserved for chapter 7.

As previously discussed, non-linearities introduced by errors in the quadrature phase difference, DC offset and gain ratio of the quadrature outputs produced by homodyne interferometers can be corrected for with the ellipse fitting Heydemann correction [21, 87], however, non-linearities introduced by the presence of higher

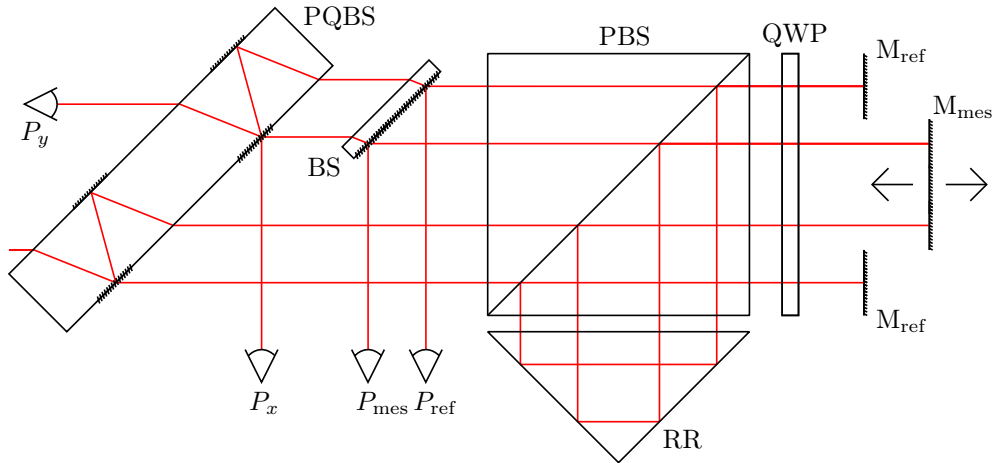


Figure 6.26: Schematic of NPL plane mirror differential interferometer design with added intensity references. PQBS, phase-quadrature beam splitter; PBS, polarising beam splitter; QWP, quarter waveplate; M_{ref} , reference mirror; M_{mes} , measurement mirror.

order harmonics in the quadrature outputs are not corrected for by the Heydemann correction. Such higher order harmonics may arise in the quadrature outputs through polarisation leakage or unwanted back reflections within the interferometer optics [84, 85]. Interferometers with single digit picometre level non-linearities have been demonstrated [11], however, as higher order non-linearities are strongly dependant on the exact alignment, or misalignment, of the interferometer optics, guaranteeing such performance for a given interferometer design is challenging [8]. Higher order non-linearities may be corrected with the two-wavelength technique described in section 6.3 [2WL], however, the two-wavelength approach introduces significant additional electronic and optical complexity, and requires a displacement range of greater than one optical fringe to be effective. Other approaches for the correction of higher order non-linearities are therefore desirable.

As demonstrated in section 3.3.3, higher order non-linearities cannot be corrected based only on the quadrature signal outputs of a homodyne interferometer. Such a correction is not possible, as an interferometer having higher order non-linearities may produce quadrature signals that form a perfectly circular Lissajous figure. This Lissajous figure, arising from a linear displacement, would be indistinguishable from that produced by a linear displacement measured with a non-linearity free interferometer. Some additional information is therefore required to correct for higher order non-linearities in homodyne interferometry. In the two-wavelength case, this additional information comes in the form of a second displacement measurement. In the case of the measurement arm normalisation technique described in this section, the additional information instead comes in the form of an intensity reference taken from the return path of the measurement arm of a homodyne interferometer.

In homodyne interferometry, normalisation of the quadrature signals to a laser intensity reference is desirable to minimise the effects of fluctuations and drift in the illuminating intensity, which may be present due to intensity instability in the

laser source or fibre coupling system, if the interferometer is fibre fed. Typically, this intensity reference is taken at the input to the interferometer, before the beam is split into the interferometer arms, however, in figure 6.26, two intensity references are taken separately from the measurement (P_{mes}) and reference (P_{ref}) arms of the interferometer. Any unwanted harmonics that affect the interferometric phase measurement must involve a reflection from the moving measurement mirror, and as such will appear in P_{mes} as a term interfering with the desired measurement arm beam. Only information regarding interference between the measurement arm beam and the unwanted harmonics is captured, with the results of interference between the unwanted harmonics and the reference arm beam being lost, and therefore in general a full correction cannot be expected. The effectiveness of the measurement arm intensity normalisation process may be investigated for individual interferometer designs with the combined scattering and transfer matrix modelling technique [85].

In order to normalise the quadrature signals the amplified voltage signals from each detector V_x , V_y , V_{mes} and V_{ref} , corresponding to outputs P_x , P_y , P_{mes} and P_{ref} , were taken, along with measurements of the dark voltages introduced by the detection and amplification electronics in the absence of laser illumination D_x , D_y , D_{mes} and D_{ref} . The dark voltages were then subtracted from each voltage signal, and the quadrature voltage signals normalised as

$$S_{x,\text{mes}} = \frac{V_x - D_x}{V_{\text{mes}} - D_{\text{mes}}}, \quad (6.87)$$

$$S_{y,\text{mes}} = \frac{V_y - D_y}{V_{\text{mes}} - D_{\text{mes}}}. \quad (6.88)$$

For the purposes of comparison, reference arm normalised signals were also calculated as

$$S_{x,\text{ref}} = \frac{V_x - D_x}{V_{\text{ref}} - D_{\text{ref}}}, \quad (6.89)$$

$$S_{y,\text{ref}} = \frac{V_y - D_y}{V_{\text{ref}} - D_{\text{ref}}}. \quad (6.90)$$

Whilst the dark voltage subtraction is included here for completeness, in the experimental work presented in this section the dark voltages were of the order of millivolts, performing, or not performing, the subtraction had no measurable effect on the non-linearity.

Following normalisation, a Heydemann correction [21, 87] was applied to compensate for offset, gain ratio and quadrature phase shift errors, resulting in corrected signals $C_{x,\text{mes}}$ and $C_{y,\text{mes}}$, and $C_{x,\text{ref}}$ and $C_{y,\text{ref}}$. The instantaneous interferometric phase was then calculated with the four quadrant arctangent function as

$$\varphi_{\text{mes}} = \text{atan2}(C_{y,\text{mes}}, C_{x,\text{mes}}), \quad (6.91)$$

$$\varphi_{\text{ref}} = \text{atan2}(C_{y,\text{ref}}, C_{x,\text{ref}}). \quad (6.92)$$

In both the measurement arm and reference arm normalised cases the phase was unwrapped at discontinuities of greater than π [22], and the corresponding displacement between two points was calculated from the change in phase between two

points ($\Delta\varphi_{\text{mes}}$ and $\Delta\varphi_{\text{ref}}$) as

$$x_{\text{mes}} = \frac{\lambda}{8\pi} \Delta\varphi_{\text{mes}}, \quad (6.93)$$

$$x_{\text{ref}} = \frac{\lambda}{8\pi} \Delta\varphi_{\text{ref}}, \quad (6.94)$$

for a wavelength λ of 632.8 nm.

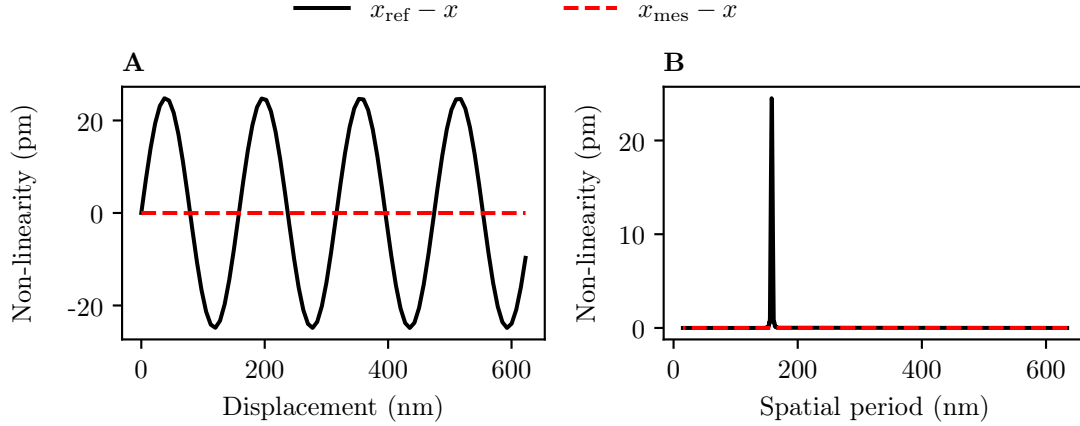


Figure 6.27: Non-linearity resulting from introducing a $\lambda/300$ retardance error in the QWP, combined with a 1000:1 extinction ratio in the PBS, of the interferometer depicted in figure 6.26, for an illuminating wavelength of 632.8 nm, after applying a Heydemann correction.

The non-linearity resulting from a $\lambda/300$ retardance error in the quarter waveplate, a common retardance specification for commercially available optics, combined with a 1000:1 extinction ratio in the PBS, as modelled with the combined scattering and transfer matrix modelling technique [85], is shown in figure 6.27 for both the measurement (x_{mes}) and reference (x_{ref}) arm intensity reference normalised cases, relative to the true displacement x . Whilst the combined scattering and transfer matrix model will tend to overestimate non-linearities, as perfect alignment of higher order reflections is assumed, the 50 pm peak-to-peak non-linearity error shown in figure 6.27 in the reference arm normalised case is a significant error source compared to the sub 10 pm non-linearities achieved by leading interferometer designs [11]. This non-linearity can be seen to be completely eliminated in the measurement arm normalised case. Non-linearities arising from QWP errors are not corrected by the Heydemann correction, and as such, reducing these non-linearities using previously published techniques requires either a more advanced correction methodology [69], introducing significant additional optical and electronic complexity, or improving the QWP retardance accuracy. Increasing QWP retardance accuracy increases the cost of the system, in particular for the large aperture waveplate required by this interferometer design.

As can be seen in figure 6.28, the measurement arm normalisation technique

cannot correct for all higher order non-linearity sources. In this case, 0.04 % back reflections have been introduced at the mirror facing surface of the QWP, and at the PBS facing surface of the QWP in the return path of both the measurement and reference arms of the interferometer. 0.28 % back reflections have been introduced at the P_y , P_{mes} and P_{ref} interferometer outputs, modelling the effects of reflection from the photodiodes, with this particular combination of back reflections selected in an attempt to model experimental results. Non-linearity errors for this case are shown in spatial frequency space in figure 6.28 **B**, where it can be seen that in this case, the measurement arm normalisation technique increases the amplitude of the non-linearity component with a spatial period of 316.4 nm, demonstrating that some care must be taken when employing the measurement arm normalisation technique to an individual interferometer design, although an overall reduction in non-linearity is still achieved.

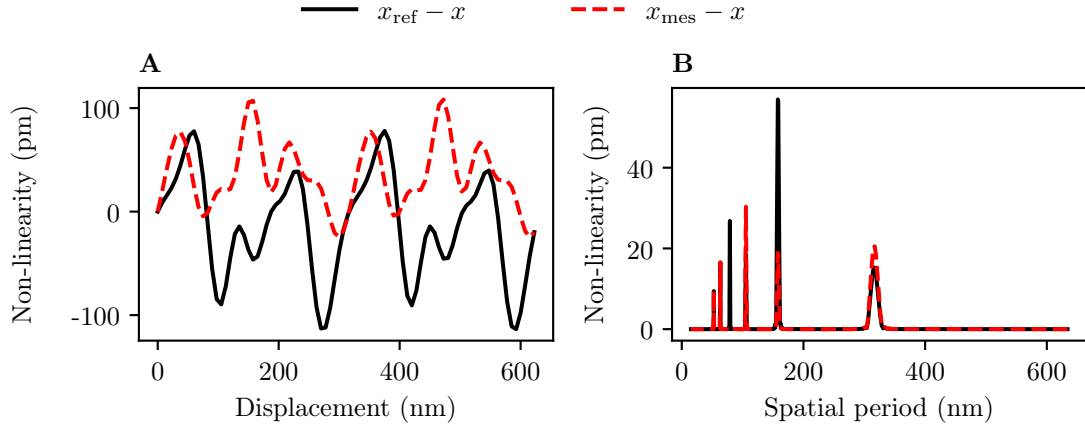


Figure 6.28: Modelled reference and measurement arm normalised non-linearities in displacement (**A**) and spatial frequency (**B**) space, calculated relative to true displacement x , including the effects of unwanted back reflections as described in the text.

The sensitivity of the measurement arm normalisation technique to unwanted back reflections is further demonstrated in figure 6.29, where back reflections 0.01 % from the four detectors are modelled, with all other conditions identical to those described for figure 6.27. Whilst non-linearities are again reduced relative to the reference arm normalised case, 11 pm peak-to-peak non-linearities are still present, highlighting the fact that even small back reflections, as may take place from an anti-reflection coated surface, may be a significant error source if they occur at some locations.

For comparison, the effects of a 0.01 % back reflection at the input to the interferometer were also modelled, with the results presented in figure 6.30. As can be seen from figure 6.30, this results in a significantly smaller non-linearity than a reflection of the same magnitude from the interferometer output, with the peak-to-peak non-linearity reduced to 0.03 pm.

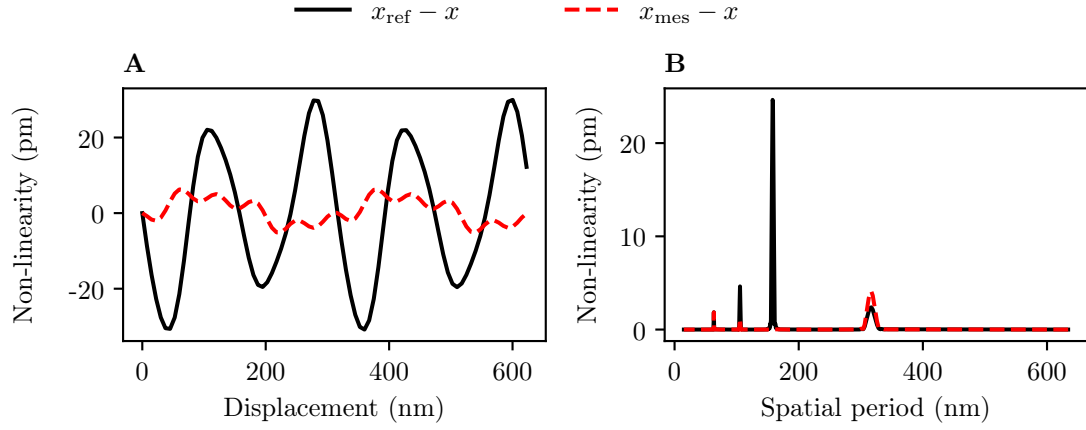


Figure 6.29: Modelled reference and measurement arm normalised non-linearities in displacement (**A**) and spatial frequency (**B**) space, calculated relative to true displacement x , including the effects of 0.01 % back reflections from all four detectors, with other conditions as described for figure 6.27.

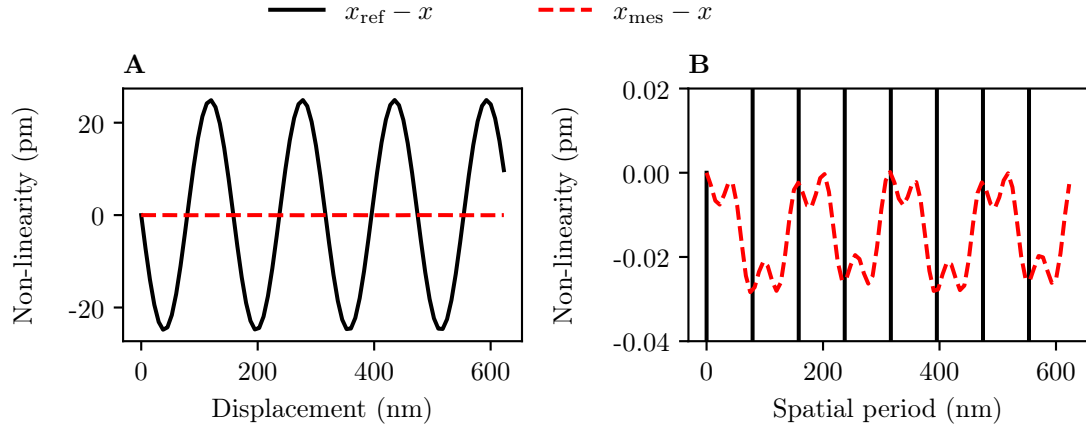


Figure 6.30: Modelled reference and measurement arm normalised non-linearities, calculated relative to true displacement x , including the effects of a 0.01 % back reflection from the interferometer input, with other conditions as described for figure 6.27. **B** shows the same data as **A**, with the ordinate axis magnified around the region of interest.

In this section, the measurement arm normalisation technique has been shown to be capable, in theory, of completely compensating for certain higher order non-linearity sources in homodyne interferometry. Whilst the measurement arm normalisation technique does have limitations, in particular the vulnerability to unwanted back reflections demonstrated in figure 6.28, the technique still provides an alternative to the two-wavelength non-linearity correction [69], that, in cases where the measurement arm normalisation technique is applicable, may be implemented with fewer additional optical and electronic components, and with reduced data processing complexity. Further work is needed to quantify the conditions under which a complete correction of higher order non-linearities is possible. A final consideration when applying the measurement arm intensity reference correction technique is that the symmetry between the reference and measurement arms of the interferometer is broken. Further work is therefore also required to characterise the implications of this break of symmetry, and to investigate whether the displacement measurement is still truly differential. Experimental results gathered with an interferometer specifically designed to utilise this technique are presented in chapter 7.

6.6 Summary

Within this chapter, four classes of non-linearity correction algorithms have been discussed in detail. The various implementations of the Heydemann correction [87] have been compared, followed by three novel corrections, all aiming to fill weaknesses in the Heydemann technique. Of these new approaches, the two-wavelength [69] and measurement arm intensity reference normalisation [172] corrections both make use of the Heydemann correction as a correction step, and the multiple intensity reference correction [69] makes use of closely related constrained ellipse fitting algorithms, demonstrating the importance of considering the benefits of different ellipse fitting algorithms.

Table 6.2: Summary of the non-linearity correction techniques discussed in this chapter. Displacement ranges are approximate, given to the nearest order of magnitude, and assume a single pass interferometer operating at 632.8 nm. It should be noted that the minimum displacement range is not a sharp cut off, and that the short range performance of all algorithms considered will be dependent on optical wavelength, interferometer pass number, and noise levels.

Correction technique	Higher order	Displacement range
Heydemann [87]	No	$\gtrsim 100$ nm
Two-wavelength [69]	Yes	$\gtrsim 1$ μ m
Multiple intensity reference [171]	No	$\gtrsim 10$ nm
Measurement arm normalisation [172]	Partial	Unlimited

Where a full fringe is available for correction, the ODR algorithm has been found to outperform the other algorithms compared in this work (presented in table 6.1),

although performance differences were found to become negligible as the number of points included in the fit increased. In the sub-fringe case, the DEF and BLUE algorithms were found to offer increased performance, although it is possible that an ellipse specific implementation of the ODR fit may improve performance in this regime. The ODR algorithm was found to provide the most accurate variance and covariance estimates of the algorithms compared in the full fringe case, however, the previously stated assumption that uncertainties in individual measurements of the instantaneous interferometric phase are uncorrelated was found to be incorrect. As such, the uncertainties calculated by all algorithms were found to become inaccurate as the fit coefficient uncertainty becomes a significant contribution. When following the uncertainty estimation approach described in section 3.3 this is not a problem, as the uncertainty introduced by the fit coefficients is accounted for in the experimental non-linearity measurement. In the sub-fringe case all the algorithms that were compared were found to produce inaccurate variance and covariance estimates, further demonstrating the requirement for experimental non-linearity measurement.

The first major weakness of the Heydemann correction is that, regardless of the ellipse fitting algorithms employed, Heydemann style corrections cannot correct for higher order non-linearities. Two algorithms that are capable of correcting for higher order non-linearities have been presented in this chapter, of which the measurement arm normalisation technique is the most straight forward to implement, requiring only that the intensity reference often already present in homodyne interferometer designs be moved to the measurement arm return path. This technique has been shown to theoretically offer a complete compensation for some higher order non-linearity sources, although notably the effects of back reflections cannot be fully corrected for. As such, in cases where modelling suggests measurement arm normalisation may offer a reduction in non-linearities, it may be implemented with few disadvantages. As an additional benefit, the measurement arm normalisation technique provides an indication of the presence of higher order non-linearities via the presence of a displacement dependent modulation in the measurement arm intensity reference output, however further work is required to quantify this effect. Further work is also required to identify the exact conditions under which the measurement arm normalisation technique produces a complete correction of higher order non-linearities.

Higher order non-linearities that cannot be compensated for by measurement arm normalisation may be corrected for using the two-wavelength non-linearity correction [69] in cases where the displacement covers multiple fringes. This technique has been shown to be effective for the correction of higher order non-linearities regardless of the optical source of the non-linearities, however implementing the two-wavelength correction requires introducing significant additional optical, electronic and data processing complexity. Beyond non-linearity correction, the two-wavelength correction also has some additional benefits that may be explored in future work. It may be possible to estimate some errors, including residual non-linearities, in the displacement measurement by observing the difference between the corrected displacement measurements for each wavelength. Additionally, two-wavelength refractive index compensation techniques may be applied [88].

The second weakness of the Heydemann correction is that when the quadrature

signal data available for correction covers less than a full ellipse, the accuracy of ellipse fitting algorithms is reduced. This limitation has been discussed with respect to the effects of the choice of ellipse fitting algorithm, however as established in figure 6.5, residual non-linearity errors increase with reduced fit range, regardless of the algorithm employed. The final new correction presented in this chapter aims to fill this performance gap in homodyne displacement measuring optical interferometry by constraining the centre point of the ellipse fit, and in some case the quadrature signal amplitudes. This approach has been shown to be capable of extending the performance of homodyne interferometry down to approximately 50 nm displacements, for an interferometer with a fringe period of 316 nm. The multiple intensity reference and constrained Heydemann algorithms have also been shown to be capable of correcting non-linearities to the sub 10 pm level. Further work in this area may explore the impact of constraining the ellipse fit on the displacement measurement uncertainty, and consider the benefits of applying more advanced fitting algorithms.

In conclusion, the algorithms presented in this chapter are not directly comparable, with novel algorithms developed specifically to fill different gaps in the performance of existing corrections. The Heydemann correction continues to be the best available first step when correcting homodyne interferometer measurements covering a displacement greater than one fringe period. In cases where higher order non-linearities cannot be eliminated through improved optical design, the applicability of the measurement arm normalisation technique should first be considered. Where this technique does not provide a complete correction, and non-linearities remain unacceptably high for the application, the two-wavelength approach may be employed. If the available displacement range covers less than one full optical fringe, the multiple intensity reference, and closely related constrained Heydemann, algorithms offer the potential for order of magnitude level reductions in residual non-linearities, depending on noise levels and the exact displacement range available. Where real time correction of interferometric measurements is required some approaches have been outlined, however, as this is not the focus of this thesis, an in depth comparison has not been made.

Chapter 7

Displacement measuring phase-quadrature coating based interferometry

7.1 Introduction

The NPL plane mirror differential interferometer is a well studied [11, 54] interferometer with an optical design shown in figure 7.1. This interferometer design makes use of a Jamin interferometer [13] style PQBS block to produce and recombine parallel reference and measurement beams. A PBS, QWP and cube corner retro-reflector are employed to allow these beams to be interfered in a differential plane mirror configuration with no optical dead path when the reference and measurement mirror surfaces are co-planar, and with balanced beam paths in air and glass. The use of a cube corner retro-reflector ensures that changes in angle of the measurement and reference mirrors do not result in changes in the relative angle between the measurement and reference beams at the sine and cosine outputs of the interferometer, with any angular motion resulting only in translation of the reference and measurement beams in the detector plane. Additionally, the closely spaced measurement and reference beam paths reduce the sensitivity of the interferometer to air turbulence or thermal gradient effects. As a result of the potential for zero dead path operation, and the reduced sensitivity to angular motion and air turbulence, this interferometer design is well suited to precision displacement measurement.

Although the general design of the interferometer is well suited to precision displacement measurement, as a result of non-idealities in the polarising optics used for beam steering in this interferometer design, non-linearities can arise [54]. Additionally, the anti-reflection coated planar optical surfaces perpendicular to the beam propagation direction introduced by the QWP, PBS and solid glass cube corner retroreflector may introduce small back reflections, which also result in higher order non-linearities. Previous measurements of the non-linearity of this interferometer design have revealed peak-to-peak non-linearities of 40 pm [54].

Interactions between these various non-linearity sources can be complex in nature, and as such in this chapter they will be examined firstly with the modelling technique proposed in chapter 5 [85], before moving onto experimental work. As

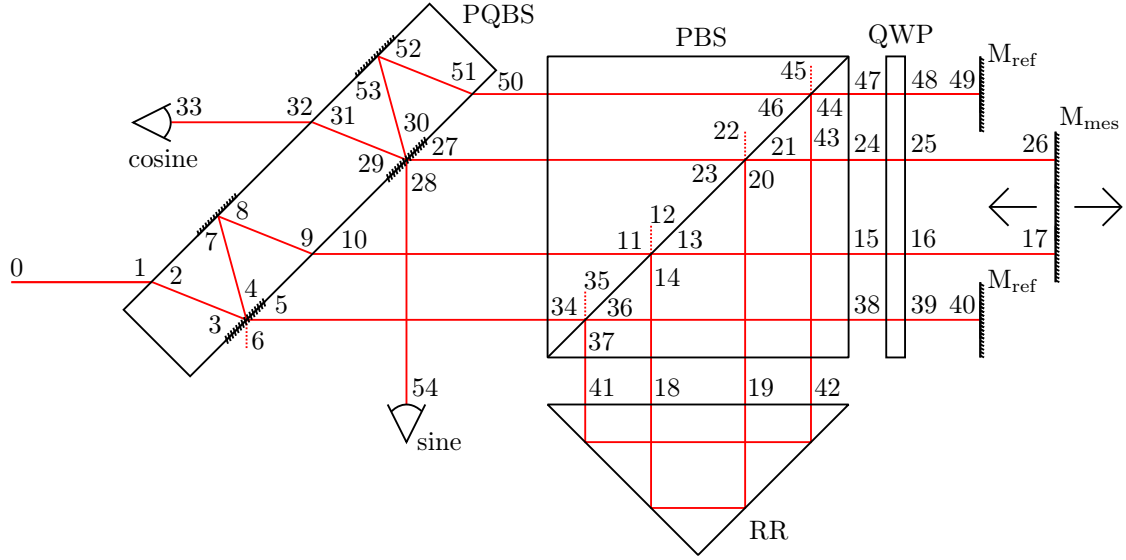


Figure 7.1: Schematic of NPL plane mirror differential interferometer with nodes (as used by the combined scattering and transfer matrix model [85]) and key optical components labelled. PQBS, phase quadrature beam splitter; PBS, polarising beam splitter; QWP, quarter wave plate; M_{ref} , reference mirror; M_{mes} , measurement mirror; RR, retro-reflector. Sine and cosine outputs are indicated.

demonstrated in chapter 6, where an ellipse fitting Heydemann style correction is applied the effects of gain ratio, offset and quadrature phase errors in the quadrature signals are negligible where large numbers of points are included in the fit, and the displacement covers ranges of greater than one optical fringe (approximately 158 nm for this interferometer design), and as a result the analysis of non-linearities will focus on higher order effects. Following this investigation into non-linearity sources, an improved design will be described, based upon the NPL plane mirror interferometer, aiming to reduce the effects of the identified non-linearity sources. The effects of these design improvements will be tested experimentally to investigate the robustness of the non-linearity mitigation methods. Finally, the performance of the final improved design will then be tested against the NPL X-ray interferometer, allowing traceable measurements of the residual non-linearities to be made.

7.2 Modelling the NPL plane mirror differential interferometer

A number of potential non-linearity sources exist in the NPL plane mirror differential interferometer design. In this section, these non-linearity sources will be analysed, both individually, and in combination.

The first non-linearity source to be considered is the polarisation based beam steering assembly, which is vulnerable to imperfections in the polarising optics. In particular, misalignments or imperfect retardance values of the QWP introduce a

complex cavity that would be challenging to model with conventional techniques. Following the path of the measurement beam through figure 7.1, light incident to the interferometer at node 0 is split into measurement and reference beams at nodes 4 and 5. The measurement beam then passes through the QWP (nodes 15 and 16), is reflected from the measurement mirror at node 17 and passes back through the QWP. The PBS and retro-reflector then steer the beam back to the measurement mirror via the QWP a second time at node 26. In the ideal case, the beam is transmitted through the PBS coating from node 44 to 46, however, if there are imperfections in the QWP a portion of the beam will instead be reflected and return to the measurement mirror a third time, repeating the latter part of the measurement beam path before finally transmitting through the PBS to the PQBS and detectors. This first order unwanted reflection travels twice the distance travelled in the interferometer measurement arm by the desired zeroth order signal, and higher order passes will also be present. The modelled effects of this non-linearity path on post Heydemann correction non-linearities are demonstrated in figure 7.2 in both displacement and spatial frequency space for a 1° QWP retardance error, and in figure 7.3 for varying QWP retardance and rotation angle for an otherwise ideal interferometer system. Non-linearities arising from imperfections in the QWP alignment or retardance were found to have a period of one optical fringe (approximately 158 nm) after applying a Heydemann correction.

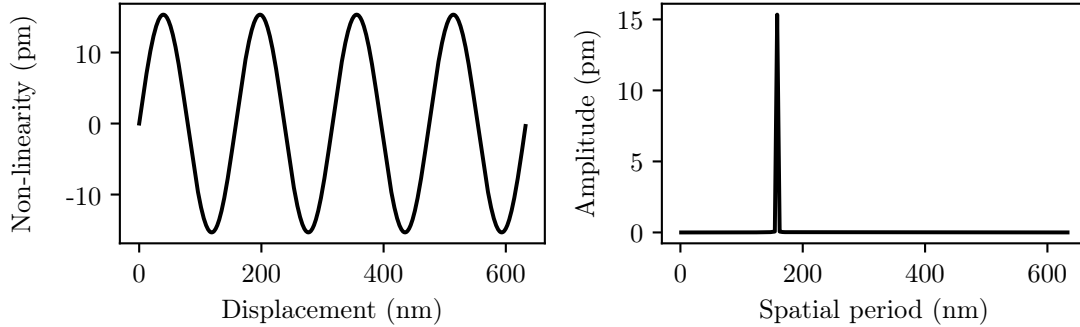


Figure 7.2: Effects of a 1° QWP retardance error on the non-linearity of a NPL plane mirror differential interferometer as modelled with the combined transfer and scattering matrix model for an otherwise ideal interferometer.

In addition to the QWP, the PBS may also introduce non-linearities if the extinction ratio is non-ideal. The beam path resulting in such non-linearities is similar to the previously described QWP non-linearity path, with light that for an ideal PBS would be transmitted from node 21 to 23 instead being reflected from 21 to 20. A significant difference is that the reflected light is in this case P polarised, whereas in the case of QWP induced non-linearities, the light is S polarised. Due to this polarisation mismatch relative to the design beam propagating between nodes 20 and 14 (with the design beam being S polarised) non-linearities introduced by non-ideal PBS extinction ratio may be eliminated in principle by placing a polariser between the PBS and retroreflector, although some residual non-linearities would remain in

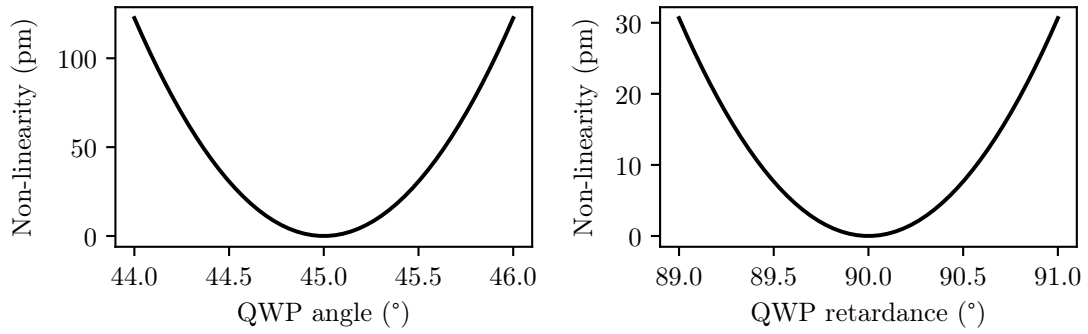


Figure 7.3: Effects of QWP rotation angle and retardance on the peak-to-peak non-linearity of a NPL plane mirror differential interferometer as modelled with the combined transfer and scattering matrix model for an otherwise ideal interferometer.

reality, due to the finite extinction ratio of real world polarisers. In the absence of a polariser, the resulting non-linearities take the same form as those introduced by the QWP with a once per fringe period after applying a Heydemann correction. As these non-linearities are introduced by non-ideal reflection from the PBS (with transmission of light of the incorrect polarisation not introducing non-linearity) the commonly specified transmission extinction ratio for thin film based PBS cannot be used to estimate non-linearities, as the extinction ratio of thin film based PBS in reflection is typically significantly worse than in transmission [186,187]. The effects of varying PBS reflection extinction ratio are shown in figure 7.4, with extinction ratios of approximately 10000:1 required to achieve non-linearities of 1 pm in the absence of a polariser. For comparison, transmission extinction ratios of typical broadband thin film coating based PBS cubes are approximately 1000:1 (ThorLabs part number PBS251 [188]), monochromatic PBS cubes 3000:1 (ThorLabs part number PBS25-633 [189]), and monochromatic plate PBS 10000:1 (ThorLabs part number PBSW-633 [190]).

Unwanted reflections, for example, from the external surfaces of the PBS and QWP, may also introduce non-linearities into the interferometer design. Although all optical surfaces are anti-reflection coated, with reflections typically below 0.5%, even these small reflections can have significant effects on the non-linearities of the interferometer. Numerous unwanted reflections are possible, with a wide range of interactions between the various reflection sources and polarisation errors. As an example, the effects of such an unwanted reflection with a power reflectivity of 0.1% are shown in figure 7.5 in both displacement and frequency space, with harmonics visible at approximately 158 nm ($\lambda/4$), 79 nm ($\lambda/8$) and 53 nm ($\lambda/12$).

To demonstrate the interactions that can arise as a result of combined reflection and polarisation effects, the simulation presented in figure 7.5 is repeated with a 1° QWP retardance error also modelled in figure 7.6. Harmonics can be seen to arise in the non-linearities that are not present in either figure 7.2 or figure 7.5, further demonstrating the capability of the combined scatter and transfer matrix model to identify complex interactions between non-linearity sources. Due to the extremely

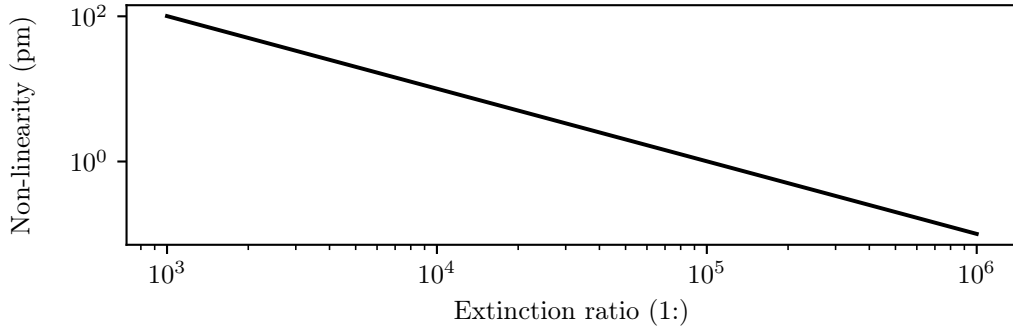


Figure 7.4: Effects of varying PBS reflection extinction ratio on the peak-to-peak non-linearity of a NPL plane mirror differential interferometer as modelled with the combined transfer and scattering matrix model for an otherwise ideal interferometer.

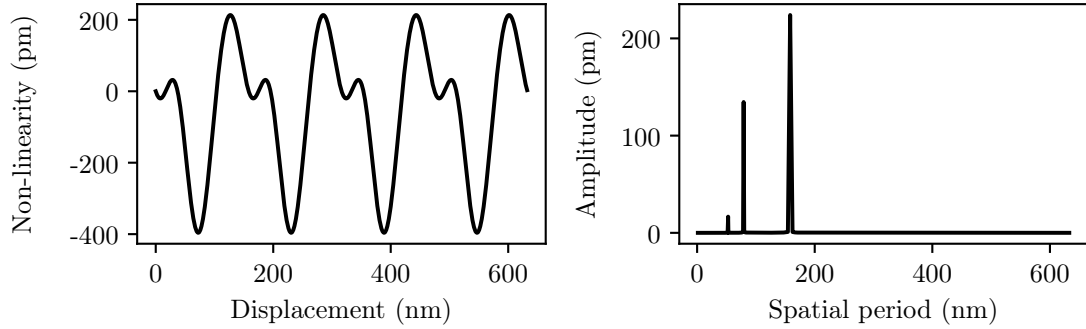


Figure 7.5: Effects of a 0.1 % reflection from the front surface of the PBS (between nodes 13 and 15, 21 and 24, 36 and 48, and 44 and 47) on the peak-to-peak non-linearity of a NPL plane mirror differential interferometer as modelled with the combined transfer and scattering matrix model for an otherwise ideal interferometer.

large numbers of possible back reflection sources, and interactions, not all can be analysed here. The simulations presented here do, however, serve to demonstrate the relatively large non-linearities that can be introduced even by small back reflections (although it should be noted that this model assumes perfect alignment as previously discussed, and therefore places an upper limit on non-linearities), and the conclusion can be drawn that that even anti-reflection coated surfaces should not be aligned to be exactly perpendicular to the beam propagation direction if picometre level non-linearities are to be achieved.

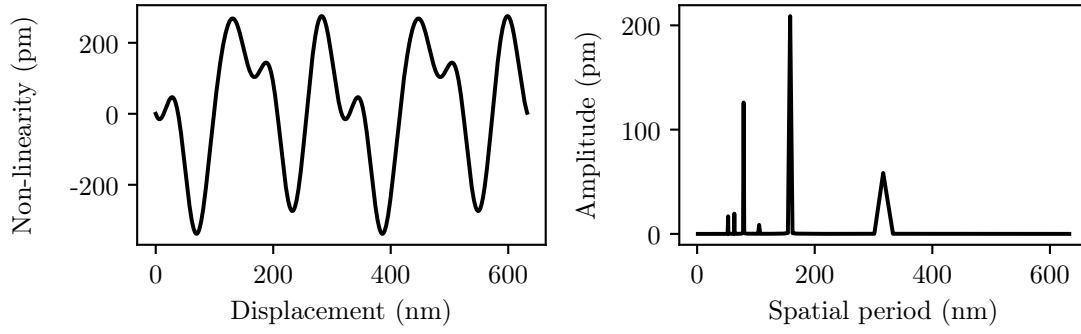


Figure 7.6: Effects of a 0.1 % reflection from the front surface of the PBS (between nodes 13 and 15, 21 and 24, 36 and 48, and 44 and 47) combined with a 1° QWP retardance error on the peak-to-peak non-linearity of a NPL plane mirror differential interferometer as modelled with the combined transfer and scattering matrix model for an otherwise ideal interferometer.

7.3 Survey of existing interferometers

From the previously presented modelling work, it is apparent that a number of potential non-linearity sources exist in the NPL plane mirror differential interferometer design. This design has been in use for a number of years, with interferometers following this design built into a number of metrological instruments [6, 55, 56, 74]. Ten NPL plane mirror differential interferometers were available for measurement, and these were tested both in order to determine a baseline level of non-linearity error, and to verify the non-linearity sources predicted by the modelling work.

In order to test the interferometers, each interferometer in turn was set up with the reference mirror mounted on a Queensgate Instruments X15-A nanopositioning stage. The interferometers were fibre fed with a fibre coupled polarisation stabilised HeNe laser (REO Model 32734) operating at approximately 632.8 nm, using polarisation maintaining fibre. The interferometer outputs were amplified with NPL designed transimpedance amplifiers, with a bandwidth of 100 kHz, and recorded. Each test interferometer was mounted in a cradle allowing the interferometer body to be rotated about all three axes, and each interferometer was aligned first to the measurement mirror by blocking the reference arm beams, and adjusting the tip and tilt of the interferometer body to maximise the quadrature signal outputs. The reference arm of the interferometer was then unblocked, and the amplitudes of the quadrature signals were maximised by adjusting the alignment of the reference mirror, whilst displacing the measurement mirror such that fringes were produced. This follows the alignment procedure generally used in practice, although it does not guarantee that the reference and measurement mirror are parallel if the internal optics of the interferometer are misaligned.

To measure non-linearities, the stage was displaced through the full 15 μm range in 1000 steps of 15 nm. To reduce the effects of both electrical noise and mechanical vibration, 100 values of the sine, cosine and intensity reference outputs were

recorded at each step, and averages were calculated. The interferometric phase was calculated from these outputs following the approach described in section 3.3, and the interferometric displacement was calculated following the approach described in section 3.2, assuming a constant optical wavelength of 632.8 nm, and therefore implicitly assuming fixed optical frequency and refractive index. The wavefront correction factor was also neglected. Assuming a fixed value for the wavelength without accounting for refractive index fluctuations may introduce a small scaling error, estimated to be of the order of parts in 10^6 in the $\pm 1^\circ\text{C}$ temperature stabilised laboratory in which the measurements were carried out, however, for the purposes of non-linearity estimation this is not significant. The difference between the interferometric displacement and the stage set position was then calculated for each stage position, and a Fourier transform of this difference was calculated. Before calculating the Fourier transform, a Hann window [185] was applied to reduce spectral leakage and edge effects, and the data was zero padded, with a number of zeros equal to the number of unpadded data points adjoined to each end of the data. The non-linearity measurement was repeated a minimum of three times for each interferometer tested, and the Fourier amplitude spectra were averaged. Averaging was performed on the amplitude spectra, rather than displacement data, so as to avoid averaging out any non-linearities that may have drifted in phase between measurements. Measurements were in some cases taken over time periods of several days, and therefore in this case the potential for phase drifts in the non-linearity was large. To compensate for the amplitude scaling effects of the Hann window function [185], the amplitudes resulting from the Fourier transform were multiplied by a scaling factor calculated as the number of displacement difference data points divided by the sum over all values in the window function. Along with the averages, standard deviations were also calculated, as were expanded uncertainties at a 95 % level of confidence, with coverage factors calculated based upon the student-t distribution and the number of repeats, as described in [89]. Averaged amplitude spectra and uncertainties are presented for each interferometer tested in figure 7.7.

From the averaged amplitude spectra, peak-to-peak non-linearities were estimated as twice the sum of the Fourier amplitudes at the 2nd, 4th, 6th, 8th and 10th spatial harmonics of 632.8 nm, corresponding to spatial periods of approximately 316.4 nm, 158.2 nm, 105.5 nm, 79.1 nm and 63.3 nm. Amplitude values at each harmonic were taken as the amplitude of the nearest peak to the spatial period of the harmonic, to avoid underestimating non-linearities where peaks may have shifted slightly due to drift effects, with a peak defined as any amplitude value higher than both neighbouring values. Estimated peak-to-peak non-linearities for each interferometer tested are presented in figure 7.8.

It should be noted that by estimating the non-linearity in this manner, an upper limit is placed on the non-linearity, which is likely to be an overestimate when non-linearities are small (tens of picometres). The estimate yields an upper limit for two reasons. Firstly, the estimate assumes all spatial harmonics add constructively, which may not be the case. Secondly, even if no non-linearity was present at a given spatial period, the noise floor of the Fourier amplitude spectrum would still be included in the estimate. The noise floor in this case is the result of the combined

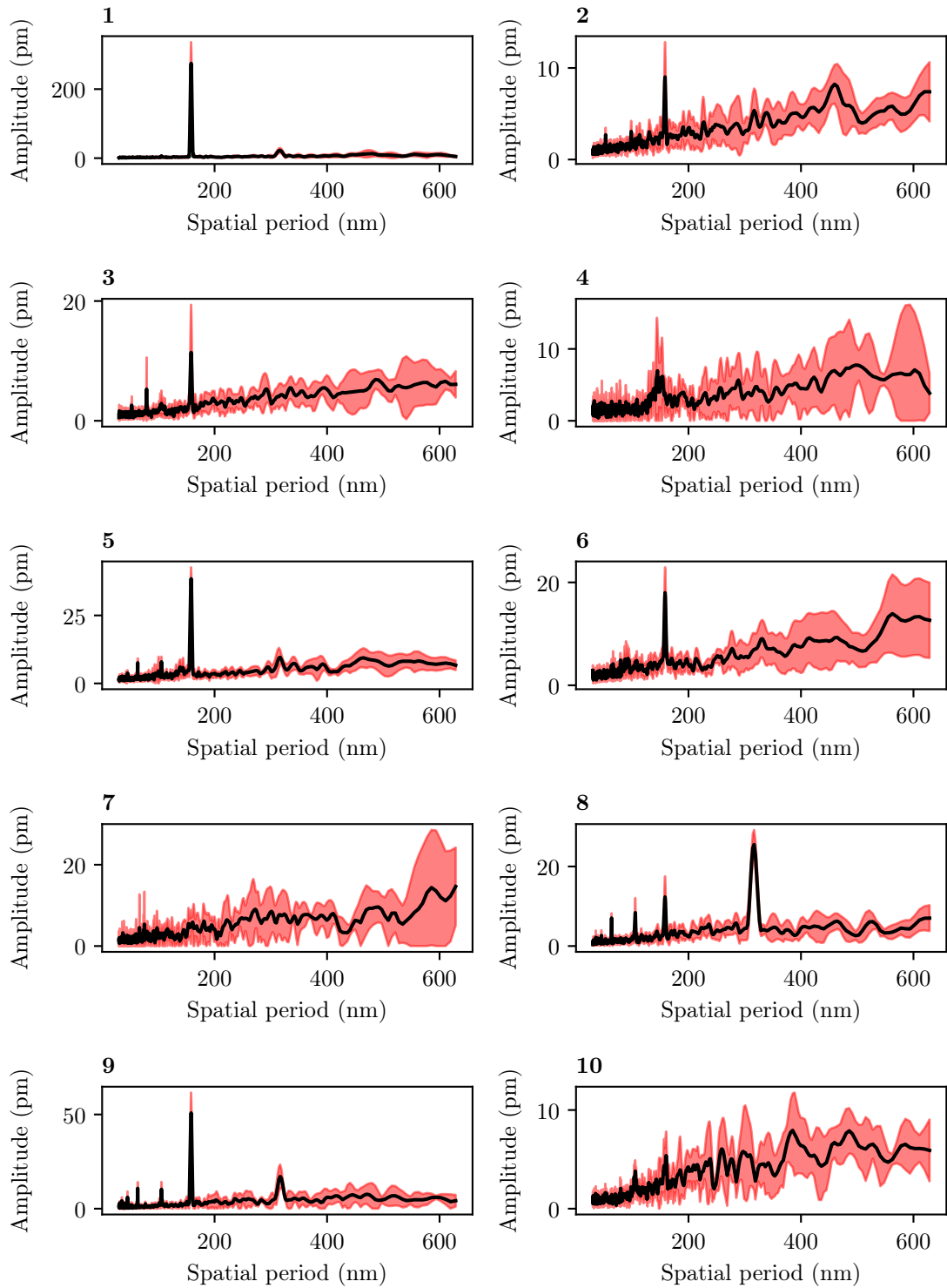


Figure 7.7: Results of a survey of the peak-to-peak non-linearity of ten NPL plane mirror differential interferometers (**1** to **10**). Non-linearity error amplitudes are shown in frequency space, calculated as described in the text. Shaded regions indicate expanded uncertainties at a 95 % level of confidence. Corresponding total non-linearity estimates are presented in figure 7.8.

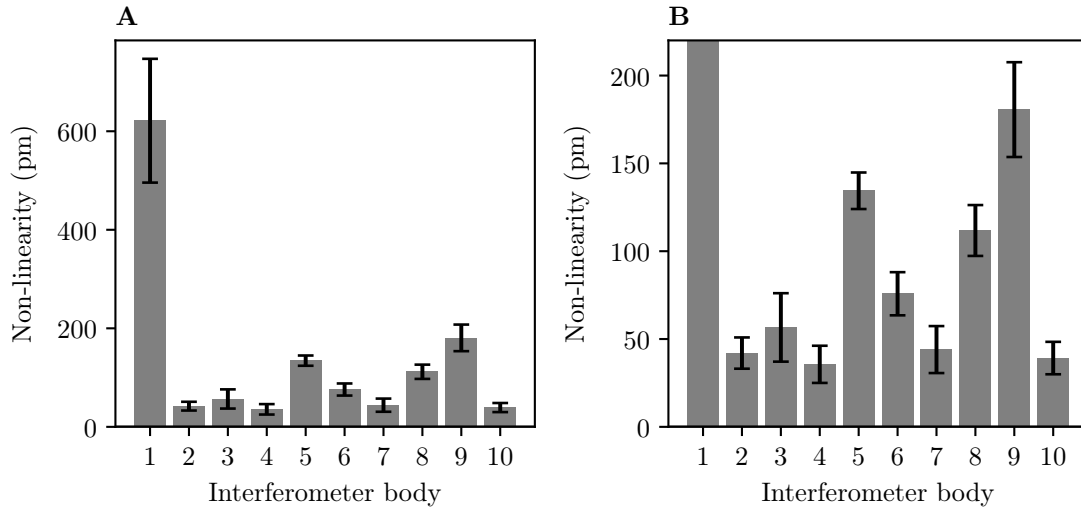


Figure 7.8: Results of a survey of the peak-to-peak non-linearity of ten NPL plane mirror differential interferometers. Non-linearities estimated as described in the text. Error bars shown expanded uncertainties at a 95 % level of confidence. **B** shows the same data as **A**, but with an adjusted ordinate axis scale.

effects of optical and electronic noise in the quadrature signals, air turbulence introducing refractive index fluctuations, and mechanical vibrations, and is therefore frequency dependent, generally increasing at longer spatial periods. As shown in figure 7.7, for the measurements presented here, the noise floor generally remained lower than 10 pm in amplitude within the region of interest (spatial periods of less than one optical wavelength). This overestimate effect should be accounted for when interpreting figure 7.8.

Significant variability was found in non-linearity between the interferometer bodies tested, as can be seen in figures 7.7 and 7.8. One clear outlier is present; interferometer body 1 was found to have significantly larger non-linearities than the other interferometers tested. This was attributed to a failed anti-reflection coating on the QWP, causing a relatively large back reflection, possibly combined with a stress induced QWP retardance error, with mounting induced stresses in the QWP hypothesised to have caused the observed failure of the anti-reflection coating.

Non-linearities in the other interferometers tested are attributed to a variety of causes. Higher order harmonics introduced by multiple reflections, as predicted in figure 7.5, are clearly visible in figure 7.7 for interferometer bodies 2, 3, 5, 8 and 9. Weak higher order harmonics are also visible for interferometer body 10, at periods of $\lambda/6$ and $\lambda/8$. Interferometer bodies 1, 8 and 9 very clearly show the combined effects of multiple reflections and polarisation errors, with a non-linearity component present at $\lambda/2$, as predicted in figure 7.6. Once per fringe errors arising from polarisation errors, predicted in figures 7.2, 7.3 and 7.4, are visible in all but one interferometer tested, with interferometer body 7 showing no non-linearities above the noise floor, although it should be noted that small non-linearities may be

masked by the noise floor.

Considered as a whole, the observed non-linearities are in line with previous measurements of the non-linearity of this interferometer design [54]. The lack of visible non-linearity for interferometer body 7, along with the very small non-linearities of interferometer bodies 4 and 10, demonstrate the potential of this interferometer design where unwanted back reflections are absent, or misaligned, and the polarising optics are close to ideal. The significant variability, however, demonstrates that further development is needed if the performance of state of the art interferometer designs is to be matched, or exceeded, consistently. Whilst the observed non-linearity, when present, is likely to be a combination of QWP errors, PBS leakage, and unwanted back reflections, with the total non-linearity being reduced compared to the model predictions as a result of small misalignments truncating higher order effects, the presence of a once-per-fringe non-linearity error in all but one of the tested interferometers does suggest that the polarisation based beam steering optics are a key weakness of this interferometer design. If residual non-linearities are to be reduced, this should therefore be a focus area, with a combination of improved optics and a novel non-linearity correction approach required.

7.4 Improved optical design

A number of non-linearity sources in the existing NPL plane mirror differential interferometer design have been identified through modelling with the combined scattering and transfer matrix model. These findings have been confirmed with experimental measurements of existing interferometers. An improved design is presented in figure 7.9, aiming to eliminate these non-linearity sources.

The improved design introduces a small tilt angle to all optical components (as shown in the schematic representation of figure 7.9), with components tilted by differing angles to minimise the number of parallel optical surfaces. Tilting the PBS has the potential to degrade the performance of the beam splitter. As thin film based PBS rely upon a series of near Brewster angle reflections to separate S and P polarisations [186], the change in extinction ratio may be expected to be asymmetric around the 45° design angle of incidence when operating at a given wavelength. As such, tilt angles either side of 45° were tested experimentally to determine the optimal tilt direction, with the extinction ratio monitored visually by observing the intensity of reflected light whilst illuminating the cube with an P polarised beam. A decreased angle of incidence, as shown in figure 7.9, was found to maximise the extinction ratio for the PBS used in this work. QWP performance may also be degraded at non-normal incidence, however, this was not tested, with any resultant errors intended to be corrected by the measurement arm normalisation technique.

In addition to tilting optical components, a non-polarising beam splitter has been added in the return path of both the measurement and reference arms allowing the non-linearity correction methodologies developed in sections 6.4 and 6.5 to be applied, with the measurement arm normalisation methodology of section 6.5 intended to minimise non-linearities introduced by the imperfect performance of the QWP and PBS. As the measurement arm normalisation methodology was found

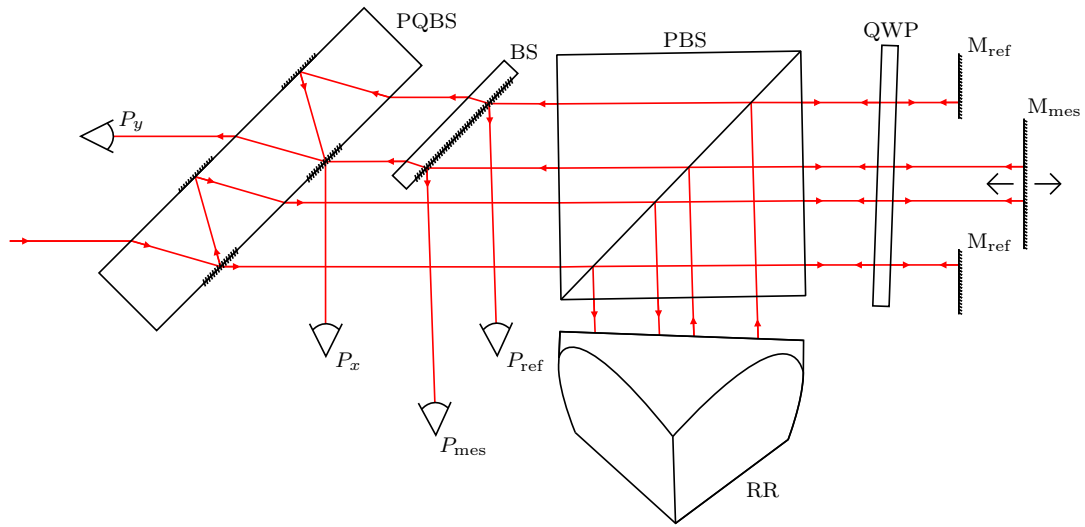


Figure 7.9: Schematic of the improved NPL plane mirror differential interferometer with key optical components labelled. PQBS, phase quadrature beam splitter; PBS, polarising beam splitter; QWP, quarter wave plate; M_{ref} , reference mirror; M_{mes} , measurement mirror; RR, solid glass retro-reflector; BS, 50:50 non-polarising beam splitter. Arrows indicate direction of propagation. Tilt angles of the optics are shown to scale.

to compensate for the effects of non-ideal PBS extinction ratio, and due to space constraints, no polariser was introduced between the PBS and retro-reflector.

A prototype interferometer following this design was constructed. The interferometer was assembled using spare optics for the existing NPL plane mirror differential interferometer stock, with the exception of the 50:50 non-polarising beam splitter (Edmund Optics #48-902). In order to ensure repeatable misalignment of back reflections, the interferometer was assembled on a custom machined aluminium baseplate, as shown in figure 7.10, with locating features for the optics. The base plate was mounted on a 6 degrees of freedom stage, allowing the interferometer base plate to be aligned relative to the reference and measurement mirrors.

The PQBS, PBS, QWP and Beam Splitter (BS) were located against steel dowels, setting the angle and position of the optics. Optics were fixed to the steel dowels with Norland 107 UV curing optical adhesive. In the case of the PBS, only two locating dowels were used as shown in figure 7.10, fixing the angle of the PBS, but allowing the PBS to be translated across the face of the corner cube retroreflector. This adjustment was used to account for manufacturing tolerances, permitting the spacing between the upper and lower beam pairs in figure 7.9 to be set to the desired 25.4 mm outer beam spacing. The outer and inner beams in this interferometer design are spaced by 10 mm, with the spacing set by the PQBS thickness, resulting in an inner beam spacing of 5.4 mm. The corner cube retroreflector was located in a V groove, with shims used to adjust the height of the corner cube, ensuring all four output beams of the interferometer remained in the same plane. The interferometer was illuminated using a fibre coupled HeNe laser operating at approximately 632.8 nm,

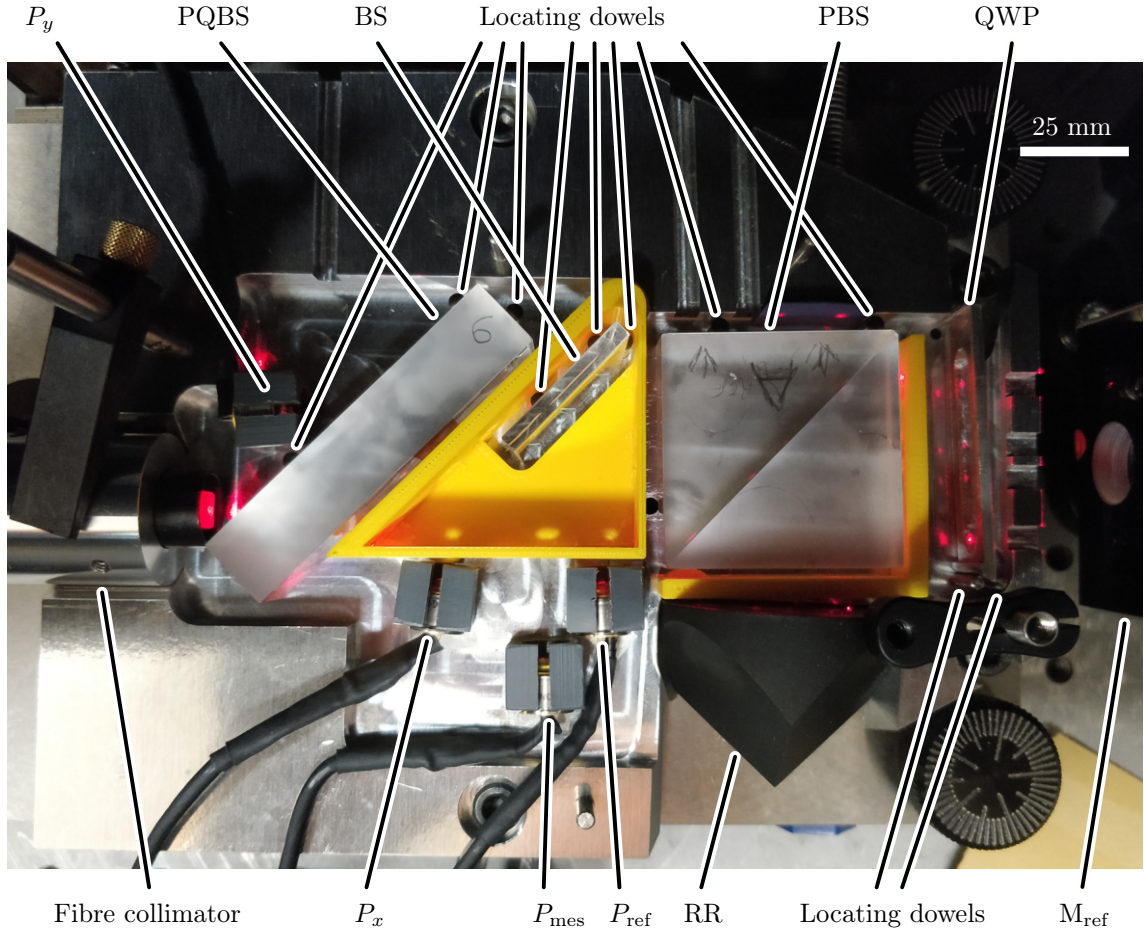


Figure 7.10: Photograph of improved NPL plane mirror differential interferometer prototype with key components labelled, as in figure 7.9. 3D printed beam baffles as discussed in the text (yellow in colour) are also visible. The Glan-Thompson polariser mounting coaxially with the fibre optic collimator in this photograph was later removed, and was not used when gathering the data presented in this work.

with the fibre collimator again located in a V groove.

During assembly of the interferometer, the reference mirror, which had a 10 mm hole trepanned in the centre allowing the measurement beams to pass through (shown to the right of figure 7.10), was mounted on a kinematic mirror mount. The reference mirror was positioned approximately 60 mm in front of the measurement mirror, which was mounted on a Queensgate X15-A nanopositioning stage, resulting in an approximately 60 mm dead path. Whilst this dead path increased the sensitivity of the instrument to changes in refractive index, non-linearities could still be resolved in the thermally stabilised ($\pm 1^\circ\text{C}$) laboratory environment. Before beginning the assembly process, the reference and measurement mirrors were first aligned to be parallel using a visual angle dekkor, with the mirrors estimated to be parallel to within one minute of arc.

Initially, all optical components were installed in their nominal positions. To align the various optical components, the reference mirror was initially blocked.

The interferometer base plate was first aligned to the measurement mirror by removing the QWP, and ensuring the back reflected beam was centred on the fibre collimator. The QWP was then replaced, and the PBS roughly aligned to give the desired 5.4 mm spacing between the inner beams. The retroreflector was then shimmed to ensure the return beam passed over the base plate at the same height as the outgoing beam. Having set the return beam height, the reference mirror was unblocked, and a final adjustment of the PBS was made by maximising the amplitudes of the outputs of the sine and cosine detectors whilst displacing the stage through multiple fringes, before cementing the PBS in place. This process was designed to ensure the interferometer would produce the maximised sine and cosine outputs when the reference and measurement mirrors were parallel, when aligned in the future. The rotation of the QWP was then aligned by adjusting set screws located under each end of the QWP mount, aiming to minimise the modulation of the measurement arm reference output whilst displacing the stage. As the QWP had to be removed to adjust these screws, the alignment was iterative, and only approximate, with residual alignment errors compensated for with the measurement arm normalisation correction. A zero order mica QWP was employed, designed for operation at 632.8 nm.

In order to reduce the potential for the many misaligned back reflections introduced by the tilted optical surfaces to reach the detectors, 3D printed beam baffles (Polylactic Acid (PLA) plastic) were included as shown in figure 7.10, with 4 mm apertures around each beam path to avoid clipping the 1.2 mm $1/e^2$ diameter beams. A 3D printed cover (PLA plastic, not pictured) was also placed over the interferometer as a whole during measurements in order to shield the beam paths from air turbulence.

The output beams of the interferometer were detected with Centronic OSD 5-5TB photodiodes, with integrated ‘bug eye’ condensing lenses. The photodiodes were operated in photovoltaic mode (with no applied bias current) to minimise dark currents, and were amplified with NPL developed electronics. The resulting voltage signals were then digitised at 200 kHz with the 16-bit ADCs of a National Instruments 7833R FPGA card, and recorded for further processing, with no data processing performed on the FPGA. The interferometer outputs were processed as described in sections 3.3 and 3.2, with a Heydemann correction applied, and either the reference arm intensity reference or measurement arm intensity reference used, separately, as described in section 6.5, with the reference arm intensity reference used only to provide a baseline non-linearity for comparison.

7.5 Non-linearity evaluation

In order to evaluate the non-linearities of the prototype improved plane mirror differential interferometer, two sets of measurements were carried out. Preliminary measurements were made against a Queensgate X15-A nanopositioning stage. More accurate, and lower noise, measurements were then made using the NPL XRI facility.

7.5.1 Methods

Nanopositioning stage methods

Non-linearities were first measured against the Queensgate Instruments X15-A nanopositioning stage used during assembly of the interferometer, with the interferometer alignment unchanged after assembly. To evaluate non-linearities, the stage was stepped through the full 15 μm range in 2000 7.5 nm steps, with 1000 data points being recorded at the full 200 kHz bandwidth at each stage position. As all four channels were detected synchronously, normalisation to the reference channels as described in section 6.5 was carried out prior to averaging at each stage position, so that any laser intensity fluctuations within the detection bandwidth were eliminated by the normalisation process. Displacements were calculated assuming a fixed wavelength of 632.8 nm as environmental parameters were not recorded. As discussed for previous measurements, whilst this may introduce a small scaling error (estimated to be parts in 10^6 in the $\pm 1^\circ\text{C}$ temperature stabilised laboratory these measurement were carried out in), a scaling error of this magnitude is insignificant when considering non-linearities of tens of picometres, corresponding to an error of tens of attometres in the non-linearity measurement. The displacement scan was repeated 30 times for each dataset, and the average calculated for both the reference and measurement arm normalised displacements at each stage position, in order to reduce the influence of mechanical vibrations and air turbulence.

Table 7.1: Key properties of Queensgate Instruments X15-A nanopositioning stage [191]. Parameters are quoted across the full closed loop range of the stage, in the ‘slow’ control loop pre-set employed in this work.

Parameter	Value
Closed loop range	15 μm
Scale factor error	0.1 % (to one standard deviation)
Position noise	0.05 nm (RMS)
Linearity error (peak)	0.005 %

It is important to note that by averaging repeated displacement measurements, it is possible for non-linearities to be averaged out if the length of the reference arm of the interferometer drifts between repeat measurements. This effect was minimised in this work by allowing the interferometer to thermalise before taking measurements, and by working inside an enclosure, in a thermally stabilised ($\pm 1^\circ\text{C}$) laboratory.

X-ray interferometer methods

Non-linearities were also measured against the NPL XRI facility, allowing for highly linear measurement in terms of the well-defined lattice parameter of silicon [8]. All XRI measurements were carried out by Dr. Andrew Yacoot, using measurement software written by Dr. Andrew Yacoot, with the unprocessed data being provided to the author for analysis.

The operation of the NPL XRI, along with the application to the measurement of non-linearities in optical interferometers, is described in detail in [11, 54]. In brief, the XRI consists of a flexure stage machined from a single crystal of silicon, with displacement of the stage generated by a Piezoelectric Transducer (PZT). Three silicon lamellae are machined into the XRI, each operating as a x-ray diffraction grating, with the final lamella machined on the moving portion of the flexure stage. The operating principle of the XRI is shown in figure 7.11. The first lamella diffracts the incident x-ray beam from the (220) lattice planes of the silicon, forming two beams. These beams are then incident on the second lamella, where they are again diffracted, each forming two additional beams. The inner beams are recombined on the third, moving, lamella, where they are again diffracted, forming a Moiré pattern between the third lamella, and the interference pattern produced by the beams incident on the third lamella. As a result, the output beams vary sinusoidally in intensity as the third lamella is displaced, with a spatial period equal to the (220) lattice plane spacing, approximately 192 pm [8]. One output beam of the XRI is detected, and the resulting intensity measurement is used to lock the XRI flexure stage to the midpoint of an x-ray fringe, on the linear portion of the sinusoidal output [12]. The XRI can be stepped between adjacent fringes, allowing highly linear motion to be produced in steps of 192 pm.

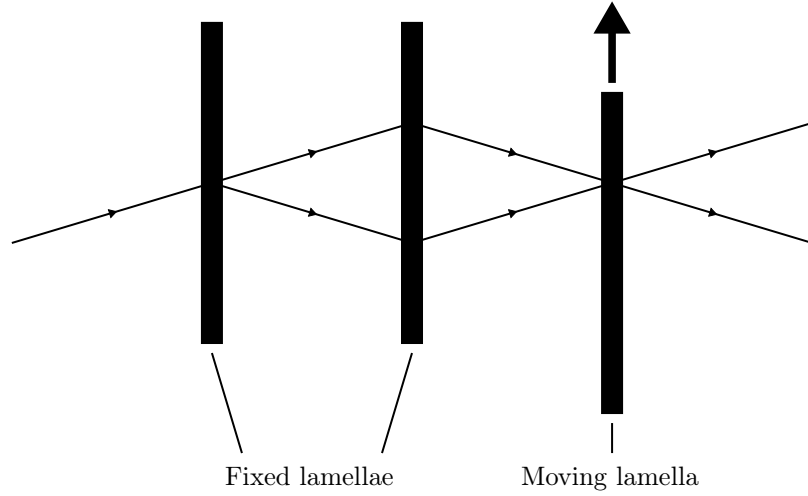


Figure 7.11: Principle of operation of the NPL XRI facility. The x-ray beam is incident from the left, and diffracted at the first and second (fixed), and third (moving) lamellae. The resulting output beams (right) are detected, as described in the text. Not to scale.

To allow for comparison with optical interferometers, three gold coated mirrors are present on the XRI, as shown in figure 7.12. The outer pair of mirrors are fixed, whereas the inner mirror forms part of the flexure stage, and moves with the third lamella. The interferometer was aligned against the XRI mirrors by first blocking the reference arm of the interferometer, and then maximising the outputs of the sine and cosine channels.

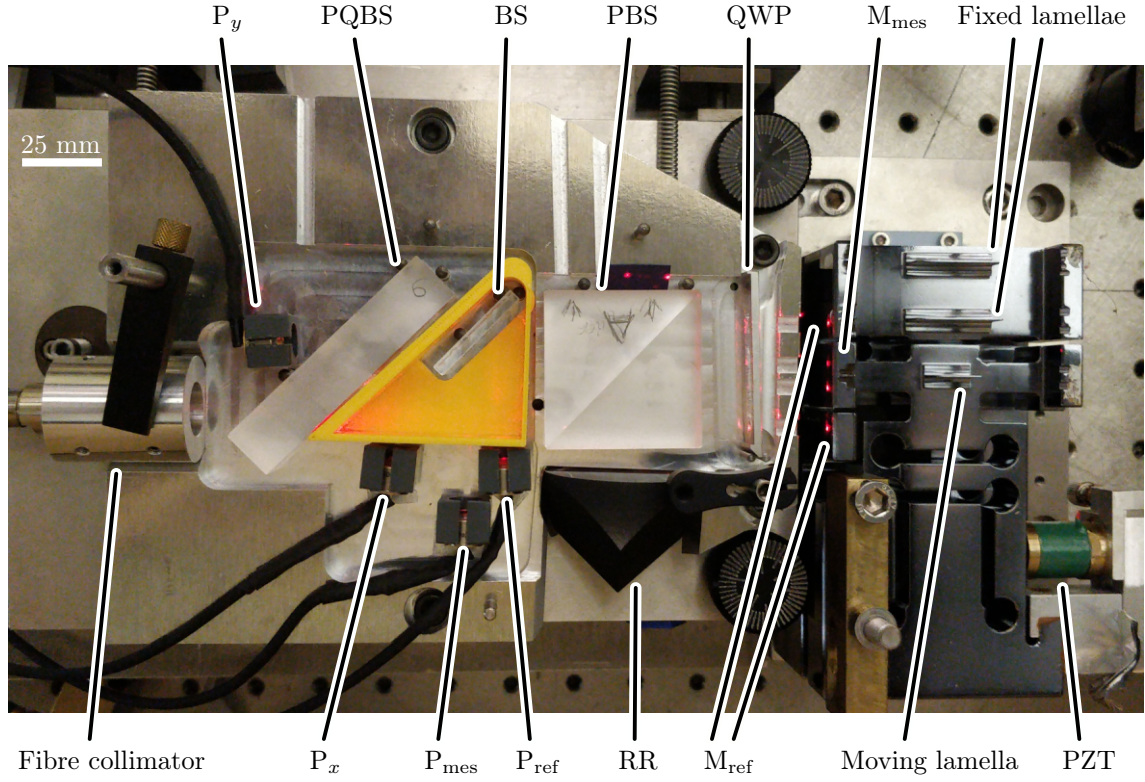


Figure 7.12: Photograph of the prototype improved plane mirror differential interferometer is operation against the NPL XRI facility.

In order to measure non-linearities, the XRI was stepped through 1651 x-ray fringes, corresponding to a displacement of 317 nm, or just over two optical fringes for the double pass plane mirror interferometer design. At each step, a number of data points were recorded, with the x-ray fringe number being recorded, along with the four interferometer outputs. The number of data points recorded at each step was determined by approximate software timing, with nominally 100 ms of data recorded, resulting in an average of 616 data points being recorded at each XRI step.

The interferometer output signals were processed as previously described, with processing conducted separately for the measurement arm intensity reference normalised and reference arm intensity reference normalised cases. Each recorded data point was propagated through to the calculation of the change in interferometric phase without averaging. Having calculated the phase change, relative to the first data point, for each data point, the phase changes were then averaged for each XRI step, and standard errors in the mean were calculated. XRI displacements were calculated from the XRI fringe number by multiplying the fringe number by 192.015 571 4 pm, the (220) lattice spacing of silicon [8]. The optical wavelength was calculated based on the XRI displacement and the measurement arm normalised

phase change averages as

$$\lambda = 8\pi \frac{\max(x_{\text{XRI}}) - \min(x_{\text{XRI}})}{\max(\Delta\varphi_{\text{mes}}) - \min(\Delta\varphi_{\text{mes}})}, \quad (7.1)$$

where x_{XRI} is the XRI displacement, and the dead path was assumed to be zero, again implicitly assuming a fixed refractive index throughout the measurement.

7.5.2 Results

Nanopositioning stage results

Experimental measurements of the non-linearity are shown in figures 7.13 and 7.14. Non-linearities errors are shown, relative to the scaled stage displacement (x), in both displacement (**A** and **B**) and spatial frequency (**C**) space in both cases. In order to clearly visualise the non-linearity in displacement space, the displacement space data was processed in two ways. Figures 7.14 **A** and 7.13 **A** show displacement errors relative to the stage position, with a fifth order polynomial subtracted, with the polynomial subtraction intended to minimise the effects of stage and wavelength errors on the plotted data, allowing non-linearities to be visualised. Figures 7.14 **B** and 7.13 **B** show the non-linearity error averaged over a range of two fringes. In order to show this, the 15 μm displacement range was broken up into 632.8 nm sections, by taking the modulo of the stage displacement with a divisor of 632.8 nm. Due to the non-integer number of data points per fringe, each successive section was slightly offset from the last. A linear fit was then made to the non-linearity errors for each 632.8 nm section, and fitted lines subtracted to minimise the effects of longer spatial period drifts introduced by air turbulence and mechanical instability. Individual data points in figures 7.14 **B** and 7.13 **B** show all of the 632.8 nm sections superposed, with a 50 point moving average applied to visualise the average non-linearity error across the displacement range. Figures 7.14 **C** and 7.13 **C** show Fourier transforms of the difference in displacement between the interferometer and stage displacements. A Hann window [185] was again applied to limit spectral leakage and edge effects, and the data was zero padded with a number of zeros equal to the number of unpadded data points adjoined to each end of the data in order to allow individual peaks to be visually distinguished more clearly, before taking the Fourier transform. The plotted amplitude data has been corrected for the scaling effects of the Hann window function [185].

Preliminary results, shown in figure 7.13, showed unexpectedly large residual non-linearities in the measurement arm normalised case, with a peak-to-peak value of approximately 44 pm. The photodiode detectors were then deliberately tilted in an attempt to reduce the effects of any back reflections from the photodiodes, resulting in the reduced non-linearities shown in figure 7.14. As non-linearities are, in the measurement arm normalised case, approaching the noise floor, it is not possible to state a value of the peak-to-peak non-linearity with a high degree of certainty. The statement of a peak-to-peak non-linearity will therefore instead be based upon the x-ray interferometry measurements, presented in the following section.

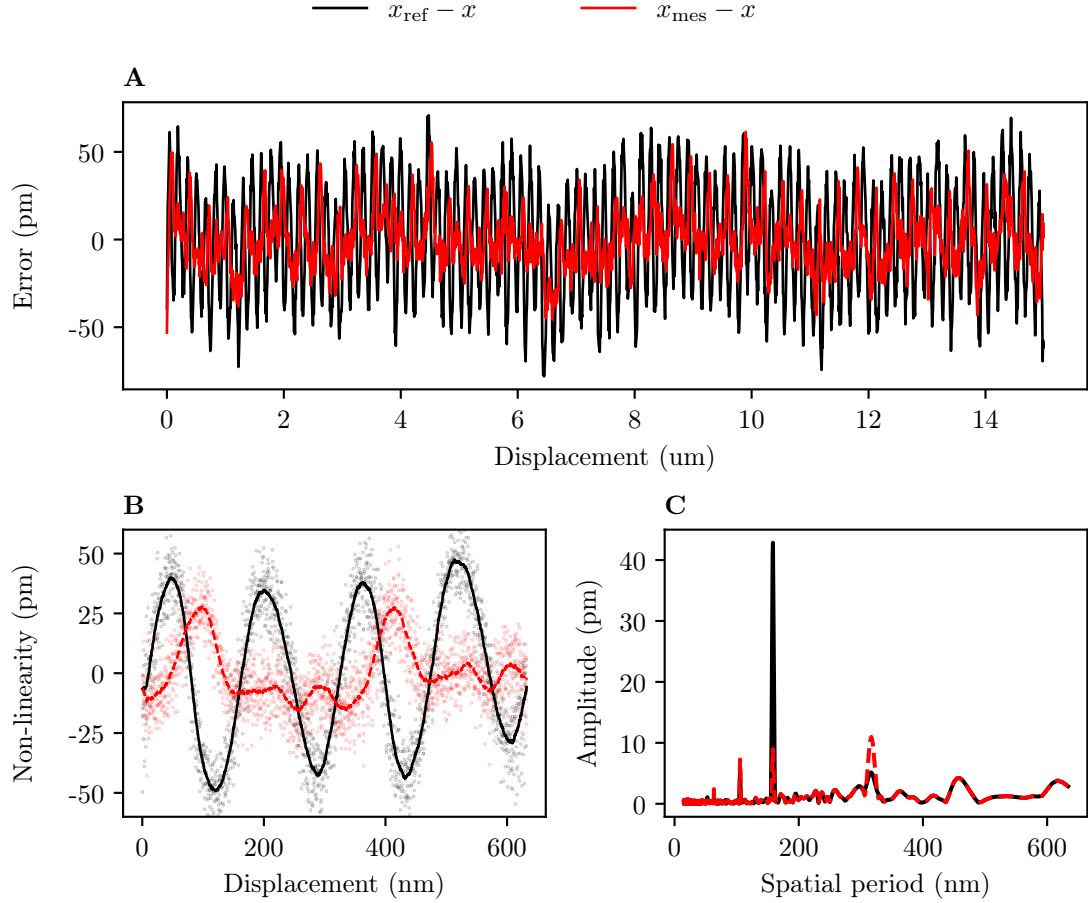


Figure 7.13: Experimental measurement of reference and measurement arm normalised non-linearities, in the presence of back reflections, calculated relative to the stage displacement x , for a $15\text{ }\mu\text{m}$ displacement made in 7.5 nm steps. **A** shows errors relative to the stage position, with a fifth order polynomial subtracted to reduce the effects of stage and wavelength errors, allowing non-linearity errors to be seen clearly. Non-linearities are also shown in displacement (**B**) and spatial frequency (**C**) space, processed as described in the text.

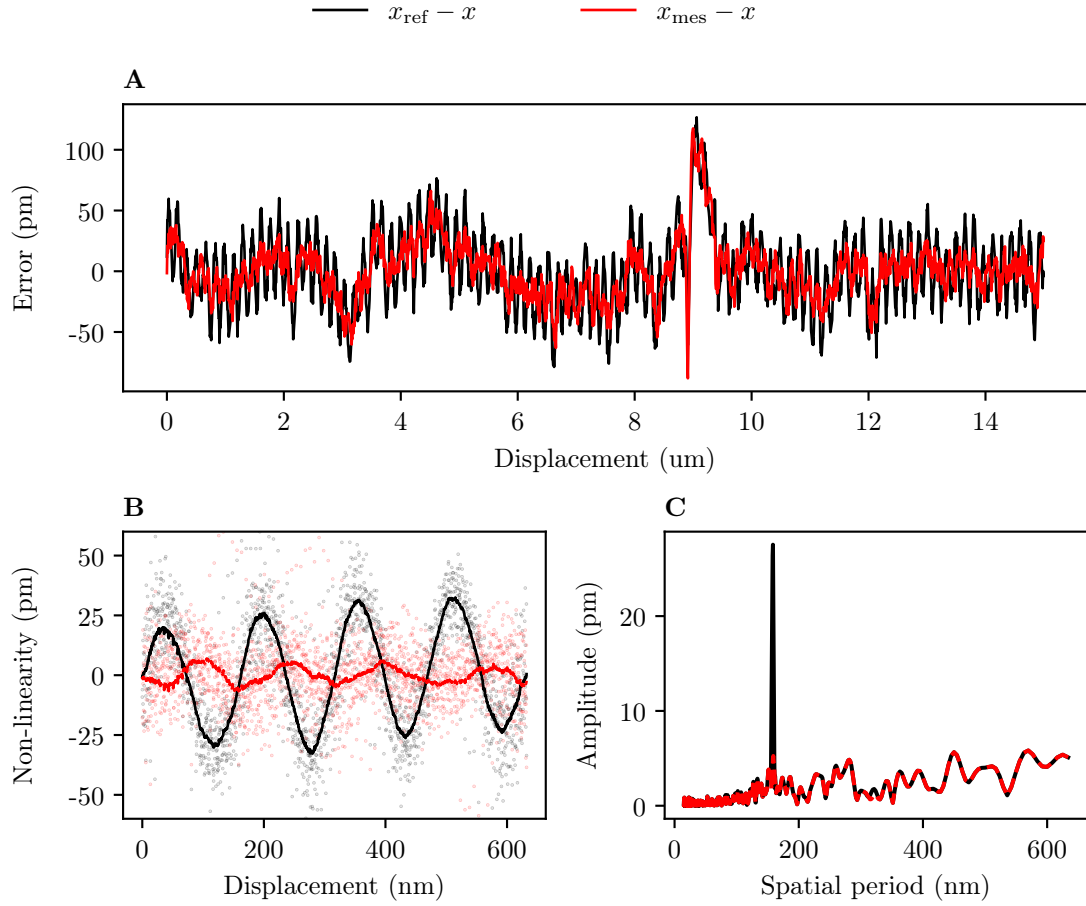


Figure 7.14: Experimental measurement of reference and measurement arm normalised non-linearities, calculated relative to the stage displacement x , for a $15\text{ }\mu\text{m}$ displacement made in 7.5 nm steps. **A** shows errors relative to the stage position, with a fifth order polynomial subtracted to reduce the effects of stage and wavelength errors, allowing non-linearity errors to be seen clearly. Non-linearities are also shown in displacement (**B**) and spatial frequency (**C**) space, processed as described in the text.

In order to investigate the source of the residual non-linearities in both cases, Fourier transforms of the four voltage signals produced by the interferometer, V_x , V_y , V_{ref} V_{mes} , were calculated. Each voltage signal was normalised to its maximum value, a Hann window was applied [185], and the data was once again zero padded with a number of zeros equal to the number of unpadded data points adjoined to each end of the data. As previously, the amplitude data was corrected for the scaling effects of the Hann window function [185]. The results of this process are plotted in figures 7.15 and 7.16, corresponding to the minimised back reflection, and back reflection, cases respectively.

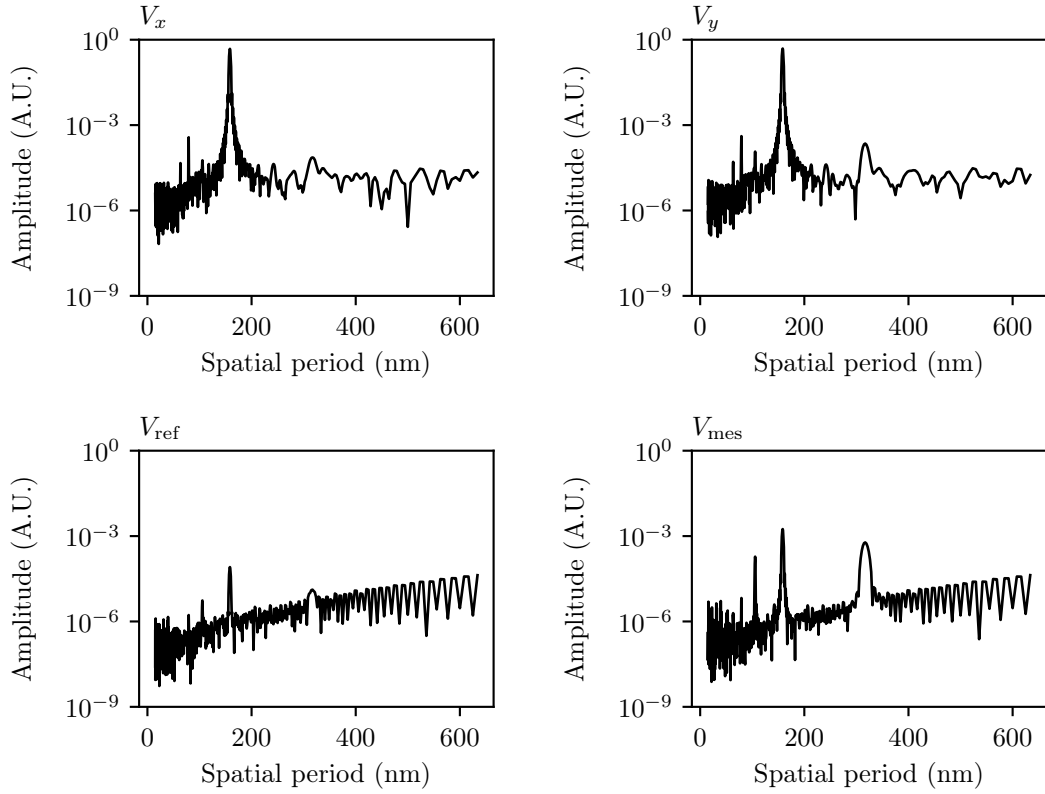


Figure 7.15: Fourier transforms of the four output voltage signals of the prototype interferometer, with back reflections present, calculated as described in the text. Corresponding non-linearities are presented in figure 7.13.

In both cases, peaks can be seen in the V_{ref} signal at harmonics of 632.8 nm. This effect has two possible causes. The first is cross talk between the channels within the electronics. This cause may be excluded, as the same electronics were used to gather the data presented in section 6.4 and [171], where no such effect was seen. The effect must therefore be a result of light from the measurement arm path of the interferometer being coupled into the reference arm. In this interferometer design, this may only take place at two locations: the input fibre collimator, and the output detectors. Considering figures 6.29 and 6.30, it can be seen that the same magnitude of back reflection has a significantly smaller effect on the non-linearity

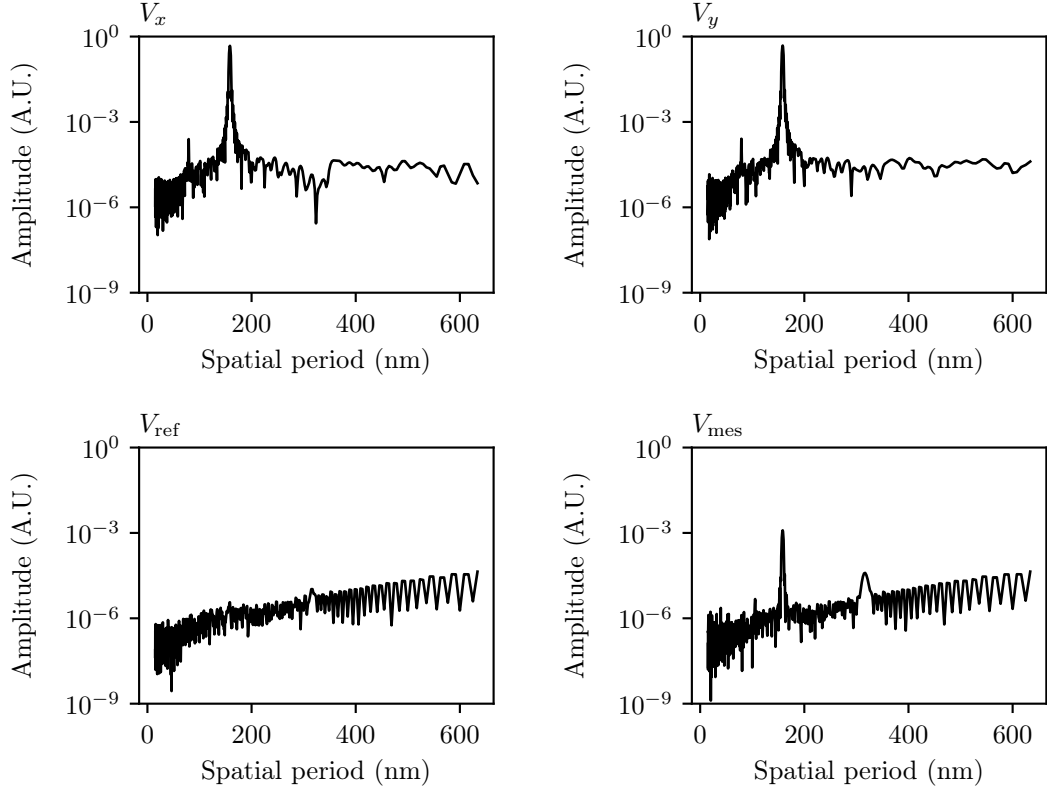


Figure 7.16: Fourier transforms of the four output voltage signals of the prototype interferometer, with back reflections minimised, calculated as described in the text. Corresponding non-linearities are presented in figure 7.14.

when it occurs at the input, therefore reflections from the photodiode detectors are suggested as the cause of the observed effect in both cases.

X-ray interferometer results

Following the nanopositioning stage measurements, the interferometer was set up against the NPL XRI, and measurements were carried out as described in section 7.5.1. The resulting non-linearity measurement is shown in figure 7.17. 7.17 **A** shows the displacement profile of the XRI in integer fringe steps. 7.17 **B** shows the error in the interferometer displacement, calculated relative the XRI displacement. Error bars show expanded uncertainties with $k = 2$, corresponding to a confidence interval of approximately 95 %.

In order to give a quantitative estimate of the peak-to-peak non-linearity, a sum of sinusoids was fitted to the errors (relative to the XRI displacement) plotted in figure 7.17 **B**, in both the reference and measurement arm intensity reference

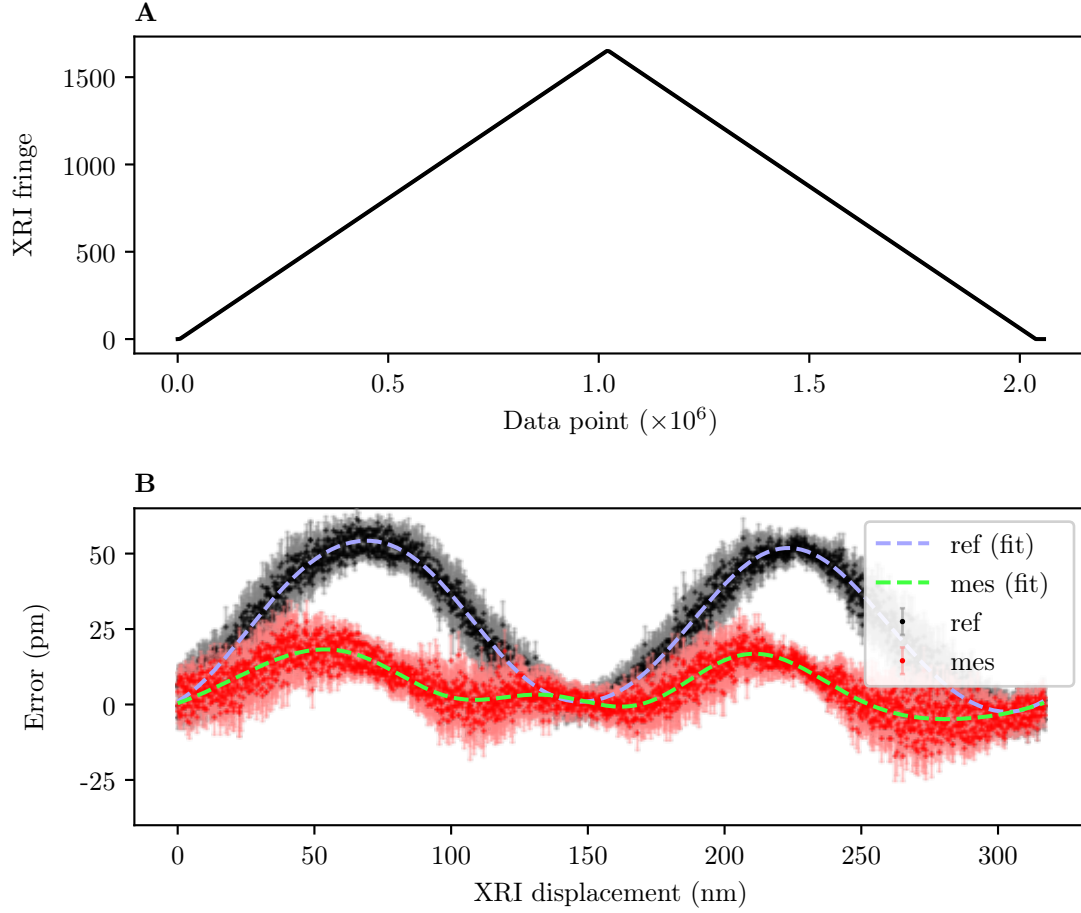


Figure 7.17: Non-linearity errors in the prototype improved plane mirror differential interferometer, measured against the NPL XRI facility. **A** shows the XRI displacement in integer x-ray fringes, with one fringe equal to approximately 192 pm. **B** shows the error in the interferometric displacement in both the reference arm intensity reference normalised (‘ref’) and measurement arm intensity reference normalised (‘mes’) cases. Fitted lines are also plotted, calculated as described in the text.

normalised cases. The function

$$f(x_{\text{XRI}}) = e_0 + \sum_{i=1}^6 A_i \sin\left(2\pi \frac{2i}{\lambda} x_{\text{XRI}} + \phi_i\right) \quad (7.2)$$

was fitted in each case, where the coefficients A_i set the amplitude of the various fitted harmonics, ϕ_i the phase, and e_0 applies an offset, required as the non-linearity errors are not centred on zero. Only even spatial harmonics are fitted, as the interferometer operates in reflection, and therefore non-linearities are odd spatial harmonics cannot be present. Fitting was performed using the trust region reflective algorithm [192] as implemented within the SciPy software [193]. In order to estimate uncertainties in the fitted curves, the fit process was repeated 1000 times following a

bootstrapping method. For each repeat, a randomly generated offset was added to each error value, with the offsets following a Gaussian distribution, with the standard deviation determined by the standard uncertainty separately for each data point. Averages (plotted as the fitted lines in figure 7.17 **B**) and standard deviations in the fitted curves (rather than the fit coefficients) were calculated across the 1000 bootstrap repeats. Uncertainty estimates in the fitted lines are not plotted, as they would not be visible. In addition, averages and standard deviations of the peak-to-peak values of the fitted lines were calculated across the 1000 bootstrap repeats, providing estimates of the peak-to-peak non-linearity. In the reference arm normalised case peak-to-peak non-linearities were estimated to be (56.8 ± 0.4) pm. In the measurement arm intensity reference normalised case peak-to-peak non-linearities were estimated to be (23.0 ± 0.6) pm. In both cases expanded uncertainties are reported with $k = 2$, corresponding to a confidence interval of approximately 95 %.

7.6 Discussion

Modelling work has identified a number of potential sources of non-linearity in the existing NPL plane mirror differential interferometer design. The existence of non-linearity in this interferometer design has been confirmed by previous work in the literature [54], a finding that has been reinforced by the work presented in this chapter. Experimental measurements of the non-linearity of ten existing NPL plane mirror differential interferometers, presented in figures 7.7 and 7.8, have found non-linearities arising from a number of sources, as predicted in section 7.2. Non-linearity errors have been found to be introduced primarily by unwanted back reflections at air-glass interfaces, and by non-ideal polarising optical components. Complex interactions between these effects may occur, making the magnitude of any non-linearities challenging to predict, and leading to the wide range of variation in non-linearity shown in figures 7.7 and 7.8. As a result, no attempt has been made to state an average, or maximum, non-linearity for the existing design.

These findings have driven the development of an improved interferometer, designed specifically to make use of the measurement arm normalisation non-linearity correction methodology introduced in section 6.5. The improved design attempts to minimise the effects of unwanted back reflections by tilting all optical surfaces by a differing amount, ensuring any multiple reflections remain misaligned relative to the primary beam path. The design attempts to eliminate the effects of non-idealities in the polarising optics by employing the measurement arm normalisation technique, which, as discussed in section 6.5, in the absence of unwanted back reflections theoretically offers a complete correction for higher non-linearities introduced by non-idealities in the QWP and PBS.

Although in theory a complete correction of higher order non-linearities introduced by non-idealities in the QWP and PBS is possible, as demonstrated in section 6.5, unwanted back reflections will still result in non-linearities. A key source of potential back reflections that was not considered in the design process for this interferometer is the photodiode detectors, with the difference made by tilting the photodiode detectors clear in the reduction in non-linearity from figure 7.13 to 7.14. The prototype interferometer was constructed using photodiodes with integrated

condensing lenses, and as such a small portion of the back reflected light may be expected to have been scattered from the uncoated lens surfaces, back along the beam path. Photodiode back reflections do not, however, in isolation introduce the non-linearities with a once-per-fringe period seen in figure 7.14 in this interferometer design, as demonstrated in figure 6.29. The presence of photodiode back reflections, even with the detectors tilted, is confirmed by figure 7.16, with a small peak visible for in V_{ref} at a spatial period of $\lambda/2$. These findings indicate that the once-per-fringe non-linearity observed in figure 7.14 is likely to be the result of interactions between some combination of photodiode back reflections and some additional back reflection source.

In order to further investigate the non-linearities seen in figure 7.14, and to allow the peak-to-peak non-linearity to be quantified in a lower noise environment, the non-linearity of the prototype improved plane mirror differential interferometer design was also measured against the NPL XRI facility. The results of this measurement are shown in figure 7.17. The measurement arm normalisation corrected was found in this case to reduce peak-to-peak non-linearities from (56.8 ± 0.4) pm to (23.0 ± 0.6) pm. When interpreting these results, it should be noted that the quoted uncertainties only given an estimate of uncertainty in the best fit of equation 7.2 to the available data; they do not describe how appropriate the fit is, and therefore the effects of other systematic errors, beyond non-linearity, are not accounted for, and may bias the result. Furthermore, these uncertainties are based upon a single alignment of the interferometer, and therefore do not quantify the repeatability of the non-linearities when the interferometer is moved, or realigned.

The effects of other systematic errors on the XRI measurement were minimised by calculating the optical wavelength based upon the XRI displacement. This process eliminates any linear errors between the x-ray and optical interferometers, and masks the effect of any geometrical cosine error that may be present. Optical wavelength fluctuations that occur within the measurement time period, along with position dependent angular errors in the XRI flexure stage, may, however, still influence the result. The data presented in figure 7.17 **B** shows a trace and retrace through the same set of XRI positions, as shown by figure 7.17 **A**. Trace and retrace errors can be seen to be effectively superposed, demonstrating that any wavelength fluctuations, or angular motions, were small, and had a minimal effect on the non-linearity measurement.

Quantifying the repeatability of non-linearities with realignment of the interferometer is more challenging, in particular when the exact source of the non-linearities is not known. The alignment procedure described for this interferometer, first maximising the outputs of the sine and cosine signals with the reference arm blocked, is expected to introduce a degree of repeatability, however, the accuracy of the alignment cannot be quantified, and variability between set ups, and operators, is therefore possible. This suggests that including the potential for angle measurement, as in [51], may be beneficial in displacement measuring optical interferometer designs, not only for the measurement of angular motion, but also to ensure repeatability of alignment.

Due to the restricted displacement range measured with the XRI, the (spatial) spectral content of the reference signals could not be analysed, however, the form

of the non-linearities observed in figure 7.17 is comparable to those seen in figure 7.14, and therefore the source of the non-linearities is hypothesised to be the same; a combination of back reflections from the photodiodes and scattered light or higher order back reflections from other optical components.

Back reflections from the photodiodes may be reduced in future iterations of this interferometer design by using photodiodes with flat windows, and tilting the photodiodes, such that any back reflections are misaligned. One interferometer design described in the literature [28] made use of QWP and polarisers in front of the photodiodes, acting as an optical isolator for first order back reflections, and this may also be a possible approach. In the prototype interferometer, the angle of the photodiodes is not constrained by the design. Any future design should ensure the photodiode angles are fixed, so as to ensure back reflection effects are minimised, and to ensure the performance of different interferometers built to the same design is repeatable.

Identifying the sources of the other suspected back reflection, or back reflections, is more challenging, as both the relative phase and location of any back reflections, along with the interactions between multiple back reflection sources, all influence the observed non-linearity. Back reflections from the QWP are able to introduce, in combination with the photodiode back reflections, the once-per-fringe non-linearity observed in figure 7.13, along with the increase in non-linearity at a spatial period of $\lambda/2$ relative to the reference arm normalised case, as demonstrated in figure 6.28. QWP back reflections therefore appear to be a possible candidate, however, the QWP was deliberately tilted, and therefore any back reflections from the QWP should not reach the detector. One possible explanation is that, although first order back reflections were misaligned, some combination of higher order reflections may reach the detectors. Baffles were included in the design to attempt to minimise the possibility of this occurring, however, the baffles had relatively large apertures (4mm) relative to the $1/e^2$ beam diameter (1.2mm). Misaligned back reflections cannot be modelled with the combined scattering and transfer matrix modelling technique [85], and therefore further experimental work is required to identify if this is the cause of the residual non-linearities observed in figure 7.14, with the next step being to place smaller apertures in front of each detector.

The work presented here has demonstrated both the strengths, and the weaknesses, of the measurement arm normalisation non-linearity correction approach described in section 6.5. The sensitivity to back reflections predicted by the modelling work has been demonstrated experimentally, and has been found to limit the performance of the interferometer, however, even in the presence of significant back reflections, a reduction in non-linearity was still achieved relative to the reference arm normalised case. An additional benefit, not previously considered, is the additional diagnostic information provided by having access to both measurement and reference arm intensity references. Frequency space representations of these outputs, as shown in figures 7.15 and 7.16, provide significant additional information about the performance of the interferometer. In this work, this information permitted the presence of back reflections from the photodiodes to be confirmed. With further development of the technique it may be possible to use these additional outputs to place an upper limit on the magnitude of any residual non-linearities,

even in cases where correction is not possible, and this would have significant benefits when estimating uncertainties in interferometric measurements of displacement. The importance of in-situ estimates of residual non-linearities, previously discussed in chapter 3, are further reinforced by the findings of the survey of existing NPL interferometers, presented in figures 7.7 and 7.8, where the lack of repeatability of the non-linearity can be clearly seen.

7.7 Summary

Sources of non-linearity errors in the NPL plane mirror differential interferometer design have been identified through a combination of modelling work, employing the combined scattering and transfer matrix model developed in chapter 5, and experimental measurement. Non-idealities in polarising optical components, and unwanted back reflections, have been identified as the main cause of non-linearities in this interferometer design. Based upon these findings, an improved optical design has been developed and tested, with measurements made using both a nanopositioning stage and the NPL XRI facility.

Measurements of the non-linearity of ten existing NPL plane mirror differential interferometers have revealed variation in the non-linearity between individual interferometers built to the same design. This variation has been attributed to back reflections from optical surfaces that are nominally perpendicular to the axis of beam propagation, but that in reality are slightly misaligned to varying degrees. This has been mitigated in the improved design by introducing deliberate and repeatable tilt angles to all optical components.

By tilting polarising optical components, the performance is likely to have been degraded. To compensate for this, along with pre-existing non-idealities in real world polarising optical components, the improved interferometer design incorporates the measurement arm normalisation non-linearity correction technique described in section 6.5. In principle, this technique allows for the complete correction of non-linearities introduced by non-ideal polarising optics in this interferometer design, in the absence of other error sources. Where other error sources, for example unwanted back reflections, are present, only a partial correction can be expected.

A prototype improved plane mirror differential interferometer has been constructed, and found to have peak-to-peak non-linearities of (23.0 ± 0.6) pm when employing the measurement arm normalisation non-linearity correction, reduced from (56.8 ± 0.4) pm when this correction is not employed. The exact source of these non-linearities could not be conclusively identified, however, the presence of back reflections from the photodiode detectors has been confirmed by the presence of a component in the reference arm intensity reference output that is periodic with harmonics of the illuminating wavelength. Photodiode back reflections alone do not fully explain the observed non-linearity. Some additional error source must be present, and scattered light or higher order misaligned reflections entering the detectors is suggested as a cause. Further work is required to verify this suggestion, which may be investigated by placing apertures in front of the photodiodes.

The work presented here has verified the potential of the measurement arm

normalisation non-linearity correction, both for the correction of higher order non-linearities, and for diagnostic purposes. The additional intensity references allowed the presence of back reflections from the photodiode detectors to be conclusively confirmed. Picking off intensity references from each arm of the interferometer is therefore suggested not only as a correction technique, but also as an approach for investigating the source of non-linearities in homodyne interferometry, even if the pick off beam splitter is only temporarily installed for diagnostic purposes.

Future work may focus both on identifying the source of the residual non-linearity in the prototype improved plane mirror differential interferometer design, and on further exploring the potential of the measurement arm normalisation technique. Two key aspects of this technique require further investigation. Firstly, the conditions under which the technique is effective should be identified. Modelling work has been used here to explore possible applications for one interferometer design, however, more general guidelines are desirable. Secondly, the potential to place an upper limit on peak-to-peak non-linearity by observing the periodic content of the intensity references should be researched. In-situ non-linearity analysis is desirable, as demonstrated by the variations observed in figure 7.7, but often challenging with existing techniques, and if the intensity references could be used to estimate non-linearities, this would allow uncertainties to be placed on nanoscale displacement measurements made with homodyne interferometers with a greater degree of confidence.

Chapter 8

Conclusions and future work

8.1 Conclusions

This thesis has explored the performance limits of homodyne displacement measuring optical interferometry, with the goal of developing techniques to reduce the measurement uncertainty of homodyne interferometers at the nanoscale.

In the course of this exploration, existing precision interferometer designs have been reviewed, and the key strengths and weaknesses of leading designs identified. Sources of uncertainty in displacement measurements made with homodyne interferometry have also been reviewed. At the nanoscale (for measurements made in air), non-linearities may dominate the measurement uncertainty if the dead path is small, and measured accurately. For measurements made in the vacuum, non-linearities may dominate over larger length scales.

Based on the findings of the review of existing interferometer designs, phase-quadrature coating based interferometers were selected as a promising candidate for further research. As such, the design and fabrication of phase-quadrature coatings was investigated, and the design of such a coating was optimised.

Another key finding of the review process was that residual non-linearities in homodyne interferometers are dominated by higher order effects. Such higher order effects, arising as a result of the formation of weak optical cavities within the interferometer optics, are challenging to model with existing techniques, and this drove the development of a novel technique for modelling higher order non-linearities in homodyne interferometers, the combined scattering and transfer matrix model. This model has been developed and published [85] alongside an associated open source software implementation [159]. The results gathered with this model have led to the identification of previously unknown non-linearity sources in the NPL plane mirror differential interferometer design.

Even when non-linearity sources are identified, it is not always possible to eliminate them completely, and as a result, three novel non-linearity correction methodologies have also been developed in the course of this PhD: the two-wavelength [69] (section 6.3), multiple intensity reference [171] (section 6.4) and measurement arm normalisation [172] (section 6.5) approaches, each aiming to fill a specific weakness in the existing ellipse fitting Heydemann correction [87]. Existing non-linearity corrections have also been reviewed, with a focus on comparing the performance of the

various ellipse fitting algorithms that may be used to implement the Heydemann correction [87].

Finally, by combining the measurement arm normalisation approach to the correction of higher order non-linearities, with design improvements driven by modelling work, an improved plane mirror differential interferometer design has been developed, based upon an existing NPL design, and the non-linearities of this improved design have been investigated. XRI measurement revealed a residual non-linearity of (23.0 ± 0.6) pm when the measurement arm normalisation technique was applied, reduced from 40 pm for the existing NPL design [54]. Further improvements are expected to be possible in the future, with measures taken to reduce the effects of scattered light.

Reviewing existing interferometer designs, it is apparent that although Fabry-Perot designs currently achieve the lowest measurement uncertainties, they are limited in practical metrological application by limited working range, optical and electronic complexity, and cost. Of the other approaches to precision displacement measuring interferometry at the nanoscale, both state of the art homodyne and heterodyne interferometers currently achieve similar performance with regards to non-linearity. Non-linearity sources in homodyne interferometry are, however, better understood, and homodyne interferometers, particularly where phase-quadrature coatings are employed, are optically simpler. A theme that runs throughout the literature, is that in general, optically simpler designs, with fewer polarising components and optical surfaces, tend to have lower non-linearities. This statement should be cautioned with the fact that there are very few interferometer designs published that are capable of achieving non-linearities in the tens of picometre range, where higher order effects begin to become significant, and therefore the sample size is small. Unwanted back reflections and polarisation leakage have, however, been suggested as causes of the residual non-linearities in a number of cases, for both homodyne and heterodyne interferometers [15, 28, 59–61, 63, 82, 84, 85], and this reinforces the idea that simpler optical designs are likely to achieve lower non-linearities, all else being equal. This conclusion is also backed up by the data presented in section 6.4 of this thesis, where in a simple corner cube based phase-quadrature Michelson interferometer, non-linearities, in the full fringe case, were reduced below the 7.5 pm noise floor. As non-linearities are not the only uncertainty source, such simple interferometers are, however, often impractical, with plane mirror differential designs required to reduce the effects of air turbulence and mechanical vibrations on the displacement measurement, minimise the dead path, and reduce the influence of thermal drifts. As phase-quadrature coating based interferometers have a simple optical design, and well understood non-linearity sources, as compared to polarisation based homodyne interferometers, this interferometer design was selected as the target for the research presented in this thesis.

A key, and novel, finding of the review of uncertainty sources in homodyne interferometry was that where higher order non-linearities occur, non-linearities can neither be corrected, or measured, based only upon the quadrature signals produced by the interferometer. The presence of higher order harmonics permit the formation of quadrature signals that, when plotted against each other, form a perfectly circular Lissajous figure, despite having a distorted phase. This is a particularly important

finding to highlight, as amplitude deviations of the Lissajous figure from the ideal circular form have previously been used to estimate residual non-linearities [50], a technique that this finding demonstrates will underestimate the non-linearity where higher order effects are present.

A less novel, but important to state, conclusion of this review was that at the nanoscale, non-linearities may dominate the measurement uncertainty, with the tens of picometres of non-linearity achieved by leading interferometer designs still in excess of the shot noise limit derived in section 3.4 at the optical powers and bandwidths typically encountered in precision interferometry. A final conclusion of this review was that although optimal values for the quadrature phase shift, and reflectivity to transmission ratio of phase-quadrature coatings exist, the additional uncertainty introduced by deviations from these ideal values is small, and does not currently limit interferometer performance. As a result, the residual variability in the phase-quadrature coating parameters discussed in chapter 4 also does not limit interferometer performance.

Following the identification of non-linearities as a factor limiting the nanoscale performance of homodyne interferometry, and before non-linearities can be eliminated, the sources of non-linearity must be identified. Although, as already discussed, the general sources of non-linearity in homodyne interferometry are well established in the literature, the specific sources for an individual design are more challenging to characterise. As discussed in chapter 3, higher order non-linearities often dominate the non-linearity in cases where a Heydemann correction is employed, and these higher order non-linearity sources are challenging to model with existing approaches. The combined scattering and transfer matrix model was developed specifically to model higher order non-linearities in homodyne displacement measuring optical interferometers. The novelty of this approach, over a related approach implemented in the FINESSE [153] software, is the ability to model the effects of cavity formation within polarising optics. By simultaneously modelling the propagation of light in two orthogonal polarisation states, chosen to be the S and P states relative to the optical components in the published implementation, coupling between the two polarisation states may be modelled. This ability proved crucial to the identification of the sources of residual non-linearity in the NPL plane mirror differential interferometer design, where complex interactions between non-idealities in various polarising optical components occur.

Some non-linearity sources identified with the combined scattering and transfer matrix model may be eliminated through improved optical design, however, in cases where this is not possible, non-linearity correction algorithms are required. The existing Heydemann correction [87] is widely used in the field of homodyne interferometry, and corrects for elliptical non-linearities introduced by errors in the signal offsets, gain ratios and quadrature phase error. This correction relies up ellipse fitting, raising the questions as to whether the specific ellipse fitting algorithm employed affects the performance of the correction. A selection of ellipse fitting algorithms were compared in section 6.2.1. The ODR algorithm was found to offer marginally better performance than the other algorithms considered in the full fringe case, however, performance improvements were minor, and negligible compared to typical levels of higher order non-linearities. Ellipse fitting corrections, as

demonstrated in section 3.3.3, cannot correct for higher order non-linearities, and therefore alternative approaches are required.

The novel two-wavelength non-linearity correction technique [69] described in this thesis represents the first technique capable of correcting for higher order non-linearities in a purely optical instrument. The two-wavelength correction is applicable to any interferometer that suffers from periodic non-linearity errors. In this work the two-wavelength correction has been demonstrated at a proof-of-concept level, however, the approach is in principle applicable to any interferometer that suffers from periodic non-linearity errors, and as such it is anticipated that the application of this correction to state of the art interferometer designs may allow non-linearities to be reduced below the 10 pm peak-to-peak level. Beyond non-linearity correction, two-wavelength interferometry may have additional benefits for refractive index compensation [88], and therefore the development of two-wavelength interferometer designs is a promising avenue for future work.

The two-wavelength correction adds significant additional cost and complexity to the interferometer, both in terms of optical and electronic design, and data processing. Displacement ranges of several optical fringes are also required to reach the maximum performance of the correction. Alternative approaches are therefore still desirable. The measurement arm normalisation approach described in section 6.5, and applied in chapter 7, is one such approach. This novel approach allows for the correction of some higher order non-linearities. Non-linearities that are suitable for correction with this approach may be identified through application of the combined scattering and transfer matrix model [85]. Simply requiring moving the existing intensity reference commonly used in homodyne interferometry to a different location within the interferometer, the measurement arm normalisation approach adds no additional optical or data processing complexity, and therefore may be implemented, where applicable, with few disadvantages. As highlighted in chapter 7, the real-world performance of this approach may be limited by unwanted back reflections, and therefore in some cases the two-wavelength correction [69] may still be required.

In addition to the inability to correct for higher order non-linearities, the Heydemann correction has another key weakness; where the displacement covers a range of less than one full optical fringe, corresponding to a full ellipse when plotted as a Lissajous figure, the performance of ellipse fitting algorithms is severely degraded. This effect was investigated in section 6.2.1, with the conclusion that even when ellipse specific fitting algorithms, designed for the sub-fringe case, were employed, performance was still poor. Specific figures are dependent on noise levels in the quadrature signals. Under the conditions tested (shot noise limited, 1 mW illuminating power, 100 kHz bandwidth), non-linearities increased above 10 pm for displacements covering less than approximately 40 nm for a double pass interferometer. Precision sub-fringe displacement measurement has a number of applications, including the out of plane axis of metrological atomic force microscopes [55], and monitoring the out of plane motion of nanopositioning stages [6], and in these cases, measurement uncertainties may be limited by the poor sub-fringe performance of the Heydemann correction. To fill this performance gap in the Heydemann correction, and therefore extend the performance envelope of precision homodyne interferometry further into the nanoscale, another novel non-linearity correction methodology was developed,

the multiple intensity reference correction [171]. This correction was found in experimental work to outperform the Heydemann correction for displacements covering less than approximately 120 nm when applied to a single pass interferometer configuration operating at approximately 632.8 nm. Non-linearities remained below 100 pm for displacement ranges greater than approximately 25 nm, again for a single pass configuration, representing a significant increase in the dynamic range of precision homodyne interferometry.

Having developed novel techniques for the identification and correction of non-linearity in homodyne interferometry, the final piece of work presented in this thesis focussed on the development of an improved plane mirror differential interferometer design, based upon an existing NPL design. Through a combination of modelling work, utilising the combined scattering and transfer matrix model [85], and experimental measurement, non-linearity sources in the existing NPL plane mirror differential interferometer design were identified. The key sources on non-linearity were found to be the unwanted back reflections, and non-idealities in the polarising optical components, both introducing higher order non-linearities, typically in the range of tens of picometres. These findings were in line with previous measurements of this interferometer design, which found peak-to-peak non-linearities of 40 pm. Significant variability was found between the individual interferometers tested, and this further reinforces the conclusion that in-situ non-linearity evaluation is desirable. Based on these findings, the improved design tilted all optical surfaces that were previously perpendicular to the beam axis by differing amounts, aiming to misalign back reflections. A beam splitter was also added to the return paths of the measurement and reference arms of the interferometer, allowing the the measurement arm normalisation non-linearity correction approach described in section 6.5 to be utilised, with modelling work indicating that in this case, the measurement arm normalisation approach should allow for the complete correction of non-linearity introduced by non-ideal performance of the QWP and PBS. A secondary benefit of this additional beam splitter is that the multiple intensity reference correction technique [69] could also be applied for sub-fringe displacements, although this was not tested. The non-linearity of the improved design was investigated, and found to be limited by back reflections from the photodiodes, an error source that had not been considered during the design process. Tilting the photodiodes to reduce the effect of any back reflections allowed the non-linearity to be reduced. Non-linearities were measured against the NPL XRI, and found to be (23.0 ± 0.6) pm when the measurement arm normalisation algorithm was applied. Without applying this correction, non-linearities rose to (56.8 ± 0.4) pm. The sources of the residual non-linearity could not be confirmed conclusively, however, the continued presence of weak back reflections from the photodiodes was confirmed. This, in combination with scattered light from higher order reflections reaching the detectors, is suggested as the cause of the residual non-linearity.

Summarising, this PhD has made a number of novel contributions to the field of homodyne interferometry, two of which have been published [69, 85], and one of which has been submitted to a journal [171]. A novel modelling technique has been developed and published [85, 159], that permits higher order non-linearities to be analysed. Three novel non-linearity corrections have been developed. Two of these,

the two-wavelength [69] and measurement arm normalisation corrections, allow for the correction of higher order non-linearities, something that was not possible with a purely optical instrument prior to this PhD. The third correction technique, the multiple intensity reference correction [171], fills an important gap in the performance of existing non-linearity correction algorithms, extending the performance of precision homodyne interferometry in the sub-fringe displacement case. Combined, these contributions form a tool kit for the development of precision homodyne displacement measuring interferometers.

8.2 Future work

A number of aspects of this thesis suggest the potential for future work, and these are outlined in this section.

8.2.1 Phase-quadrature coatings

Although the performance of the phase-quadrature coatings used in NPL interferometer designs does not currently limit the performance of the interferometers, this may become the case in the future if non-linearities are further reduced. Improving the repeatability of the phase-quadrature coating process is therefore desirable. Modifying the existing design for use at wavelengths other than 632.8 nm is also desirable, and improved repeatability would simplify the re-design and re-optimisation process.

8.2.2 Uncertainties

In chapter 3, the importance of measuring the dead path with a low uncertainty for measurements made in air was highlighted. Accurate dead path measurement is often challenging in practice. In some cases mechanical measurement of the separation of the measurement and reference mirrors may be sufficient, however, particularly in less stable environments, or over longer time periods, where the refractive index is likely to vary more, and hence the dead path uncertainty is a larger contributing factor, optical techniques may be required. This is essentially an absolute distance measurement problem, and may be accomplished optically through a number of methods, including frequency scanning interferometry [194], frequency modulated interferometry [32, 33], and synthetic wavelength approaches [184]. The application of these techniques to the measurement of dead path in displacement measuring interferometry should be considered in future work. Where the dead path may be controlled, and can be set to zero, simplified frequency scanning or modulating approaches may be possible, and these should also be investigated.

8.2.3 Two-wavelength non-linearity correction

In this thesis the two-wavelength non-linearity correction [69] has been applied only as a proof of concept, with the corrected non-linearities deliberately introduced into a simple corner cube base Michelson interferometer. A number of improvements

could be made to the implementation. Firstly, the two-wavelengths were multiplexed together in the time domain, limiting the sampling rate of the interferometer, and increasing the complexity of data processing. Future work could instead spatially multiplex the two-interferometers. As one wavelength sets the displacement scale, any small misalignments between the measurement axes of the two interferometers are not critical, and therefore, at the cost of doubling the required detection channels, truly simultaneous sampling could be achieved between the two interferometers, with no risk of cross talk as could occur with polarisation or dichroic based separation approaches.

Secondly, the corner cube Michelson configuration may be impractical for many metrological applications, where plane mirror differential configurations are often preferred. Taking the NPL plane mirror differential interferometer as an example, the wavelength dependence of the PBS and QWP present challenges. Achromatic and QWP are commercially available, however, as are QWP designed for operation at two specific wavelengths. PBS designed for broadband operation are also available, however, the extinction ratio is likely to be degraded. As discussed in section 6.5, errors in both the QWP and PBS may be compensated with the measurement arm normalisation technique. This approach could therefore be combined with the two-wavelength correction in future work to correct residual non-linearities in further developments of the NPL plane mirror differential interferometer design. A schematic of a proposed spatially separated two-wavelength interferometer, following the improved NPL plane mirror differential interferometer design, is shown in figure 8.1.

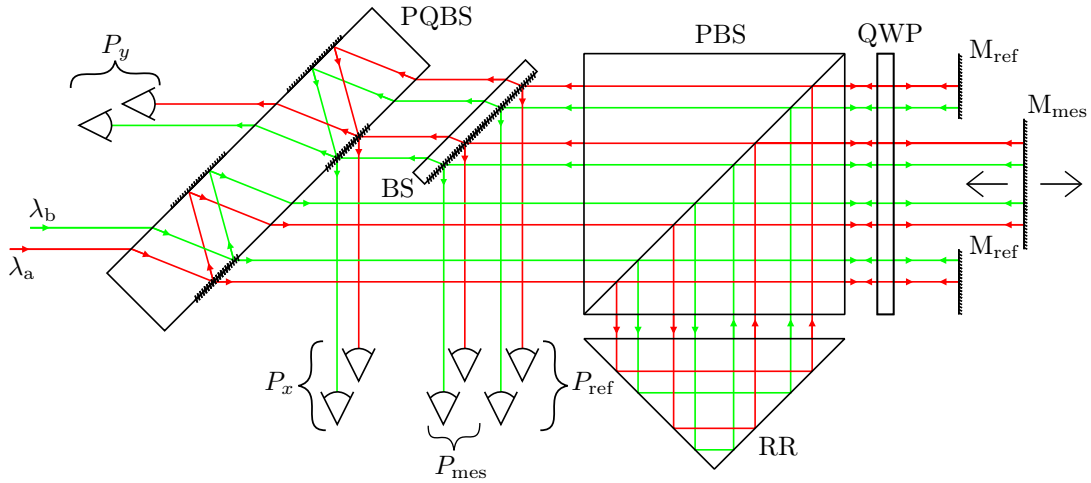


Figure 8.1: Proposed spatially separated two-wavelength interferometer, following the improved NPL plane mirror differential interferometer design.

Finally, the two-wavelength non-linearity correction may compliment the application of multiple wavelength refractive index compensation methods. Both two [88] and three [195] wavelength based refractive index compensation methods have been proposed in the literature, and both would be compatible with the two-wavelength non-linearity correction. The main factor limiting the practical application of these

techniques is the scaling of the interferometric phase change uncertainty by a large factor determined by the wavelength spacing [195]. Combining high precision interferometers employing the two-wavelength non-linearity correction, which have the potential for very low non-linearities, and therefore phase uncertainties, with multiple wavelength refractive index compensation methods may therefore lead to improved refractive index compensation.

Relating back to the previously discussed dead path uncertainty problem, the potential application of multiple wavelength interferometry for non-linearity correction [69], refractive index compensation [88, 195] and dead path measurement [184] has been discussed. The development of a metrological light source having a high degree of frequency stability, as required for precision interferometry, whilst also generating multiple wavelengths, potentially with a degree of frequency tunability, is therefore suggested.

8.2.4 Measurement arm normalisation non-linearity correction

The second higher order non-linearity correction technique proposed in this thesis, the measurement arm normalisation technique, also suggests further work. Currently, although specific scenarios may be tested with the combined scattering and transfer matrix model [85], no general guidelines have been developed for when the measurement arm normalisation non-linearity correction may offer a complete correction. The development of such guidelines is an important target for future work.

An additional aspect of the measurement arm normalisation non-linearity correction that is of interest for future work is the diagnostic application of intensity references taken separately from each arm of the interferometer. The utility of this approach was demonstrated in chapter 7, where the presence of components that were periodic with harmonics of the illuminating wavelength in the reference output of the interferometer allowed the presence of back reflections from the photodiodes to be confirmed. In combination with the combined scattering and transfer matrix model, with further research, similar effects may provide a useful diagnostic tool for identifying non-linearity sources in homodyne interferometry.

In addition to diagnostic applications, it may be possible to place an upper limit on non-linearities based upon the intensity references. If this were to be the case, the ability to place limits on the non-linearity of the interferometer would provide an approach to in-situ non-linearity analysis, the importance of which has been noted throughout this thesis.

8.2.5 Improved NPL plane mirror differential interferometer

Residual non-linearities are still present in the improved NPL plane mirror differential interferometer prototype described in chapter 7. Future work should identify the source of these non-linearities, and attempt to eliminate them.

Notably, the presence of residual back reflections from the photodiodes has been confirmed, and these should be eliminated by moving to use photodiodes with flat

windows, and ensuring the photodiodes remain tilted, such that any back reflections are misaligned. Another likely source of residual non-linearity is scattered light. Whilst, due to the tilted optics, first order back reflections from the various optical surfaces cannot reach the detectors, is it possible higher order reflections may do. This problem could be eliminated by placing apertures comparable in size to the beam diameter in front of each photodiode, and by reducing the size of the apertures in the beam baffles currently installed in the interferometer in an attempt to minimise scattered light.

Other improvements to the design are possible, including the use of a hollow cube corner retroreflector, as opposed to the current solid glass retroreflector. Although silver coated, polarisation rotation still occurs during propagation through the retroreflector, and this should be reduced for a hollow retroreflector. Additionally, a possible source of back reflections, the front surface of the retroreflector, would be eliminated. The addition of angle measurement, as in [51], may also be beneficial, for direct angle measurement, for repeatability of alignment, and for the minimisation or quantification of geometric errors.

List of publications

Journal papers

Bridges A, Yacoot A, Kissinger T, Tatam RP. Correction of periodic non-linearities in homodyne optical interferometry by measurement arm intensity normalisation. Submitted to Optics Letters. 2021.

Bridges A, Yacoot A, Kissinger T, Tatam RP. Multiple intensity reference interferometry for the correction of sub-fringe displacement non-linearities. Submitted to Measurement Science and Technology. 2021.

Bridges A, Yacoot A, Kissinger T, Humphreys D, Tatam RP. Correction of periodic displacement non-linearities by two-wavelength interferometry. Measurement Science and Technology. 2021;32(12):125202.

Bridges A, Yacoot A, Kissinger T, Tatam RP. Polarization-sensitive transfer matrix modeling for displacement measuring interferometry. Applied Optics. 2020;59(25):7694-704.

Conference presentations

Bridges A, Yacoot A, Kissinger T and Tatam R P. A homodyne displacement measuring interferometer with real time non-linearity correction. 5th Postgraduate Institute for Measurement Science Conference. Teddington, UK; 2020 (Poster presentation).

Bridges A, Yacoot A, Kissinger T and Tatam R P. A transfer and scattering matrix model of the NPL plane mirror differential interferometer. 4th Postgraduate Institute for Measurement Science Conference. Teddington, UK; 2019 (Poster presentation).

Bridges A, Yacoot A, Kissinger T and Tatam R P. A matrix modelling approach for displacement measuring interferometer design. The Enlighten Conference, Photonex Europe. Coventry, UK; 2019 (Oral presentation).

Bridges A, Yacoot A, Kissinger T and Tatam R P. Transfer matrix modelling for displacement measuring interferometer design. 8th Seminar on Nanoscale Calibration, Standards and Methods. Braunschweig, Germany; 2019 (Poster presentation).

Bridges A, Yacoot A, Kissinger T and Tatam R P. A comparison of non-linearity correction algorithms in optical interferometry. 19th International Conference and Exhibition, euspen. Bilbao, Spain; 2019 (Poster presentation).

Software publications

strapy (Scattering TRAnsfer matrix model, PYthon) <https://github.com/strapy-project/strapy> released under the MIT licence.

ctmm (C Transfer Matrix Model) <https://github.com/strapy-project/ctmm> released under the MIT licence.

Appendix A

Derivation of the covariance of Heydemann corrected quadrature signals

The covariance between the Heydemann corrected quadrature signals, as defined by equations 3.26 and 3.27 may be calculated from the definition of the covariance [112] as

$$\sigma_{C_x, C_y} = \langle C_x C_y \rangle - \langle C_x \rangle \langle C_y \rangle, \quad (\text{A.1})$$

where angle brackets denote the expectation value of the contained expression. Evaluating the expectation values making use of the law of the unconscious statistician [119] results in

$$\begin{aligned} \langle C_x \rangle &= \int (S_x - \alpha_x) f(S_x, S_y) dS_x dS_y, \\ &= \mu_x - \alpha_x, \end{aligned} \quad (\text{A.2})$$

$$\begin{aligned} \langle C_y \rangle &= \frac{1}{\cos(\delta')} \int \left((S_x - \alpha_x) \sin(\delta') + G_{xy} (S_y - \alpha_y) \right) f(S_x, S_y) dS_x dS_y, \\ &= \frac{1}{\cos(\delta')} \left((\mu_x - \alpha_x) \sin(\delta') + G_{xy} (\mu_y - \alpha_y) \right), \end{aligned} \quad (\text{A.3})$$

$$\begin{aligned} \langle C_x C_y \rangle &= \frac{1}{\cos(\delta')} \int (S_x - \alpha_x) \left((S_x - \alpha_x) \sin(\delta') \right. \\ &\quad \left. + G_{xy} (S_y - \alpha_y) \right) f(S_x, S_y) dS_x dS_y, \end{aligned} \quad (\text{A.4})$$

$$\begin{aligned} &= \frac{1}{\cos(\delta')} \left(\alpha_x^2 \sin(\delta') + \alpha_x \alpha_y G_{xy} + \sin(\delta') \langle S_x^2 \rangle \right. \\ &\quad \left. - \left(2\alpha_x \sin(\delta') + \alpha_y G_{xy} \right) \langle S_x \rangle - \alpha_x G_{xy} \langle S_y \rangle, \right. \\ &\quad \left. + G_{xy} \int S_x S_y f(S_x, S_y) dS_x dS_y \right) \end{aligned} \quad (\text{A.5})$$

$$\begin{aligned} &= \frac{1}{\cos(\delta')} \left(\alpha_x^2 \sin(\delta') + \alpha_x \alpha_y G_{xy} + \sin(\delta') (\mu_x^2 + \sigma_x^2) \right. \\ &\quad \left. - \left(2\alpha_x \sin(\delta') + \alpha_y G_{xy} \right) \mu_x - \alpha_x G_{xy} \mu_y \right. \\ &\quad \left. + G_{xy} (\mu_x \mu_y + \sigma_{x,y}) \right) \end{aligned} \quad (\text{A.6})$$

$$\begin{aligned} &= \frac{G_{xy}}{\cos(\delta')} \left((\alpha_x - \mu_x) (\alpha_y - \mu_y) + \sigma_{x,y} \right) \\ &\quad + \left((\alpha_x - \mu_x)^2 + \sigma_x^2 \right) \tan(\delta'), \end{aligned} \quad (\text{A.7})$$

where a bivariate Gaussian distribution [52] has been assumed such that

$$\begin{aligned} f(S_x, S_y) &= \frac{1}{2\pi\sigma_x\sigma_y\sqrt{1-\frac{\sigma_{x,y}}{\sigma_x\sigma_y}}} \exp \left(-\frac{1}{2\left(1-\frac{\sigma_{x,y}}{\sigma_x\sigma_y}\right)} \left(\frac{(S_x - \mu_x)^2}{\sigma_x^2} \right. \right. \\ &\quad \left. \left. + \frac{(S_y - \mu_y)^2}{\sigma_y^2} - \frac{2\sigma_{x,y}(S_x - \mu_x)(S_y - \mu_y)}{\sigma_x^2\sigma_y^2} \right) \right), \end{aligned} \quad (\text{A.8})$$

and the definition of the variance has been used to establish

$$\langle S_x^2 \rangle = \langle S_x \rangle^2 + \sigma_x^2, \quad (\text{A.9})$$

$$= \mu_x^2 + \sigma_x^2. \quad (\text{A.10})$$

Finally, equations A.1 to A.8 yield a covariance between the corrected quadrature signals of

$$\sigma_{C_x, C_y} = \frac{G_{xy}}{\cos(\delta')} \sigma_{x,y} + \tan(\delta') \sigma_x^2. \quad (\text{A.11})$$

Appendix B

Phase uncertainty propagation calculations

The elliptical non-linearity corrected phase of a quadrature interferometer may be evaluated directly from the ellipse fit coefficients A to E and the quadrature signals S_x and S_y as

$$\varphi = \tan^{-1} \left(\frac{\sqrt{4AB - C^2} (E + CS_x + 2BS_y)}{2B(D + 2AS_x) - C(E + CS_x)} \right) + e_{\text{NL}}. \quad (\text{B.1})$$

Propagating uncertainties through this equation using standard partial derivative techniques [89], neglecting covariances between the quadrature signals and the fit coefficients, results in an expression for the variance in the interferometric phase, given by

$$\begin{aligned} \sigma_{\varphi,\varphi} = & \left(\sigma_{\varphi,\varphi}|_{S_x,S_x} \sigma_{S_x,S_x} + \sigma_{\varphi,\varphi}|_{S_y,S_y} \sigma_{S_y,S_y} + \sigma_{\varphi,\varphi}|_{S_x,S_y} \sigma_{S_x,S_y} \right. \\ & + \sigma_{\varphi,\varphi}|_{A,A} \sigma_{A,A} + \sigma_{\varphi,\varphi}|_{A,B} \sigma_{A,B} + \sigma_{\varphi,\varphi}|_{A,C} \sigma_{A,C} + \sigma_{\varphi,\varphi}|_{A,D} \sigma_{A,D} \\ & + \sigma_{\varphi,\varphi}|_{A,E} \sigma_{A,E} + \sigma_{\varphi,\varphi}|_{B,B} \sigma_{B,B} + \sigma_{\varphi,\varphi}|_{B,C} \sigma_{B,C} + \sigma_{\varphi,\varphi}|_{B,D} \sigma_{B,D} \\ & + \sigma_{\varphi,\varphi}|_{B,E} \sigma_{B,E} + \sigma_{\varphi,\varphi}|_{C,C} \sigma_{C,C} + \sigma_{\varphi,\varphi}|_{C,D} \sigma_{C,D} + \sigma_{\varphi,\varphi}|_{C,E} \sigma_{C,E} \\ & \left. + \sigma_{\varphi,\varphi}|_{D,D} \sigma_{D,D} + \sigma_{\varphi,\varphi}|_{D,E} \sigma_{D,E} + \sigma_{\varphi,\varphi}|_{E,E} \sigma_{E,E} \right) \\ & \left/ \left(64B^4 (4AB - C^2) \left(-4AB^2 S_y^2 + B(-4AS_y(CS_x + E) - (2AS_x + D)^2 \right. \right. \right. \\ & \left. \left. \left. + C^2 S_y^2) + (CS_x + E)(C(AS_x + CS_y + D) - AE) \right)^4 \right) + \sigma_{\text{NL,NL}} \right) \end{aligned} \quad (\text{B.2})$$

where $\sigma_{\varphi,\varphi}|_{i,j}$ is the numerator component of the coefficient of $\sigma_{i,j}$. The forms of these coefficients are given in the following equations:

$$\sigma_{\varphi,\varphi}|_{S_x,S_x} = B^2 (C^2 - 4AB)^2 (C(CS_y + D) - 2A(2BS_y + E))^2 ((4AB - C^2) (2BS_y + CS_x + E)^2 + (C(CS_x + E) - 2B(2AS_x + D))^2)^2, \quad (\text{B.3})$$

$$\begin{aligned} \sigma_{\varphi,\varphi}|_{S_y,S_y} = & 16B^4 (C^2 - 4AB)^2 (C(CS_x + E) - 2B(2AS_x + D))^2 \\ & \left(4AB^2S_y^2 + B \left(4AS_y(CS_x + E) + (2AS_x + D)^2 - C^2S_y^2 \right) \right. \\ & \left. - (CS_x + E)(C(AS_x + CS_y + D) - AE) \right)^2, \quad (\text{B.4}) \end{aligned}$$

$$\begin{aligned} \sigma_{\varphi,\varphi}|_{S_x,S_y} = & 2B^2 (C^2 - 4AB)^2 (2B(2AS_x + D) - C(CS_x + E))(C(CS_y + D) \\ & - 2A(2BS_y + E)) \left((4AB - C^2) (2BS_y + CS_x + E)^2 \right. \\ & \left. + (C(CS_x + E) - 2B(2AS_x + D))^2 \right)^2, \quad (\text{B.5}) \end{aligned}$$

$$\begin{aligned} \sigma_{\varphi,\varphi}|_{A,A} = & B^2 (2BS_y + CS_x + E)^2 (2B(D - 2AS_x) + C(CS_x - E))^2 \\ & \left((4AB - C^2) (2BS_y + CS_x + E)^2 + (C(CS_x + E) - 2B(2AS_x + D))^2 \right)^2, \quad (\text{B.6}) \end{aligned}$$

$$\begin{aligned} \sigma_{\varphi,\varphi}|_{A,B} = & 2B(2BS_y + CS_x + E)(2B(D - 2AS_x) + C(CS_x - E)) \\ & \left((4AB - C^2) (2BS_y + CS_x + E)^2 + (C(CS_x + E) - 2B(2AS_x + D))^2 \right)^2 \\ & \left(4A^2BS_x(2BS_y - CS_x - E) + A(4B^2DS_y - 2B(3CS_y + D)(CS_x + E) \right. \\ & \left. + C^3S_x^2 - CE^2) + C^2(CS_y + D)(CS_x + E) \right), \quad (\text{B.7}) \end{aligned}$$

$$\begin{aligned} \sigma_{\varphi,\varphi}|_{A,C} = & -2B^2(2BS_y + CS_x + E)(2B(D - 2AS_x) + C(CS_x - E)) \\ & \left((4AB - C^2) (2BS_y + CS_x + E)^2 + (C(CS_x + E) - 2B(2AS_x + D))^2 \right)^2 \\ & (C(D - 2AS_x)(2BS_y + E) - 2A(2BS_x(2AS_x + D) + 2BES_y + E^2) \\ & + 2C^2S_x(AS_x + D) + C^3S_xS_y), \quad (\text{B.8}) \end{aligned}$$

$$\begin{aligned} \sigma_{\varphi,\varphi}|_{A,D} = & -2B^2 (4AB - C^2) (2BS_y + CS_x + E)^2 (2B(D - 2AS_x) \\ & + C(CS_x - E)) \left((4AB - C^2) (2BS_y + CS_x + E)^2 \right. \\ & \left. + (C(CS_x + E) - 2B(2AS_x + D))^2 \right)^2, \quad (\text{B.9}) \end{aligned}$$

$$\begin{aligned}\sigma_{\varphi,\varphi}|_{A,E} = & 2B^2 (4AB - C^2) (2AS_x + CS_y + D)(2BS_y + CS_x + E) \\ & (2B(D - 2AS_x) + C(CS_x - E)) \left((4AB - C^2) (2BS_y + CS_x + E)^2 \right. \\ & \left. + (C(CS_x + E) - 2B(2AS_x + D))^2 \right)^2, \quad (\text{B.10})\end{aligned}$$

$$\begin{aligned}\sigma_{\varphi,\varphi}|_{B,B} = & \left((4AB - C^2) (2BS_y + CS_x + E)^2 + (C(CS_x + E) \right. \\ & \left. - 2B(2AS_x + D))^2 \right)^2 \left(4A^2BS_x(2BS_y - CS_x - E) + A(4B^2DS_y \right. \\ & \left. - 2B(3CS_y + D)(CS_x + E) + C^3S_x^2 - CE^2) + C^2(CS_y + D)(CS_x + E) \right)^2, \quad (\text{B.11})\end{aligned}$$

$$\begin{aligned}\sigma_{\varphi,\varphi}|_{B,C} = & -2B \left((4AB - C^2) (2BS_y + CS_x + E)^2 + (C(CS_x + E) \right. \\ & \left. - 2B(2AS_x + D))^2 \right)^2 \left(C(D - 2AS_x)(2BS_y + E) - 2A(2BS_x(2AS_x + D) \right. \\ & \left. + 2BES_y + E^2) + 2C^2S_x(AS_x + D) + C^3S_xS_y \right) \left(4A^2BS_x(2BS_y \right. \\ & \left. - CS_x - E) + A(4B^2DS_y - 2B(3CS_y + D)(CS_x + E) \right. \\ & \left. + C^3S_x^2 - CE^2) + C^2(CS_y + D)(CS_x + E) \right), \quad (\text{B.12})\end{aligned}$$

$$\begin{aligned}\sigma_{\varphi,\varphi}|_{B,D} = & -2B (4AB - C^2) (2BS_y + CS_x + E) \left((4AB - C^2) \right. \\ & \left. (2BS_y + CS_x + E)^2 + (C(CS_x + E) - 2B(2AS_x + D))^2 \right)^2 \\ & \left(4A^2BS_x(2BS_y - CS_x - E) + A(4B^2DS_y - 2B(3CS_y + D)(CS_x + E) \right. \\ & \left. + C^3S_x^2 - CE^2) + C^2(CS_y + D)(CS_x + E) \right), \quad (\text{B.13})\end{aligned}$$

$$\begin{aligned}\sigma_{\varphi,\varphi}|_{B,E} = & 2B (4AB - C^2) (2AS_x + CS_y + D) \left((4AB - C^2) \right. \\ & \left. (2BS_y + CS_x + E)^2 + (C(CS_x + E) - 2B(2AS_x + D))^2 \right)^2 \\ & \left(4A^2BS_x(2BS_y - CS_x - E) + A(4B^2DS_y - 2B(3CS_y + D) \right. \\ & \left. (CS_x + E) + C^3S_x^2 - CE^2) + C^2(CS_y + D)(CS_x + E) \right), \quad (\text{B.14})\end{aligned}$$

$$\begin{aligned}\sigma_{\varphi,\varphi}|_{C,C} = & B^2 \left((4AB - C^2) (2BS_y + CS_x + E)^2 + (C(CS_x + E) \right. \\ & \left. - 2B(2AS_x + D))^2 \right)^2 \left(C(D - 2AS_x)(2BS_y + E) - 2A \right. \\ & \left. (2BS_x(2AS_x + D) + 2BES_y + E^2) + 2C^2S_x(AS_x + D) + C^3S_xS_y \right)^2, \quad (\text{B.15})\end{aligned}$$

$$\begin{aligned}
\sigma_{\varphi,\varphi}|_{C,D} &= 2B^2 (4AB - C^2) (2BS_y + CS_x + E) \\
&\quad \left((4AB - C^2) (2BS_y + CS_x + E)^2 + (C(CS_x + E) - 2B(2AS_x + D))^2 \right)^2 \\
&\quad \left(C(D - 2AS_x)(2BS_y + E) - 2A(2BS_x(2AS_x + D) + 2BES_y + E^2) \right. \\
&\quad \left. + 2C^2S_x(AS_x + D) + C^3S_xS_y \right), \quad (\text{B.16})
\end{aligned}$$

$$\begin{aligned}
\sigma_{\varphi,\varphi}|_{C,E} &= -2B^2 (4AB - C^2) (2AS_x + CS_y + D) \\
&\quad \left((4AB - C^2) (2BS_y + CS_x + E)^2 + (C(CS_x + E) - 2B(2AS_x + D))^2 \right)^2 \\
&\quad \left(C(D - 2AS_x)(2BS_y + E) - 2A(2BS_x(2AS_x + D) + 2BES_y + E^2) \right. \\
&\quad \left. + 2C^2S_x(AS_x + D) + C^3S_xS_y \right), \quad (\text{B.17})
\end{aligned}$$

$$\begin{aligned}
\sigma_{\varphi,\varphi}|_{D,D} &= 16B^4 (C^2 - 4AB)^2 (2BS_y + CS_x + E)^2 \\
&\quad \left(4AB^2S_y^2 + B(4AS_y(CS_x + E) + (2AS_x + D)^2 - C^2S_y^2) \right. \\
&\quad \left. - (CS_x + E)(C(AS_x + CS_y + D) - AE) \right)^2, \quad (\text{B.18})
\end{aligned}$$

$$\begin{aligned}
\sigma_{\varphi,\varphi}|_{D,E} &= -2B^2 (C^2 - 4AB)^2 (2AS_x + CS_y + D)(2BS_y + CS_x + E) \\
&\quad \left((4AB - C^2) (2BS_y + CS_x + E)^2 + (C(CS_x + E) - 2B(2AS_x + D))^2 \right)^2, \quad (\text{B.19})
\end{aligned}$$

$$\begin{aligned}
\sigma_{\varphi,\varphi}|_{E,E} &= B^2 (C^2 - 4AB)^2 (2AS_x + CS_y + D)^2 \\
&\quad \left((4AB - C^2) (2BS_y + CS_x + E)^2 + (C(CS_x + E) - 2B(2AS_x + D))^2 \right)^2. \quad (\text{B.20})
\end{aligned}$$

Appendix C

Derivation of the covariance of the normalised quadrature signals

The normalised quadrature signals S_x and S_y are given in terms of the measure quadrature and reference voltages by

$$S_x = \frac{V_x}{V_r}, \quad (\text{C.1})$$

$$S_y = \frac{V_y}{V_r}. \quad (\text{C.2})$$

$$(\text{C.3})$$

Assuming V_x , V_y and V_r follow a trivariate Gaussian distribution with a probability density function given by [52]

$$f(V_x, V_y, V_r) = \frac{1}{\sqrt{8\pi^3|\Sigma|}} e^{-\frac{1}{2}(\mathbf{V}-\boldsymbol{\mu})^T \Sigma^{-1}(\mathbf{V}-\boldsymbol{\mu})}, \quad (\text{C.4})$$

where

$$\mathbf{V} = \begin{pmatrix} V_x \\ V_y \\ V_r \end{pmatrix}, \quad (\text{C.5})$$

$$\boldsymbol{\mu} = \begin{pmatrix} \mu_{V_x} \\ \mu_{V_y} \\ \mu_{V_r} \end{pmatrix}, \quad (\text{C.6})$$

$$\Sigma = \begin{pmatrix} \sigma_{V_x, V_x} & \sigma_{V_x, V_y} & \sigma_{V_x, V_r} \\ \sigma_{V_x, V_y} & \sigma_{V_y, V_y} & \sigma_{V_y, V_r} \\ \sigma_{V_x, V_r} & \sigma_{V_y, V_r} & \sigma_{V_r, V_r} \end{pmatrix}, \quad (\text{C.7})$$

and $|\Sigma|$ is the matrix determinant of Σ , the covariance matrix, the expectation values required to calculate the covariance [112] are given by

$$\langle S_x \rangle = \int S_x f(V_x, V_y, V_r), \quad (\text{C.8})$$

$$= \int \frac{V_x}{V_r} f(V_x, V_y, V_r), \quad (\text{C.9})$$

$$\langle S_y \rangle = \int S_y f(V_x, V_y, V_r), \quad (\text{C.10})$$

$$= \int \frac{V_y}{V_r} f(V_x, V_y, V_r), \quad (\text{C.11})$$

$$\langle S_x S_y \rangle = \int S_x S_y f(V_x, V_y, V_r), \quad (\text{C.12})$$

$$= \int \frac{V_x V_y}{V_r^2} f(V_x, V_y, V_r). \quad (\text{C.13})$$

The integrals in equations C.8 to C.13 cannot be evaluated due to the non-physical (but approximately valid) assumption of a Gaussian distribution for the reference voltage, which would permit the voltage, and therefore (with the previously made assumption that a dark voltage measurement has been made, and subtracted) optical power to cross zero and become negative, whilst having a non-zero probability density at zero, leading to a division by zero in the integral. Making the assumption that noise contributions will be small, second order Taylor expansions may be made around the point $V_r = \mu_{V_r}$, resulting in

$$S_x \approx V_x \left(\frac{1}{\mu_{V_r}} - \frac{(V_r - \mu_{V_r})}{\mu_{V_r}^2} + \frac{(V_r - \mu_{V_r})^2}{\mu_{V_r}^3} \right), \quad (\text{C.14})$$

$$S_y \approx V_y \left(\frac{1}{\mu_{V_r}} - \frac{(V_r - \mu_{V_r})}{\mu_{V_r}^2} + \frac{(V_r - \mu_{V_r})^2}{\mu_{V_r}^3} \right). \quad (\text{C.15})$$

Using these approximations, the integrals in equations C.8 to C.13 can be evaluated to calculate the expectation values, resulting in

$$\langle S_x \rangle = \frac{\mu_{V_x} (\mu_{V_r}^2 + \sigma_{V_r, V_r}) - \mu_{V_r} \sigma_{V_x, V_r}}{\mu_{V_r}}, \quad (\text{C.16})$$

$$\langle S_y \rangle = \frac{\mu_{V_y} (\mu_{V_r}^2 + \sigma_{V_r, V_r}) - \mu_{V_r} \sigma_{V_y, V_r}}{\mu_{V_r}}, \quad (\text{C.17})$$

$$\begin{aligned} \langle S_x S_y \rangle = & \frac{1}{\mu_{V_r}^6} \left(\mu_{V_r}^4 (\mu_{V_x} \mu_{V_y} + \sigma_{V_x, V_y}) - 2\mu_{V_r}^3 (\mu_{V_x} \sigma_{V_y, V_r} + \mu_{V_y} \sigma_{V_x, V_r}) \right. \\ & + 3\mu_{V_r}^2 (\sigma_{V_r, V_r} (\mu_{V_x} \mu_{V_y} + \sigma_{V_x, V_y}) + 2\sigma_{V_x, V_r} \sigma_{V_y, V_r}) \\ & - 6\mu_{V_r} \sigma_{V_r, V_r} (\mu_{V_x} \sigma_{V_y, V_r} + \mu_{V_y} \sigma_{V_x, V_r}) \\ & \left. + 3\sigma_{V_r, V_r}^2 (\mu_{V_x} \mu_{V_y} + \sigma_{V_x, V_y}) + 12\sigma_{V_r, V_r} \sigma_{V_x, V_r} \sigma_{V_y, V_r} \right). \end{aligned} \quad (\text{C.18})$$

Equations C.16 to C.18 can be combined to calculate the covariance of S_x and

S_y as [112]

$$\sigma_{S_x, S_y} = \langle S_x S_y \rangle - \langle S_x \rangle \langle S_y \rangle, \quad (\text{C.19})$$

$$\begin{aligned} &= \frac{1}{V_r^6} \left(V_r^4 \sigma_{V_x, V_y} - V_r^3 (V_x \sigma_{V_y, V_r} + V_y \sigma_{V_x, V_r}) \right. \\ &\quad + V_r^2 (\sigma_{V_r, V_r} (V_x V_y + 3\sigma_{V_x, V_y}) + 5\sigma_{V_x, V_r} \sigma_{V_y, V_r}) \\ &\quad - 5V_r \sigma_{V_r, V_r} (V_x \sigma_{V_y, V_r} + V_y \sigma_{V_x, V_r}) + \sigma_{V_r, V_r}^2 (2V_x V_y + 3\sigma_{V_x, V_y}) \\ &\quad \left. + 12\sigma_{V_r, V_r} \sigma_{V_x, V_r} \sigma_{V_y, V_r} \right), \end{aligned} \quad (\text{C.20})$$

where mean values μ_{V_x} , μ_{V_y} and μ_{V_r} have been replaced with the measured voltages V_x , V_y and V_r to allow equation C.20 to be used for uncertainty estimation in experimental work.

Appendix D

Derivation of the outputs of a phase-quadrature coating based homodyne interferometer

Three homodyne quadrature interferometer designs are presented in figure 4.2. In order to apply conservation of energy principles the optical intensity at the interferometer outputs must be found in terms of the intensity at the laser source. The amplitude of the electric field vector at the source, assuming unity refractive index for simplicity, may be written in terms of source intensity as

$$A_0 = \sqrt{\frac{2}{c\epsilon_0} I_0}. \quad (\text{D.1})$$

Making the plane wave assumption, the corresponding electric field vector is then given by

$$\mathbf{E}_0 = A_0 \cos(\mathbf{k} \cdot \mathbf{r} - \omega t) \hat{\mathbf{E}}_0, \quad (\text{D.2})$$

where $\mathbf{k}$ is the wavevector, r is a position vector, ω is the optical angular frequency and $\hat{\mathbf{E}}$ is a unit vector orthogonal to $\mathbf{k}$ describing the polarisation direction, and in general may be a function of t . Working in a coordinate system where the beam always propagates along the x axis, the electric field vector may be rewritten as

$$\mathbf{E}_0 = A_0 \cos(kx - \omega t) \hat{\mathbf{E}}_0, \quad (\text{D.3})$$

where k is the wavenumber. Taking the phase-quadrature coating interferometer schematic (figure 4.2 **A**), two beams, one from the right hand retro-reflector in figure 4.2 (beam 1) and one from the upper retro-reflector (beam 2), are superposed at the outputs I_a and I_b and the corresponding electric field vectors are given, neglecting phase changes upon reflection and transmission, for a Michelson beam splitter with

intensity reflection and transmission coefficients $\Re$ and $\Im$, by

$$\mathbf{E}_{a,1} = \sqrt{\frac{2\Re\Im}{c\epsilon_0}} I_0 \cos(kx_1 - \omega t) \hat{\mathbf{E}}_0, \quad (\text{D.4})$$

$$\mathbf{E}_{a,2} = \sqrt{\frac{2\Re\Im}{c\epsilon_0}} I_0 \cos(kx_2 - \omega t) \hat{\mathbf{E}}_0, \quad (\text{D.5})$$

$$\mathbf{E}_{b,1} = \Im \sqrt{\frac{2}{c\epsilon_0}} I_0 \cos(kx_1 - \omega t) \hat{\mathbf{E}}_0, \quad (\text{D.6})$$

$$\mathbf{E}_{b,2} = \Re \sqrt{\frac{2}{c\epsilon_0}} I_0 \cos(kx_2 - \omega t) \hat{\mathbf{E}}_0, \quad (\text{D.7})$$

$$(\text{D.8})$$

where x_1 and x_2 are the path lengths of the first and second beam paths respectively. Electric field vectors at detectors a and b are therefore given by

$$\mathbf{E}_a = \mathbf{E}_{a,1} + \mathbf{E}_{a,2}, \quad (\text{D.9})$$

$$= \sqrt{\frac{2\Re\Im}{c\epsilon_0}} I_0 (\cos(kx_1 - \omega t) + \cos(kx_2 - \omega t)) \hat{\mathbf{E}}_0, \quad (\text{D.10})$$

$$\mathbf{E}_b = \mathbf{E}_{b,1} + \mathbf{E}_{b,2}, \quad (\text{D.11})$$

$$= \sqrt{\frac{2}{c\epsilon_0}} I_0 (\Im \cos(kx_1 - \omega t) + \Re \cos(kx_2 - \omega t)) \hat{\mathbf{E}}_0. \quad (\text{D.12})$$

The corresponding Poynting vector at each output is then given by

$$\mathbf{S}_a = \frac{1}{\mu_0} \mathbf{E}_a \times \mathbf{B}_a, \quad (\text{D.13})$$

$$= \frac{1}{c\mu_0} |\mathbf{E}_a|^2 \hat{\mathbf{x}}, \quad (\text{D.14})$$

$$= 2\Re\Im I_0 (\cos^2(kx_1 - \omega t) + \cos^2(kx_2 - \omega t) + \cos(k(x_1 - x_2)) + \cos(k(x_1 + x_2) - 2\omega t)) \hat{\mathbf{x}}, \quad (\text{D.15})$$

$$\mathbf{S}_b = \frac{1}{\mu_0} \mathbf{E}_b \times \mathbf{B}_b, \quad (\text{D.16})$$

$$= \frac{1}{c\mu_0} |\mathbf{E}_b|^2 \hat{\mathbf{x}}, \quad (\text{D.17})$$

$$= 2I_0 (\Im^2 \cos^2(kx_1 - \omega t) + \Re^2 \cos^2(kx_2 - \omega t) + \Re\Im (\cos(k(x_1 - x_2)) + \cos(k(x_1 + x_2) - 2\omega t))) \hat{\mathbf{x}}, \quad (\text{D.18})$$

where in the beam following coordinate system the Poynting vector lies along the respective x axis of the coordinate system of each beam for both outputs, and the trigonometric identity

$$\cos(x) \cos(y) = \frac{1}{2} (\cos(x - y) + \cos(x + y)) \quad (\text{D.19})$$

has been applied. The detected, time averaged, optical intensity may then be found from the magnitude of the Poynting vector by integrating over one full temporal period and normalising as

$$I_a = \frac{1}{2\pi} \int_0^{2\pi} |\mathbf{S}_a| d(\omega t), \quad (\text{D.20})$$

$$= 2\Re\Im I_0 (1 + \cos(k(x_1 - x_2))), \quad (\text{D.21})$$

$$I_b = \frac{1}{2\pi} \int_0^{2\pi} |\mathbf{S}_b| d(\omega t), \quad (\text{D.22})$$

$$= I_0 (\Re^2 + \Im^2 + 2\Re\Im \cos(k(x_1 - x_2))). \quad (\text{D.23})$$

Phase differences between the outputs have up to this point been neglected. For the phase-quadrature beamsplitter based interferometer design, phase changes take place upon reflection and transmission from the phase-quadrature beamsplitter. As the interfering beams at output I_a have each reflected and transmitted from the coating once, no relative phase change will take place, however at output I_b a phase difference between the interfering beams equal to twice the sum of the phase change upon reflection and transmission is present. Introducing this phase difference as δ results in

$$I_a = 2\Re\Im I_0 (1 + \cos(\varphi)), \quad (\text{D.24})$$

$$I_b = I_0 (\Re^2 + \Im^2 + 2\Re\Im \cos(\varphi + \delta)), \quad (\text{D.25})$$

where the interferometric phase

$$\varphi = k(x_1 - x_2). \quad (\text{D.26})$$

Appendix E

Outline of the thin film transfer matrix model

A transfer matrix modelling methodology based upon that of Yeh, [130] with the required adjustments for absorbing materials given explicitly, will be outlined here, with the aim of constructing a transfer matrix that links the electric field on one side of a multi-layer thin film stack to the other.

The fundamental assumption of the transfer matrix methodology is that the electric field at any point in the multi-layer structure may be described as the sum of forward and backward propagating plane waves at that point, where the backward propagating wave is a result of back reflections of the forward propagating wave. Considering a multi-layer structure consisting of optically isotropic materials, as shown in figure 4.3, that is homogeneous in the x - y plane, with all material interfaces therefore falling in this plane, a coordinate system may be chosen such that the wavevector, and therefore the direction of propagation, falls in the x - z plane, without loss of generality. Applying these conditions, the electric field vector may be written as

$$\mathbf{E} = A_{\mathbf{F}} \cos(k_x x + k_z z - \omega t) + A_{\mathbf{B}} \cos(k_x x - k_z z - \omega t), \quad (\text{E.1})$$

where $A_{\mathbf{F}}$ and $A_{\mathbf{B}}$ are the amplitudes of the forward and backward propagating fields respectively, k_x and k_z are the x and z components of the wavevector and ω is the angular optical frequency. As the backward propagating wave is a result of reflections of the forward propagating wave from material interfaces lying in the x - y plane, the z component of the wavevector of the backward propagating wave has been reversed. Rewriting equation E.1 in complex exponential form in order to allow phase changes to be applied with linear operations for computational simplicity yields

$$\mathbf{E} = \frac{1}{2} A_{\mathbf{F}} \left(e^{i(k_x x + k_z z - \omega t)} + e^{-i(k_x x + k_z z - \omega t)} \right) + \frac{1}{2} A_{\mathbf{B}} \left(e^{i(k_x x - k_z z - \omega t)} + e^{-i(k_x x - k_z z - \omega t)} \right), \quad (\text{E.2})$$

which may alternatively be written

$$\tilde{\mathbf{E}} = \tilde{\mathbf{A}}_F \left(e^{i(k_x x + k_z z - \omega t)} \right) + \tilde{\mathbf{A}}_B \left(e^{i(k_x x - k_z z - \omega t)} \right), \quad (\text{E.3})$$

where $\tilde{\mathbf{E}}$ and $\tilde{\mathbf{A}}$ represent complex valued electric field and electric field amplitude vectors. The physical electric field is recovered by taking the real part of equation E.3

$$\mathbf{E} = \text{Re} \left(\tilde{\mathbf{E}} \right). \quad (\text{E.4})$$

Finally the electric field may be decomposed into S and P polarised components, projecting $\tilde{\mathbf{E}}$ onto the $\hat{\mathbf{y}}$ and $\hat{\mathbf{x}}$ directions respectively

$$\tilde{E}^{(S)} = \tilde{A}_F^{(S)} e^{i(k_x x + k_z z - \omega t)} + \tilde{A}_B^{(S)} e^{i(k_x x - k_z z - \omega t)}, \quad (\text{E.5})$$

$$\tilde{E}^{(P)} = \tilde{A}_F^{(P)} e^{i(k_x x + k_z z - \omega t)} + \tilde{A}_B^{(P)} e^{i(k_x x - k_z z - \omega t)}. \quad (\text{E.6})$$

Returning to figure 4.3, it can be seen that in order for the electric field to be continuous over the layer boundaries, k_x must remain constant between layers, and therefore must remain constant throughout the multi-layer structure. Due to the previously stated constraint that the beam must propagate in the x - z plane

$$k_y = 0, \quad (\text{E.7})$$

permitting the z component of the wavevector in layer m to be calculated as

$$k_{z,m} = \sqrt{\tilde{k}_m^2 - k_x^2}, \quad (\text{E.8})$$

where the complex wavevector magnitude $\tilde{k}_m$ is given in terms of the complex refractive index $\tilde{n}_m$ by

$$\tilde{k}_m = \frac{2\pi}{\lambda_0} \tilde{n}_m. \quad (\text{E.9})$$

Applying the complex valued k_z to equations E.5 and E.6 results in

$$\tilde{E}^{(S)} = \tilde{A}_F^{(S)} e^{i(k_x x + \text{Re}(k_z)z - \omega t)} e^{-\text{Im}(k_z)z} + \tilde{A}_B^{(S)} e^{i(k_x x - \text{Re}(k_z)z - \omega t)} e^{\text{Im}(k_z)z}, \quad (\text{E.10})$$

$$\tilde{E}^{(P)} = \tilde{A}_F^{(P)} e^{i(k_x x + \text{Re}(k_z)z - \omega t)} e^{-\text{Im}(k_z)z} + \tilde{A}_B^{(P)} e^{i(k_x x - \text{Re}(k_z)z - \omega t)} e^{\text{Im}(k_z)z}, \quad (\text{E.11})$$

introducing an exponential decay with increasing z for the forward propagating field component for positive values of the imaginary component of the refractive index, and vice versa for the backward propagating field component. The complex refractive index is therefore defined as

$$\tilde{n} = n + i\kappa, \quad (\text{E.12})$$

where n is the real valued refractive index, and κ is a loss or gain coefficient, positive for lossy (metallic) optical media.

The required wavevector components may now be calculated throughout the multi-layer structure, however before the electric field components can be evaluated the behaviour at the layer boundaries must be described. For brevity of notation,

equations E.5 and E.6 may be rewritten in terms of forward and backward propagating electric field components as functions of z

$$\tilde{E}^{(S)} = \tilde{E}_F^{(S)}(z) + \tilde{E}_B^{(S)}(z), \quad (\text{E.13})$$

$$\tilde{E}^{(P)} = \tilde{E}_F^{(P)}(z) + \tilde{E}_B^{(P)}(z), \quad (\text{E.14})$$

making the implicit assumption that the temporal component may be considered constant throughout the multi-layer structure. This assumption restricts the transfer matrix methodology described here to steady state solutions on optical timescales. Considering these forward and backward propagating field components on the low z (z_m^-) and high z (z_m^+) sides of the layer boundary between layers $m-1$ and m , a system of simultaneous equations arises, given by

$$\tilde{E}_F^{(S/P)}(z_m^+) = \mathbf{t}_+^{(S/P)} \tilde{E}_F^{(S/P)}(z_m^-) + \mathbf{r}_-^{(S/P)} \tilde{E}_B^{(S/P)}(z_m^+), \quad (\text{E.15})$$

$$\tilde{E}_B^{(S/P)}(z_m^-) = \mathbf{r}_+^{(S/P)} \tilde{E}_F^{(S/P)}(z_m^-) + \mathbf{t}_-^{(S/P)} \tilde{E}_B^{(S/P)}(z_m^+), \quad (\text{E.16})$$

where the superscript (S/P) represents S or P polarisation and $\mathbf{r}_\pm^{(S/P)}$ and $\mathbf{t}_\pm^{(S/P)}$ are the amplitude reflectivity and transmission coefficients in the positive and negative z directions respectively for each polarisation.

In order to avoid equations involving complex angles, the amplitude reflectivity and transmission coefficients may be rewritten in terms of wavevector components. The classical form of the amplitude (Fresnel) coefficients is

$$\mathbf{r}_+^{(S)} = \frac{\tilde{n}_{m-1} \cos(\theta_{m-1}) - \tilde{n}_m \cos(\theta_m)}{\tilde{n}_{m-1} \cos(\theta_{m-1}) + \tilde{n}_m \cos(\theta_m)}, \quad (\text{E.17})$$

$$\mathbf{t}_+^{(S)} = \frac{2\tilde{n}_{m-1} \cos(\theta_{m-1})}{\tilde{n}_{m-1} \cos(\theta_{m-1}) + \tilde{n}_m \cos(\theta_m)}, \quad (\text{E.18})$$

$$\mathbf{r}_+^{(P)} = \frac{\tilde{n}_m \cos(\theta_{m-1}) - \tilde{n}_{m-1} \cos(\theta_m)}{\tilde{n}_m \cos(\theta_{m-1}) + \tilde{n}_{m-1} \cos(\theta_m)}, \quad (\text{E.19})$$

$$\mathbf{t}_+^{(P)} = \frac{2\tilde{n}_{m-1} \cos(\theta_{m-1})}{\tilde{n}_m \cos(\theta_{m-1}) + \tilde{n}_{m-1} \cos(\theta_m)}, \quad (\text{E.20})$$

where θ_{m-1} and θ_m are in non-absorbing materials the angles of incidence and ex-
 idence respectively on the low z and high z sides of layer boundary. For plane
 waves propagating in absorbing media the planes of constant phase and planes of
 constant amplitude are no long aligned, and therefore the direction of propagation
 and resulting angle of incidence at a layer boundary are ill defined. Returning to
 non-absorbing materials, the cosine of the angle of propagation in layer m may be
 written in terms of wavevector components as

$$\cos(\theta_m) = \frac{k_{z,m} \lambda_0}{2\pi \tilde{n}_m} \quad (\text{E.21})$$

by application of trigonometry. Substituting equation E.21 into equations E.17 to

E.20 results in the wavevector forms of the forward amplitude reflectivity coefficients,

$$\mathbf{r}_+^{(S)} = \frac{k_{z,m-1} - k_{z,m}}{k_{z,m-1} + k_{z,m}}, \quad (\text{E.22})$$

$$\mathbf{r}_+^{(P)} = \frac{\tilde{n}_{m-1}^2 k_{z,m} - \tilde{n}_m^2 k_{z,m-1}}{\tilde{n}_{m-1}^2 k_{z,m} + \tilde{n}_m^2 k_{z,m-1}}, \quad (\text{E.23})$$

and by observing that

$$\mathbf{t}_+^{(S)} = 1 + \mathbf{r}_+^{(S)}, \quad (\text{E.24})$$

$$\mathbf{t}_+^{(P)} = \frac{\tilde{n}_{m-1}}{\tilde{n}_m} \left(1 + \mathbf{r}_+^{(P)} \right), \quad (\text{E.25})$$

the required forward amplitude coefficients can now be calculated in terms of wavevector components and complex refractive indices. The backward coefficients are then given in terms of the forward coefficients by

$$\mathbf{r}_-^{(S)} = \frac{k_{z,m} - k_{z,m-1}}{k_{z,m-1} + k_{z,m}}, \quad (\text{E.26})$$

$$= -\mathbf{r}_+^{(S)}, \quad (\text{E.27})$$

$$\mathbf{r}_-^{(P)} = \frac{\tilde{n}_m^2 k_{z,m-1} - \tilde{n}_{m-1}^2 k_{z,m}}{\tilde{n}_{m-1}^2 k_{z,m} + \tilde{n}_m^2 k_{z,m-1}}, \quad (\text{E.28})$$

$$= -\mathbf{r}_+^{(P)}, \quad (\text{E.29})$$

$$\mathbf{t}_-^{(S)} = 2 - \mathbf{t}_+^{(S)}, \quad (\text{E.30})$$

$$\mathbf{t}_-^{(P)} = 2 \frac{n_{m-1}}{n_m} - \mathbf{t}_+^{(P)}. \quad (\text{E.31})$$

Substituting these equations back into equations E.15 and E.16 and rearranging, expressions for the low z field components in terms of the high z field components and forward amplitude coefficients may be found, resulting in

$$\tilde{E}_F^{(S/P)}(z_m^-) = \frac{1}{\mathbf{t}_+^{(S/P)}} \tilde{E}_F^{(S/P)}(z_m^+) + \frac{\mathbf{r}_+^{(S/P)}}{\mathbf{t}_+^{(S/P)}} \tilde{E}_B^{(S/P)}(z_m^+), \quad (\text{E.32})$$

$$\tilde{E}_B^{(S/P)}(z_m^-) = \frac{\mathbf{r}_+^{(S/P)}}{\mathbf{t}_+^{(S/P)}} \tilde{E}_F^{(S/P)}(z_m^+) + \frac{1}{\mathbf{t}_+^{(S/P)}} \tilde{E}_B^{(S/P)}(z_m^+), \quad (\text{E.33})$$

or alternatively in matrix form

$$\begin{pmatrix} \tilde{E}_F^{(S)}(z_m^-) \\ \tilde{E}_B^{(S)}(z_m^-) \\ \tilde{E}_F^{(P)}(z_m^-) \\ \tilde{E}_B^{(P)}(z_m^-) \end{pmatrix} = \begin{pmatrix} \frac{1}{t_+^{(S)}} & \frac{r_+^{(S)}}{t_+^{(S)}} & 0 & 0 \\ \frac{r_+^{(S)}}{t_+^{(S)}} & \frac{1}{t_+^{(S)}} & 0 & 0 \\ 0 & 0 & \frac{1}{t_+^{(P)}} & \frac{r_+^{(P)}}{t_+^{(P)}} \\ 0 & 0 & \frac{r_+^{(P)}}{t_+^{(P)}} & \frac{1}{t_+^{(P)}} \end{pmatrix} \begin{pmatrix} \tilde{E}_F^{(S)}(z_m^+) \\ \tilde{E}_B^{(S)}(z_m^+) \\ \tilde{E}_F^{(P)}(z_m^+) \\ \tilde{E}_B^{(P)}(z_m^+) \end{pmatrix}, \quad (\text{E.34})$$

$$\begin{pmatrix} \tilde{E}_F^{(S)}(z_m^-) \\ \tilde{E}_B^{(S)}(z_m^-) \\ \tilde{E}_F^{(P)}(z_m^-) \\ \tilde{E}_B^{(P)}(z_m^-) \end{pmatrix} = \mathbf{T}_{m-1,m} \begin{pmatrix} \tilde{E}_F^{(S)}(z_m^+) \\ \tilde{E}_B^{(S)}(z_m^+) \\ \tilde{E}_F^{(P)}(z_m^+) \\ \tilde{E}_B^{(P)}(z_m^+) \end{pmatrix}, \quad (\text{E.35})$$

where $\mathbf{T}_{m,m-1}$ is a transfer matrix from layer m to $m-1$.

In order to link the electric field components across a layer m , an expression for the field at z_{m-1}^+ in terms of z_m^- is required. This may be achieved with a propagation matrix applying a phase change given by the product of the z component of the wavevector and the layer thickness in the appropriate direction for the forward and backward propagating beams, resulting in

$$\begin{pmatrix} \tilde{E}_F^{(S)}(z_{m-1}^+) \\ \tilde{E}_B^{(S)}(z_{m-1}^+) \\ \tilde{E}_F^{(P)}(z_{m-1}^+) \\ \tilde{E}_B^{(P)}(z_{m-1}^+) \end{pmatrix} = \begin{pmatrix} e^{-ik_{z,1}d_m} & 0 & 0 & 0 \\ 0 & e^{ik_{z,1}d_m} & 0 & 0 \\ 0 & 0 & e^{-ik_{z,1}d_m} & 0 \\ 0 & 0 & 0 & e^{ik_{z,1}d_m} \end{pmatrix} \begin{pmatrix} \tilde{E}_F^{(S)}(z_m^-) \\ \tilde{E}_B^{(S)}(z_m^-) \\ \tilde{E}_F^{(P)}(z_m^-) \\ \tilde{E}_B^{(P)}(z_m^-) \end{pmatrix}, \quad (\text{E.36})$$

$$\begin{pmatrix} \tilde{E}_F^{(S)}(z_{m-1}^+) \\ \tilde{E}_B^{(S)}(z_{m-1}^+) \\ \tilde{E}_F^{(P)}(z_{m-1}^+) \\ \tilde{E}_B^{(P)}(z_{m-1}^+) \end{pmatrix} = \mathbf{P}_m \begin{pmatrix} \tilde{E}_F^{(S)}(z_m^-) \\ \tilde{E}_B^{(S)}(z_m^-) \\ \tilde{E}_F^{(P)}(z_m^-) \\ \tilde{E}_B^{(P)}(z_m^-) \end{pmatrix}, \quad (\text{E.37})$$

where $\mathbf{P}_m$ is a propagation matrix across layer m .

A global transfer matrix $\mathbf{T}$ from layer N to layer 0 may now be defined in terms

of the individual transfer and propagation matrices as

$$\begin{pmatrix} \tilde{E}_F^{(S)}(0) \\ \tilde{E}_B^{(S)}(0) \\ \tilde{E}_F^{(P)}(0) \\ \tilde{E}_B^{(P)}(0) \end{pmatrix} = \mathbf{T}_{0,1} \left(\prod_{m=1}^{N-1} \mathbf{P}_m \mathbf{T}_{m,m+1} \right) \begin{pmatrix} \tilde{E}_F^{(S)}(N) \\ \tilde{E}_B^{(S)}(N) \\ \tilde{E}_F^{(P)}(N) \\ \tilde{E}_B^{(P)}(N) \end{pmatrix}, \quad (\text{E.38})$$

$$= \mathbf{T} \begin{pmatrix} \tilde{E}_F^{(S)}(N) \\ \tilde{E}_B^{(S)}(N) \\ \tilde{E}_F^{(P)}(N) \\ \tilde{E}_B^{(P)}(N) \end{pmatrix}. \quad (\text{E.39})$$

References

- [1] The International System of Units. 8th ed. Paris, France: Bureau International des Poids et Mesures; 2006.
- [2] Schattenburg ML, Smith HI. The critical role of metrology in nanotechnology. In: Peckerar MC, Postek Jr MT, editors. Proc. SPIE, Nanostructure Science, Metrology, and Technology. vol. 4608. SPIE Press; 2002. p. 116-24.
- [3] Taniguchi N. Current status in, and Future Trends of, Ultraprecision Machining and Ultrafine Materials Processing. *Annals of the CIRP*. 1983;32(2):573-82.
- [4] Leach RK, Boyd R, Burke T, Danzebrink HU, Dirscherl K, Dziomba T, Gee M, Koenders L, Morazzani V, Pidduck A, Roy D, Unger WES, Yacoot A. The European nanometrology landscape. *Nanotechnology*. 2011;22(6):062001.
- [5] IEEE. International Roadmap for Devices and Systems 2020 (Metrology); 2020.
- [6] Yacoot A, Klapetek P, Valtr M, Grolich P, Dongmo H, Lazzerini GM, Bridges A. Design and performance of a test rig for evaluation of nanopositioning stages. *Measurement Science and Technology*. 2019;30(3):035002.
- [7] Mise en pratique for the definition of the metre in the SI, Appendix 2, SI Brochure. 9th ed. Consultative Committee for Length, Bureau International des Poids et Mesures; 2019.
- [8] Yacoot A, Bosse H, Dixon R. The lattice parameter of silicon: A secondary realisation of the metre. *Measurement Science and Technology*. 2020;31(12):121001.
- [9] Basile G, Bergamin A, Cavagnero G, Mana G. Measurement of the Silicon (220) lattice Spacing. *Physical Review Letters*. 1994;72(20):3133-6.
- [10] Massa E, Sasso CP, Mana G, Palmisano C. A more accurate measurement of the ^{28}Si lattice parameter. *Journal of Physical and Chemical Reference Data*. 2015;44(3):031208.
- [11] Pisani M, Yacoot A, Balling P, Bancone N, Birlikseven C, Çelik M, Flügge J, Hamid R, Köchert P, Kren P, Kuertgens U, Lassila A, Picotto GB, Şahin E, Seppä J, Tedaldi M, Weichert C. Comparison of the performance of the next generation of optical interferometers. *Metrologia*. 2012;49(4):455-67.

- [12] Yacoot A, Kuetgens U. Sub-atomic dimensional metrology: developments in the control of x-ray interferometers. *Measurement Science and Technology*. 2012;23(7):074003.
- [13] Hariharan P. Interferometers. In: Bass M, DeCusatis C, Enoch JM, Lakshminarayanan V, Li G, MacDonald C, Mahajan VN, Stryland EV, editors. *Handbook of Optics, Volume II: Design, Fabrication and Testing, Sources and Detectors, Radiometry and Photometry*. 3rd ed. New York, USA: McGraw-Hill Professional Publishing; 2009. p. 21.1-21.28.
- [14] Stone JA, Howard LP. Simple technique for observing periodic nonlinearities in Michelson interferometers. *Precision Engineering*. 1998;22(4):220-32.
- [15] Lawall J, Kessler E. Michelson interferometry with 10 pm accuracy. *Review of Scientific Instruments*. 2000;71(7):2669-76.
- [16] Cosijns SJAG. Displacement laser interferometry with sub-nanometer uncertainty. Technische Universiteit Eindhoven; 2004.
- [17] Joo KN, Clark E, Zhang Y, Ellis JD, Guzmán F. A compact high-precision periodic-error-free heterodyne interferometer. *Journal of the Optical Society of America A*. 2020;37(9):B11-8.
- [18] SIOS Meßtechnik GmbH, SP 5000 DI; 2019. Available at: <https://sios-de.com/products/length-measurement/differential-interferometer/> (Accessed: 16 September 2021).
- [19] Michelson AA, Morley EW. On the Relative Motion of the Earth and the Luminiferous Ether. *American Journal of Science*. 1887;34(203):333-45.
- [20] Born M, Wolf E. *Principles of Optics*. 6th ed. Oxford, UK: Pergamon Press; 1980.
- [21] Birch KP. Optical fringe subdivision with nanometric accuracy. *Precision Engineering*. 1990;12(4):195-8.
- [22] Itoh K. Analysis of the phase unwrapping algorithm. *Applied Optics*. 1982;21(14):2470.
- [23] Morokuma T. Interference Comparator for Routine Measurement of Length (II). *Oyo Buturi*. 1962;31(3):201-2206.
- [24] Peck ER, Obetz SW. Wavelength or length measurement by reversible fringe counting. *Journal of the Optical Society of America*. 1953;43(6):505-9.
- [25] Raine KW, Downs MJ. Beam-splitter coatings for producing phase quadrature interferometer outputs. *Optica Acta*. 1978;25(7):549-58.
- [26] Badami VG, de Groot PJ. Displacement Measuring Interferometry. In: Harding K, editor. *Handbook of Optical Dimensional Metrology*. Boca Raton, USA: CRC Press; 2016. p. 157-238.

- [27] de Lang H, Bouwhuis G. Displacement measurement with a laser interferometer. *Phillips Technical Review*. 1969;30(6):160-5.
- [28] Weichert C, Köchert P, Köning R, Flügge J, Andreas B, Kuetgens U, Yacoot A. A heterodyne interferometer with periodic nonlinearities smaller than ± 10 pm. *Measurement Science and Technology*. 2012;23(9):094005.
- [29] Wang C, Ellis JD. Dynamic Doppler Frequency Shift Errors: Measurement, Characterization, and Compensation. *IEEE Transactions on Instrumentation and Measurement*. 2015;64(7):1994-2004.
- [30] Horowitz P, Hill W. *The Art of Electronics*. 3rd ed. New York, USA: Cambridge University Press; 2015.
- [31] Zhang S, Tan Y, Li Y. Orthogonally polarized dual frequency lasers and applications in self-sensing metrology. *Measurement Science and Technology*. 2010;21(5):054016.
- [32] Jackson DA, Kersey AD, Corke M, Jones JDC. Pseudoheterodyne detection scheme for optical interferometers. *Electronics Letters*. 1982;18(25-26):1081-3.
- [33] Kissinger T, Charrett TOH, Tatam RP. Range-resolved interferometric signal processing using sinusoidal optical frequency modulation. *Optics Express*. 2015;23(7):9415-31.
- [34] Gerberding O. Deep frequency modulation interferometry. *Optics Express*. 2015;23(11):14753-62.
- [35] Zheng J. Analysis of optical frequency-modulated continuous-wave interference. *Applied Optics*. 2004;43(21):4189-98.
- [36] Zheng J. Optical frequency-modulated continuous-wave interferometers. *Applied Optics*. 2006;45(12):2723-30.
- [37] Wiseman KB, Kissinger T, Tatam RP. Two-dimensional remote interferometric stage encoder through a single access port using range-resolved interferometry. In: *Optical Measurement Systems for Industrial Inspection XI: SPIE Optical Metrology*. vol. 11056. Munich, Germany: SPIE Press; 2019. p. 110560C.
- [38] Howard L, Stone J, Fu J. Real-time displacement measurements with a Fabry-Perot cavity and a diode laser. *Precision Engineering*. 2001;25(4):321-35.
- [39] Bitou Y. High-accuracy displacement metrology and control using a dual Fabry-Perot cavity with an optical frequency comb generator. *Precision Engineering*. 2009;33(2):187-93.
- [40] Durand M, Lawall J, Wang Y. High-accuracy Fabry-Perot displacement interferometry using fiber lasers. *Measurement Science and Technology*. 2011;22(9):094025.

- [41] Durand M, Lawall J, Wang Y. Fabry-Perot displacement interferometry for next-generation calculable capacitor. *IEEE Transactions on Instrumentation and Measurement*. 2011;60(7):2673-7.
- [42] Çelik M, Hamid R, Kuetgens U, Yacoot A. Picometre displacement measurements using a differential Fabry-Perot optical interferometer and an x-ray interferometer. *Measurement Science and Technology*. 2012;23(8):085901.
- [43] Martincek I, Kacik D, Tarjanyi N, Schuster K. Determination of sub-nanometer distance from reflective surface by Fabry-Perot interferometer. *Applied Surface Science*. 2018;461:139-42.
- [44] Drever RWP, Hall JL, Kowalski FV, Hough J, Ford GM, Munley AJ, Ward H. Laser phase and frequency stabilization using an optical resonator. *Applied Physics B*. 1983;31(2):97-105.
- [45] Pisani M. Multiple reflection Michelson interferometer with picometer resolution. *Optics Express*. 2008;16(26):21558-63.
- [46] Pisani M. A homodyne Michelson interferometer with sub-picometer resolution. *Measurement Science and Technology*. 2009;20(8):084008.
- [47] Pisani M, Giugni A. A portable picometer reference actuator with 100 μ m range , picometer resolution , subnanometer accuracy and submicroradian tip-tilt error for the characterization of measuring instruments at the nanoscale. *Metrologia*. 2018;55(4):541-57.
- [48] Křen P, Balling P. Common path two-wavelength homodyne counting interferometer development. *Measurement Science and Technology*. 2009;20(8):084009.
- [49] Stone JA, Phillips SD. Corrections for wavelength variations in precision interferometric displacement measurements. *Journal of Research of the National Institute of Standards and Technology*. 1996;101(5):671-4.
- [50] Hori Y, Gonda S, Bitou Y, Watanabe A, Nakamura K. Periodic error evaluation system for linear encoders using a homodyne laser interferometer with 10 picometer uncertainty. *Precision Engineering*. 2018;51(July 2017):388-92.
- [51] Řeřucha Š, Holá M, Šarbort M, Hrabina J, Oulehla J, Číp O, Lazar J. Compact differential plane interferometer with in-axis mirror tilt detection. *Optics and Lasers in Engineering*. 2021;141:106568.
- [52] Forbes C, Evans M, Hastings N, Peacock B. *Statistical Distributions*. 4th ed. John Wiley & Sons, Inc., Hoboken, New Jersey; 2011.
- [53] Řeřucha S, Yacoot A, Pham TM, Cizek M, Hucl V, Lazar J, Číp O. Laser source for dimensional metrology: investigation of an iodine stabilized system based on narrow linewidth 633 nm DBR diode. *Measurement Science and Technology*. 2017;28(4):045204.

- [54] Yacoot A, Downs MJ. The use of x-ray interferometry to investigate the linearity of the NPL Differential Plane Mirror Optical Interferometer. *Measurement Science and Technology*. 2000;11(8):1126-30.
- [55] Haycocks J, Jackson K. Traceable calibration of transfer standards for scanning probe microscopy. *Precision Engineering*. 2005;29(2):168-75.
- [56] Leach RK, Giusca CL, Naoi K. Development and characterization of a new instrument for the traceable measurement of areal surface texture. *Measurement Science and Technology*. 2009;20(12):125102.
- [57] Rosenbluth AE, Bobroff N. Optical sources of non- linearity in heterodyne interferometers. *Precision Engineering*. 1990;12(1):7-11.
- [58] Bobroff N. Recent advances in displacement measuring interferometry. *Measurement Science and Technology*. 1993;4(9):907-26.
- [59] Wu CM, Deslattes RD. Analytical modeling of the periodic nonlinearity in heterodyne interferometry. *Applied Optics*. 1998;37(28):6696-700.
- [60] Wu CM. Periodic nonlinearity resulting from ghost reflections in heterodyne interferometry. *Optics Communications*. 2002;215(1-3):17-23.
- [61] Fu H, Wang Y, Hu P, Tan J, Fan Z. Nonlinear Errors Resulting from Ghost Reflection and Its Coupling with Optical Mixing in Heterodyne Laser Interferometers. *Sensors*. 2018;18(3):758.
- [62] Tanaka M, Yamagami T, Nakayama K. Linear Interpolation of Periodic Error in a Heterodyne Laser Interferometer at Subnanometer Levels. *IEEE Transactions on Instrumentation and Measurement*. 1989;38(2):552-4.
- [63] Wu Cm, Lawall J, Deslattes RD. Heterodyne interferometer with subatomic periodic nonlinearity. *Applied Optics*. 1999;38(19):4089-94.
- [64] Joo KN, Ellis JD, Buice ES, Spronck JW, Schmidt RHM. High resolution heterodyne interferometer without detectable periodic nonlinearity. *Optics Express*. 2010;18(2):1159-65.
- [65] Joo KN, Ellis JD, Spronck JW, van Kan PJM, Munnig Schmidt RH. Simple heterodyne laser interferometer with subnanometer periodic errors. *Optics Letters*. 2009;34(3):386-8.
- [66] Ellis JD, Meskers AJH, Spronck JW, Schmidt RHM. Fiber-coupled displacement interferometry without periodic nonlinearity. *Optics Letters*. 2011;36(18):3584-6.
- [67] McRae TG, Hsu MTL, Freund CH, Shaddock DA, Herrmann J, Gray MB. Linearization and minimization of cyclic error with heterodyne laser interferometry. *Optics Letters*. 2012;37(13):2448.

- [68] Seppä J, Korpelainen V, Merimaa M, Picotto GB, Lassila A. A method for linearization of a laser interferometer down to the picometre level with a capacitive sensor. *Measurement Science and Technology*. 2011;22(9):094027.
- [69] Bridges A, Yacoot A, Kissinger T, Humphreys D, Tatam RP. Correction of periodic displacement non-linearities by two-wavelength interferometry. *Measurement Science and Technology*. 2021;32(12):125202.
- [70] Basile G, Bergamin A, Cavagnero G, Mana G. Phase Modulation in High-resolution Optical Interferometry. *Metrologia*. 1991;28(6):455-61.
- [71] Basile G, Becker P, Bergamin A, Cavagnero G, Franks A, Jackson K, Kuetgens U, Mana G, Palmer EW, Robbie CJ, Stednan M, Stümpel J, Yacoot A, Zosi G. Combined optical and X-ray interferometry for high-precision dimensional metrology. *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences*. 2000;456(1995):701-29.
- [72] Thurner K, Quacquarelli FP, Braun PF, Savio CD, Karrai K. Fiber-based distance sensing interferometry. *Applied Optics*. 2015;54(10):3051-63.
- [73] Danzebrink HU, Koenders L, Wilkening G, Yacoot A, Kunzmann H. Advances in scanning force microscopy for dimensional metrology. *Annals of the CIRP*. 2006;55(2):841-78.
- [74] Yacoot A, Koenders L, Wolff H. An atomic force microscope for the study of the effects of tip-sample interactions on dimensional metrology. *Measurement Science and Technology*. 2007;18(2):350-9.
- [75] Yacoot A, Koenders L. Recent developments in dimensional nanometrology using AFMs. *Measurement Science and Technology*. 2011;22(12):122001.
- [76] Schibli TR, Minoshima K, Bitou Y, Hong FL, Inaba H, Onae A, Matsumoto H. Displacement metrology with sub-pm resolution in air based on a fs-comb wavelength synthesizer. *Optics Express*. 2006;14(13):5984-93.
- [77] Schibli TR, Minoshima K, Hong FL, Inaba H, Bitou Y, Onae A, Matsumoto H. An arbitrary optical single-frequency generator based on a femtosecond frequency comb. *Optics Letters*. 2005;30(17):2323-5.
- [78] Lawall JR. Fabry – Perot metrology for displacements up to 50 mm. *Journal of the Optical Society of America A*. 2005;22(12):2786-98.
- [79] Durand M, Wang Y, Lawall J. Accurate Gouy phase measurement in an astigmatic optical cavity. *Applied Physics B: Lasers and Optics*. 2012;108(4):749-53.
- [80] Wang Y, Guzmán Cervantes F, Stambaugh C, Šmíd R, Zúñiga Calvo H, Koffman A, Pratt JR, Lawall J. Alignment and testing of the NIST Calculable Capacitor. In: 29th Conference on Precision Electromagnetic Measurements (CPEM 2014). Rio de Janeiro, Brazil: IEEE; 2014. p. 476-7.

- [81] Stambaugh C, Durand M, Lawall J. Displacement measurements in air for the calculable capacitor. In: CPEM Digest (Conference on Precision Electromagnetic Measurements). IEEE; 2014. p. 112-3.
- [82] De Freitas JM, Player MA. Importance of rotational beam alignment in the generation of second harmonic errors in laser heterodyne interferometry. *Measurement Science and Technology*. 1993;4(10):1173-6.
- [83] Wu Cm, Su Cs, Peng GS. Correction of nonlinearity in one-frequency optical interferometry. *Measurement Science and Technology*. 1996;7(4):520-4.
- [84] Hu P, Wang Y, Fu H, Zhu J, Tan J. Nonlinearity error in homodyne interferometer caused by multi-order Doppler frequency shift ghost reflections. *Optics Express*. 2017;25(4):3605-12.
- [85] Bridges A, Yacoot A, Kissinger T, Tatam RP. Polarization-sensitive transfer matrix modeling for displacement measuring interferometry. *Applied Optics*. 2020;59(25):7694-704.
- [86] Bobroff N. Residual errors in laser interferometry from air turbulence and nonlinearity. *Applied Optics*. 1987;26(13):2676.
- [87] Heydemann PLM. Determination and correction of quadrature fringe measurement errors in interferometers. *Applied Optics*. 1981;20(19):3382-4.
- [88] Meiners-Hagen K, Abou-Zeid A. Refractive index determination in length measurement by two-colour interferometry. *Measurement Science and Technology*. 2008;19(8):084004.
- [89] JCGM 100:2008, Evaluation of measurement data — Guide to the expression of uncertainty in measurement. Joint Committee for Guides in Metrology, Bureau International des Poids et Mesures; 2008.
- [90] ISO 1: Geometrical product specifications (GPS) – Standard reference temperature for the specification of geometrical and dimensional properties. 3rd ed.; 2016.
- [91] Monchalin JP, Kelly MJ, Thomas JE, Kurnit NA, Szöke A, Zernike F, Lee PH, Javan A. Accurate laser wavelength measurement with a precision two-beam scanning Michelson interferometer. *Applied Optics*. 1981;20(5):736-57.
- [92] Edlén B. The Dispersion of Standard Air. *Journal of the Optical Society of America*. 1953;43(5):339-44.
- [93] Edlén B. The refractive index of air. *Metrologia*. 1966;2(2):71-80.
- [94] Birch KP, Downs MJ. An Updated Edlen Equation for the Refractive Index of Air. *Metrologia*. 1993;30(3):155-62.
- [95] Birch KP, Downs MJ. Correction to the Updated Edlén Equation for the Refractive Index of Air. *Metrologia*. 1994;31(4):315-6.

- [96] Ciddor PE. Refractive index of air: new equations for the visible and near infrared. *Applied Optics*. 1996;35(9):1566.
- [97] Bonsch G, Potulski E. Measurement of the refractive index of air and comparison with modified Edlen's formulae. *Metrologia*. 2003;35(2):133-9.
- [98] Schellekens P, Wilkening G, Reinboth F, Downs MJ, Birch KP, Spronck J. Measurements of the Refractive Index of Air Using Interference Refractometers. *Metrologia*. 1986;22(4):279-87.
- [99] Niebauer TM, Faller JE, Godwin HM, Hall JL, Barger RL. Frequency stability measurements on polarization-stabilized He-Ne lasers. *Applied Optics*. 1988;27(7):1285-9.
- [100] Bergamin A, Cavagnero G, Cordiali L, Mana G. A Fourier optics model of two-beam scanning laser interferometers. *European Physical Journal D*. 1999;5(3):433-40.
- [101] van Westrum D, Niebauer TM. The Diffraction Correction for Interferometric Absolute Gravimeters. *Metrologia*. 2003;40(5):258-63.
- [102] Cavagnero G, Mana G, Massa E. Aberration effects in two-beam laser interferometers. *Journal of the Optical Society of America A*. 2006;23(8):1951-9.
- [103] Robertsson L. On the diffraction correction in absolute gravimetry. *Metrologia*. 2007;44(1):35-9.
- [104] Wanner G, Heinzl G. Analytical description of interference between two misaligned and mismatched complete Gaussian beams. *Applied Optics*. 2014;53(14):3043-8.
- [105] Mana G, Massa E, Sasso CP, Andreas B, Kuetgens U. A new analysis for diffraction correction in optical interferometry. *Metrologia*. 2017;54(4):559-65.
- [106] Mana G, Massa E, Sasso CP. Wavefront errors in a two-beam interferometer. *Metrologia*. 2018;55(4):535-40.
- [107] Mana G, Sasso CP. Corrections of the travelling-fringe period for the interference of aberrated beams. *Metrologia*. 2019;56(5):055004.
- [108] Köning R, Flügge J, Bosse H. A method for the in situ determination of Abbe errors and their correction. *Measurement Science and Technology*. 2007;18(2):476-81.
- [109] Schödel R, Bönsch G. Highest-accuracy interferometer alignment by retroreflection scanning. *Applied Optics*. 2004;43(31):5738-43.
- [110] Köning R, Wimmer G, Witkovský V. The statistical uncertainty of the Heydemann correction: A practical limit of optical quadrature homodyne interferometry. *Measurement Science and Technology*. 2015;26(8):084004.

- [111] Köning R, Wimmer G, Witkovský V. Ellipse fitting by nonlinear constraints to demodulate quadrature homodyne interferometer signals and to determine the statistical uncertainty of the interferometric phase. *Measurement Science and Technology*. 2014;25(11):115001.
- [112] Everitt BS, Skrondal A. *The Cambridge dictionary of statistics*. 4th ed. Cambridge, United Kingdom: Cambridge University Press; 2010.
- [113] Schottky W. On spontaneous current fluctuations in various electrical conductors (Translated by Burkhardt, M and Yen, A). *Journal of Micro/Nanolithography, MEMS, and MOEMS*. 2018;17(4):041001.
- [114] Fujimoto H, Mana G, Nakayama K. Light bounces in two-beam scanning laser interferometers. *Japanese Journal of Applied Physics, Part 1: Regular Papers and Short Notes and Review Papers*. 2000;39(5 A):2870-5.
- [115] Cavagnero G, Mana G, Massa E. Effect of recycled light in two-beam interferometry. *Review of Scientific Instruments*. 2005;76(5):053106.
- [116] Caves CM. Quantum-mechanical noise in an interferometer. *Physical Review D*. 1981;28(8):1693-708.
- [117] Johnson JB. Thermal agitation of electricity in conductors. *Physical Review*. 1928;32(1):97-109.
- [118] Hobbs PC. Reaching the shot noise limit for \$10. *Optics & Photonics News*. 1991;2(4):17-23.
- [119] Griffiths DJ. *Introduction to quantum mechanics*. Upper Saddle River, New Jersey (USA): Prentice Hall; 1994.
- [120] Hamamatsu Photonics K.K., Si Photodiodes - S2386 series (KSPD1035E09); Available at: <https://www.hamamatsu.com/eu/en/product/type/S2386-8K/index.html> (Accessed: 20 September 2021).
- [121] JCGM 101:2008, Evaluation of measurement data — Supplement 1 to the “Guide to the expression of uncertainty in measurement” — Propagation of distributions using a Monte Carlo method. Joint Committee for Guides in Metrology, Bureau International des Poids et Mesures; 2008.
- [122] Monzón JJ, Sánchez-Soto LL. Absorbing beam splitter in a Michelson interferometer. *Applied Optics*. 1995 dec;34(34):7834-9.
- [123] Monzón JJ, Sánchez-Soto LL. Characterization of symmetric, absorbing 50-50 beam splitters. *Applied Optics*. 1996;35(1):106-12.
- [124] Yeh P, Yariv A, Hong CS. Electromagnetic propagation in periodic stratified media. I. General theory. *Journal of the Optical Society of America*. 1977;67(4):423-38.

- [125] Vukusic P, Stavenga DG. Physical methods for investigating structural colours in biological systems. *Journal of the Royal Society Interface*. 2009;6(January):S133-48.
- [126] Yakubovsky DI, Arsenin AV, Stebunov YV, Fedyanin DY, Volkov VS. Optical constants and structural properties of thin gold films. *Optics Express*. 2017 oct;25(21):25574.
- [127] Clapham PB, Downs MJ, King RJ. Some Applications of Thin Films to Polarization Devices. *Applied Optics*. 1969;8(10):1965-74.
- [128] Boden SA, Bagnall DM. Bio-mimetic subwavelength surfaces for near-zero reflection sunrise to sunset. In: *IEEE 4th World Conference on Photovoltaic Energy*. vol. 2. IEEE; 2006. p. 1358-61.
- [129] Jacobs M, Lopez-Garcia M, Phrathep OpP, Lawson T, Oulton R, Whitney HM. Photonic multilayer structure of Begonia chloroplasts enhances photosynthetic efficiency. *Nature Plants*. 2016;2(November):16162.
- [130] Yeh P. *Optical Waves in Layered Media*. Hoboken, USA: John Wiley and Sons Inc; 1998.
- [131] Johnson PB, Christy RW. Optical constants of transition metals: Ti, V, Cr, Mn, Fe, Co, Ni, and Pd. *Physical Review B*. 1974;9(12):5056-70.
- [132] Johnson PB, Christy RW. Optical Constants of the Noble Metals. *Physical Review B*. 1972;6(12):4370-9.
- [133] Schott AG, N-BK 7 (517642.251); 2014. Available at: https://shop.schott.com/advanced_optics/en/Optical-Glass/SCHOTT-N-BK7/c/optical-glass/glass-SCHOTT%20N-BK7%C2%AE (Accessed: 20 September 2021).
- [134] Karakouz T, Holder D, Goomanovsky M, Vaskevich A, Rubinstein I. Morphology and refractive index sensitivity of gold island films. *Chemistry of Materials*. 2009 dec;21(24):5875-85.
- [135] Petrov I, Barna PB, Hultman L, Greene JE. Microstructural evolution during film growth. *Journal of Vacuum Science & Technology A: Vacuum, Surfaces, and Films*. 2003 sep;21(5):S117-28.
- [136] Yakubovsky DI, Stebunov YV, Kirtaev RV, Ermolaev GA, Mironov MS, Novikov SM, Arsenin AV, Volkov VS. Ultrathin and Ultrasoother Gold Films on Monolayer MoS₂. *Advanced Materials Interfaces*. 2019 jul;6(13):1900196.
- [137] Rosenblatt G, Simkhovich B, Bartal G, Orenstein M. Nonmodal Plasmonics: Controlling the Forced Optical Response of Nanostructures. *Physical Review X*. 2020 mar;10(1):011071.
- [138] Vossen JL, Kern W. *Thin film processes*. New York: Academic Press; 1978.

- [139] Vossen JL, Kern W. Thin film processes II. San Diego, CA: Academic Press, Inc; 1991.
- [140] Bergamin A, Cavagnero G, Mana G. A displacement and angle interferometer with subatomic resolution. *Review of Scientific Instruments*. 1993;64(11):3076-81.
- [141] Büchner HJJ, Jäger G. A novel plane mirror interferometer without using corner cube reflectors. *Measurement Science and Technology*. 2006;17(4):746-52.
- [142] Hu P, Zhu J, Zhai X, Tan J. DC-offset-free homodyne interferometer and its nonlinearity compensation. *Optics Express*. 2015;23(7):8399-408.
- [143] Hosoe S. Highly precise and stable laser displacement measurement interferometer with differential optical passes in practical use. *Nanotechnology*. 1993;4(2):81-5.
- [144] Cui J, He Z, Jiu Y, Tan J, Sun T. Homodyne laser interferometer involving minimal quadrature phase error to obtain subnanometer nonlinearity. *Applied Optics*. 2016;55(25):7086-92.
- [145] Liu J, Azzam RM. Polarization properties of corner-cube retroreflectors: theory and experiment. *Applied optics*. 1997;36(7):1553-9.
- [146] Jones RC. A New Calculus for the Treatment of Optical Systems I . Description and Discussion of the Calculus. *Journal of the Optical Society of America*. 1941;31(7):488-93.
- [147] Hurwitz H, Jones RC. A New Calculus for the Treatment of Optical Systems II. Proof of Three General Equivalence Theorems. *Journal of the Optical Society of America*. 1941;31(7):493-9.
- [148] Jones RC. A New Calculus for the Treatment of Optical Systems III. The Sohncke Theory of Optical Activity. *Journal of the Optical Society of America*. 1941;31(7):500-3.
- [149] Jones RC. A New Calculus for the Treatment of Optical Systems. IV. *Journal of the Optical Society of America*. 1942;32(8):486-93.
- [150] Jones RC. A New Calculus for the Treatment of Optical Systems V. A More General Formulation, and Description of Another Calculus. *Journal of the Optical Society of America*. 1947;37(2):107-10.
- [151] Jones RC. A New Calculus for the Treatment of Optical Systems VI. Experimental Determination of the Matrix. *Journal of the Optical Society of America*. 1947;37(2):110-2.
- [152] Jones RC. A New Calculus for the Treatment of Optical Systems. VII. Properties of the N-Matrices. *Journal of the Optical Society of America*. 1948;38(8):671-85.

- [153] Brown DD, Freise A. FINESSE (<http://www.gwoptics.org/finesse>); 2014.
- [154] Freise A, Heinzl G, Lück H, Schilling R, Willke B, Danzmann K. Frequency-domain interferometer simulation with higher-order spatial modes. *Classical and Quantum Gravity*. 2004;21(5):S1067-74.
- [155] Harry GM. Advanced LIGO: The next generation of gravitational wave detectors. *Classical and Quantum Gravity*. 2010;27(8):084006.
- [156] Yun G, Crabtree K, Chipman RA. Three-dimensional polarization ray-tracing calculus I: Definition and diattenuation. *Applied Optics*. 2011;50(18):2855-65.
- [157] Yun G, McClain SC, Chipman RA. Three-dimensional polarization ray-tracing calculus II: retardance. *Applied optics*. 2011;50(18):2866-74.
- [158] Yeh P. Electromagnetic propagation in birefringent layered media. *Journal of the Optical Society of America*. 1979;69(5):742-56.
- [159] Bridges A, Yacoot A, Kissinger T, Tatam RP. Strap - scattering and transfer matrix modelling model python - <https://github.com/strapy-project/>; 2020.
- [160] Mana G. Diffraction effects in optical interferometers illuminated by laser sources. *Metrologia*. 1989;26(2):87-93.
- [161] Yeh P. Optics of anisotropic layered media: A new 4×4 matrix algebra. *Surface Science*. 1980;96(1-3):41-53.
- [162] Schubert M. Polarization-dependent optical parameters of arbitrarily anisotropic homogeneous layered systems. *Physical Review B*. 1996;53(8):4265-74.
- [163] Požar T, Možina J. Enhanced ellipse fitting in a two-detector homodyne quadrature laser interferometer. *Measurement Science and Technology*. 2011 aug;22(8):085301.
- [164] Watkins LR, Collet MJ. Ellipse fitting for interferometry Part 2: experimental realization. *Applied Optics*. 2014;53(32):7697-703.
- [165] Collett MJ, Tee GJ. Ellipse fitting for interferometry Part 1: static methods. *Journal of the Optical Society of America A*. 2014;31(12):2573-83.
- [166] Collett MJ, Watkins LR. Ellipse fitting for interferometry Part 3: dynamic method. *Journal of the Optical Society of America A*. 2015;32(3):491-6.
- [167] Hu P, Zhu J, Guo X, Tan J. Compensation for the variable cyclic error in homodyne laser interferometers. *Sensors*. 2015;15(2):3090-106.
- [168] Kim JA, Kim JW, Kang CS, Eom TB, Ahn J. A digital signal processing module for real-time compensation of nonlinearity in a homodyne interferometer using a field-programmable gate array. *Measurement Science and Technology*. 2009 jan;20(1):017003.

- [169] Zhu J, Hu P, Tan J. Homodyne laser vibrometer capable of detecting nanometer displacements accurately by using optical shutters. *Applied Optics*. 2015;54(34):10196-9.
- [170] Wang C, Burnham-Fay ED, Ellis JD. Real-time FPGA-based Kalman filter for constant and non-constant velocity periodic error correction. *Precision Engineering*. 2017;48:133-43.
- [171] Bridges A, Yacoot A, Kissinger T, Tatam RP. Multiple intensity reference interferometry for the correction of sub-fringe displacement non-linearities. Submitted to *Measurement Science and Technology*. 2021.
- [172] Bridges A, Yacoot A, Kissinger T, Tatam RP. Correction of periodic non-linearities in homodyne optical interferometry by measurement arm intensity normalisation. Submitted to *Optics Letters*. 2021.
- [173] Bookstein FL. Fitting conic sections to scattered data. *Computer Graphics and Image Processing*. 1979;9(1):56-71.
- [174] Fitzgibbon AW, Maurizio P, Fisher RB. Direct Least Squares Fitting of Ellipses. *IEEE Transactions on Pattern Analysis and Machine Intelligence*. 1996;21(5):476-80.
- [175] Harker M, O’Leary P, Zsombor-Murray P. Direct type-specific conic fitting and eigenvalue bias correction. *Image and Vision Computing*. 2008;26(3):372-81.
- [176] Prasad DK, Leung MKH, Quek C. ElliFit: An unconstrained, non-iterative, least squares based geometric Ellipse Fitting method. *Pattern Recognition*. 2013;46(5):1449-65.
- [177] Szpak ZL, Chojnacki W, van den Hengel A. Guaranteed Ellipse Fitting with a Confidence Region and an Uncertainty Measure for Centre, Axes, and Orientation. *Journal of Mathematical Imaging and Vision*. 2015;52(2):173-99.
- [178] Li Z, Herrmann K, Pohlenz F. A neural network approach to correcting non-linearity in optical interferometers. *Measurement Science and Technology*. 2003;14(3):376-81.
- [179] Boggs PT, Donaldson JR, Byrd Rh, Schnabel RB. Algorithm 676: ODR-PACK: Software for Weighted Orthogonal Distance Regression. *ACM Transactions on Mathematical Software (TOMS)*. 1989;15(4):348-64.
- [180] Walck C. Hand-book on statistical distributions for experimentalists. Stockholm University; 1996.
- [181] Ahn J, Kim JAJW, Kang CS, Kim JAJW, Kim S. A passive method to compensate nonlinearity in a homodyne interferometer. *Optics Express*. 2009;17(25):23299-308.

- [182] Wu H, Zhang F, Liu T, Li J, Qu X. Absolute distance measurement with correction of air refractive index by using two-color dispersive interferometry. *Optics Express*. 2016;24(21):24361-76.
- [183] Liu Z, Li W, Bayanheshig, Li X, Jiang S, Song Y, Lv Q. Two-color heterodyne laser interferometry for long-distance stage measurement with correction of uncertainties in measured optical distances. *Scientific Reports*. 2017;7(1):8173.
- [184] Meiners-Hagen K, Schödel R, Pollinger F, Abou-Zeid A. Multi-Wavelength Interferometry for Length Measurements Using Diode Lasers. *Measurement Science Review*. 2009;9(1):16-26.
- [185] Harris FJ. On the Use of Windows for Harmonic Analysis with the Discrete Fourier Transform. *Proceedings of the IEEE*. 1978;66(1):51-83.
- [186] MacNeille SM. Beam Splitter, US Patent No. 2403731; 1946.
- [187] Guanliang P, Ruiquan Y, Jiankun Y, Shengli C, Juncal Y. Design of MacNeille polarizing beam splitter. In: 2009 Symposium on Photonics and Optoelectronics. Wuhan, China: IEEE eXpress Conference Publishing; 2009. p. 1-4.
- [188] ThorLabs, Inc., PBS251; Available at: https://www.thorlabs.com/newgrouppage9.cfm?objectgroup_id=739 (Accessed: 22 September 2021).
- [189] ThorLabs, Inc., PBS25-633; Available at: https://www.thorlabs.com/newgrouppage9.cfm?objectgroup_id=7490 (Accessed: 22 September 2021).
- [190] ThorLabs, Inc., PBSW-633; Available at: https://www.thorlabs.com/newgrouppage9.cfm?objectgroup_id=6004 (Accessed: 22 September 2021).
- [191] Queensgate Instruments Ltd., NPS-X-15A (Version issue 1); Available at: https://www.nanopositioning.com/wp-content/uploads/2018/05/NPS-X-15A_NPS-X-15B-Datasheet.pdf (Accessed: 22 September 2021).
- [192] Branch MA, Coleman TF, Li Y. A Subspace, Interior, and Conjugate Gradient Method for Large-Scale Bound-Constrained Minimization Problems. *SIAM Journal on Scientific Computing*. 1999;21(1):1-29.
- [193] Virtanen P, Gommers R, Oliphant TE, Haberland M, Reddy T, Cournapeau D, Burovski E, Peterson P, Weckesser W, Bright J, van der Walt SJ, Brett M, Wilson J, Millman KJ, Mayorov N, Nelson ARJ, Jones E, Kern R, Larson E, Carey CJ, Polat I, Feng Y, Moore EW, VanderPlas J, Laxalde D, Perktold J, Cimrman R, Henriksen I, Quintero EA, Harris CR, Archibald AM, Ribeiro AH, Pedregosa F, van Mulbregt P, Vijaykumar A, Bardelli AP, Rothberg A, Hilboll A, Kloeckner A, Scopatz A, Lee A, Rokem A, Woods CN, Fulton C, Masson C, Häggström C, Fitzgerald C, Nicholson DA, Hagen DR, Pasechnik DV, Olivetti E, Martin E, Wieser E, Silva F, Lenders F, Wilhelm F, Young G, Price GA, Ingold GL, Allen GE, Lee GR, Audren H, Probst I, Dietrich JP, Silterra J, Webber JT, Slavič J, Nothman J, Buchner J, Kulick J, Schönberger JL, de Miranda Cardoso JV, Reimer J, Harrington J, Rodríguez JLC, Nunez-Iglesias

J, Kuczynski J, Tritz K, Thoma M, Newville M, Kümmerer M, Bolingbroke M, Tartre M, Pak M, Smith NJ, Nowaczyk N, Shebanov N, Pavlyk O, Brodtkorb PA, Lee P, McGibbon RT, Feldbauer R, Lewis S, Tygier S, Sievert S, Vigna S, Peterson S, More S, Pudlik T, Oshima T, Pingel TJ, Robitaille TP, Spura T, Jones TR, Cera T, Leslie T, Zito T, Krauss T, Upadhyay U, Halchenko YO, Vázquez-Baeza Y. SciPy 1.0: fundamental algorithms for scientific computing in Python. *Nature Methods*. 2020;17(3):261-72.

- [194] Stone JA, Stejskal A, Howard L. Absolute interferometry with a 670-nm external cavity diode laser. *Applied Optics*. 1999;38(28):5981-94.
- [195] Golubev AN, Chekhovsky AM. Three-color optical range finding. *Applied Optics*. 1994;33(31):7511-7.