Robust reference algorithms for form metrology: Application to aspherical and freeform optics

Thèse préparée à l'Ecole Normale Supérieure de Paris-Saclay

École doctorale n°579 Sciences mécaniques et énergétiques, matériaux et géosciences (SMEMAG)
Spécialité de doctorat: Génie Mécanique

Thèse présentée et soutenue à Cachan, le 05 Décembre 2019, par

Yassir AREZKI

Composition du Jury :

Jean-François Fontaine
Professeur, Université de Bourgogne  
Rapporteur
Denis Teissandier
Professeur, Université de Bordeaux  
Rapporteur
Jean-Marc Linarès
Professeur, Aix Marseille Université  
Examinateur
Fengzhou Fang
Professeur, University College Dublin & Tianjin University  
Examinateur
Olivier Bruneau
Professeur, Université Paris-Sud  
Examinateur
Nabil Anwer
Professeur, Université Paris-Sud  
Directeur de thèse
Hichem Nouira
Chercheur (Dr. HDR.), LNE/Cnam  
Co-Directeur de thèse
Charyar Mehdi-Souzani
MCF, Université Paris 13  
Co-Encadrant de thèse
Muriel Thomasset
Chercheuse, Synchrotron SOLEIL  
Invitée
Contents

Nomenclature v
List of Figures vii
List of Tables xi

General introduction 1

1 Metrology of aspherical and freeform surfaces 5
1.1 Introduction ............................................ 6
1.2 Mathematical descriptions of aspherical and freeform surfaces ......... 7
    1.2.1 Mathematical descriptions of aspherical surfaces ............ 8
    1.2.2 Mathematical descriptions of freeform surfaces in optics .... 12
1.3 Manufacturing of aspherical and freeform surfaces .................. 16
1.4 Measurement of aspherical and freeform surfaces ................... 22
1.5 Conclusion ............................................. 25

2 Fitting algorithms of aspherical and freeform surfaces 27
2.1 Introduction ............................................ 28
2.2 Form error assessment .................................. 30
    2.2.1 Waviness, roughness and form error ....................... 31
    2.2.2 Minimum Zone vs. Least Squares ....................... 32
2.3 Literature review ...................................... 35
    2.3.1 Canonical forms .................................. 35
    2.3.2 FreeForm and aspherical shapes ...................... 38
    2.3.3 Summary ......................................... 39
2.4 Form assessment using EPF and PDIP .......................... 40
2.4.1 Exponential Penalty Function (EPF) .......... 40
2.4.2 Primal-Dual Interior Point method (PDIP) .......... 41
2.4.3 Comparison of EPF and PDIP using reference data .......... 42
2.4.4 Comparison of EPF and PDIP using measured data .......... 47

2.5 The Hybrid Trust Region algorithm (HTR) .......... 49
2.5.1 Application on reference data .......... 52
2.5.2 Application on benchmark data .......... 55
2.5.3 Experimental investigation .......... 55

2.6 Numerical uncertainties .......... 58
2.7 Conclusion .......... 62

3 Validation of fitting algorithms ........ 65
3.1 Introduction .......... 66
3.2 Literature review on algorithms' validation .......... 67
3.3 Reference data generation .......... 69
  3.3.1 Least Squares .......... 69
  3.3.2 Minimum Zone .......... 71
3.4 Extension to non-vertex solutions for aspherical surfaces .......... 72
  3.4.1 Vertex vs. non-vertex solutions .......... 73
  3.4.2 General approach .......... 75
  3.4.3 Case study .......... 77
3.5 Beyond data generation .......... 78
  3.5.1 Requirements on generated reference data .......... 78
  3.5.2 Reference data uncertainty .......... 83
  3.5.3 Difficulty vs. performance diagram .......... 85
  3.5.4 Case Study .......... 92
3.6 Conclusion .......... 94

4 Artefacts' design and inter-laboratory comparison ........ 97
4.1 Introduction .......... 98
4.2 Thermo-invariant material standards .......... 98
  4.2.1 Literature review on existing material standards .......... 98
  4.2.2 Design of the material standards .......... 100
4.3 Measurements .......... 106
## Nomenclature

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>BCR</td>
<td>Community Bureau of Reference</td>
</tr>
<tr>
<td>BFGS</td>
<td>Broyden-Fletcher-Goldfarb-Shanno</td>
</tr>
<tr>
<td>CAD</td>
<td>Computer-aided design</td>
</tr>
<tr>
<td>CGH</td>
<td>Computer Generated Hologram</td>
</tr>
<tr>
<td>CMM</td>
<td>Coordinate Measuring Machine</td>
</tr>
<tr>
<td>CNC</td>
<td>Computer Numerical Control</td>
</tr>
<tr>
<td>CSI</td>
<td>Coherence Scanning Interferometer</td>
</tr>
<tr>
<td>DE</td>
<td>Differential Evolution</td>
</tr>
<tr>
<td>DI</td>
<td>Designated Institute</td>
</tr>
<tr>
<td>EMPR</td>
<td>European Metrology Research Programme</td>
</tr>
<tr>
<td>EOS</td>
<td>European Optical Society</td>
</tr>
<tr>
<td>EPF</td>
<td>Exponential Penalty Function</td>
</tr>
<tr>
<td>FTS</td>
<td>Fast Tool Servo</td>
</tr>
<tr>
<td>GPS</td>
<td>Geometrical and Product Specification</td>
</tr>
<tr>
<td>GUM</td>
<td>Guide to the Expression of Uncertainty in Measurement</td>
</tr>
<tr>
<td>IPP</td>
<td>Institute of Plasma Physics</td>
</tr>
<tr>
<td>ITO</td>
<td>Institute of Applied Optics</td>
</tr>
<tr>
<td>KCRV</td>
<td>Key Comparison Reference Value</td>
</tr>
<tr>
<td>KKT</td>
<td>Karush-Kuhn-Tucker</td>
</tr>
<tr>
<td>LICQ</td>
<td>Linear Independence Constraint Qualification</td>
</tr>
<tr>
<td>LM</td>
<td>Levenberg-Marquardt</td>
</tr>
<tr>
<td>LS</td>
<td>Least Squares</td>
</tr>
<tr>
<td>MC</td>
<td>Minimum Circumscribed</td>
</tr>
<tr>
<td>MFG</td>
<td>Modular Freeform Gauge</td>
</tr>
<tr>
<td>MI</td>
<td>Maximum Inscribed</td>
</tr>
<tr>
<td>MPA</td>
<td>Molded Polymer Asphere</td>
</tr>
<tr>
<td>Acronym</td>
<td>Description</td>
</tr>
<tr>
<td>---------</td>
<td>-------------</td>
</tr>
<tr>
<td>MRF</td>
<td>Magnetorheological Finishing</td>
</tr>
<tr>
<td>MZ</td>
<td>Minimum Zone</td>
</tr>
<tr>
<td>NIST</td>
<td>National Institute of Standards and Technology</td>
</tr>
<tr>
<td>NMI</td>
<td>National Metrology Institute</td>
</tr>
<tr>
<td>NMIJ</td>
<td>National Metrology Institute of Japan</td>
</tr>
<tr>
<td>NPL</td>
<td>National Physical Laboratory</td>
</tr>
<tr>
<td>PDIP</td>
<td>Primal-Dual Interior Point method</td>
</tr>
<tr>
<td>PGM</td>
<td>Precision Glass Molding</td>
</tr>
<tr>
<td>PMD</td>
<td>Phase Measuring Deflectometry</td>
</tr>
<tr>
<td>PTB</td>
<td>Physikalisch-Technische Bundesanstalt</td>
</tr>
<tr>
<td>PV</td>
<td>Peak-to-Valley</td>
</tr>
<tr>
<td>RMS</td>
<td>Root of Mean Squares</td>
</tr>
<tr>
<td>SPDT</td>
<td>Single Point Diamond Turning</td>
</tr>
<tr>
<td>SSI</td>
<td>Subaperture Stitching Interferometer</td>
</tr>
<tr>
<td>STS</td>
<td>Slow Tool Servo</td>
</tr>
<tr>
<td>SVD</td>
<td>Singular Value Decomposition</td>
</tr>
<tr>
<td>SVM</td>
<td>Support Vector Machine</td>
</tr>
<tr>
<td>TNO</td>
<td>Netherlands Organisation of Applied Scientific Research</td>
</tr>
<tr>
<td>uhp-CMMs</td>
<td>ultra-high precision Coordinate Measuring Machines</td>
</tr>
<tr>
<td>UNOTT</td>
<td>University of Nottingham</td>
</tr>
<tr>
<td>UPM</td>
<td>Ultra-precision machining</td>
</tr>
<tr>
<td>UPOB</td>
<td>Ultra-Precise Surface Manufacturing</td>
</tr>
<tr>
<td>WLI</td>
<td>White Light Interferometry</td>
</tr>
</tbody>
</table>
### List of Figures

<table>
<thead>
<tr>
<th>Figure</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Metrology components of aspherical and freeform surfaces</td>
<td>2</td>
</tr>
<tr>
<td>1.1</td>
<td>A photograph of aspherical surfaces</td>
<td>6</td>
</tr>
<tr>
<td>1.2</td>
<td>A photograph of a freeform surface used in lighting systems</td>
<td>7</td>
</tr>
<tr>
<td>1.3</td>
<td>A photograph of a freeform with regular surface texture</td>
<td>7</td>
</tr>
<tr>
<td>1.4</td>
<td>Aspherical surface</td>
<td>9</td>
</tr>
<tr>
<td>1.5</td>
<td>Plot of the orthogonal basis elements $m = 1, ..., 5$</td>
<td>11</td>
</tr>
<tr>
<td>1.6</td>
<td>A best fit sphere of a mild asphere</td>
<td>12</td>
</tr>
<tr>
<td>1.7</td>
<td>Classification of optical components regarding quality, quantity, price and production methods</td>
<td>16</td>
</tr>
<tr>
<td>1.8</td>
<td>Process sequences from idea to product with classical standard process chain for freeform manufacturing</td>
<td>17</td>
</tr>
<tr>
<td>1.9</td>
<td>SPDT machine from Apollo Optical Systems</td>
<td>17</td>
</tr>
<tr>
<td>1.10</td>
<td>Ultra-precision machining (UPM) methods</td>
<td>19</td>
</tr>
<tr>
<td>1.11</td>
<td>PGM process</td>
<td>19</td>
</tr>
<tr>
<td>1.12</td>
<td>Polymer molding technique</td>
<td>20</td>
</tr>
<tr>
<td>1.13</td>
<td>MR fluid – Working principle</td>
<td>20</td>
</tr>
<tr>
<td>1.14</td>
<td>Manufacturing of an aspherical lens using a MR process</td>
<td>21</td>
</tr>
<tr>
<td>1.15</td>
<td>Principle of the Q22 multiple-axis computer-controlled MRF machine</td>
<td>21</td>
</tr>
<tr>
<td>1.16</td>
<td>Modified ink-jet 3D printing process</td>
<td>22</td>
</tr>
<tr>
<td>1.17</td>
<td>Measurement of an optical lens using CMM</td>
<td>23</td>
</tr>
<tr>
<td>1.18</td>
<td>Principle of deflectometry</td>
<td>24</td>
</tr>
<tr>
<td>1.19</td>
<td>The Michelson interferometer</td>
<td>25</td>
</tr>
<tr>
<td>2.1</td>
<td>Form tolerance definition for : (a) line profile, (b) surface profile</td>
<td>30</td>
</tr>
<tr>
<td>2.2</td>
<td>Form error, waviness and roughness</td>
<td>31</td>
</tr>
<tr>
<td>2.3</td>
<td>Definition of form deviation. Measured data are supposed to lie exactly on the manufactured shape</td>
<td>33</td>
</tr>
<tr>
<td>Figure</td>
<td>Description</td>
<td></td>
</tr>
<tr>
<td>--------</td>
<td>-------------</td>
<td></td>
</tr>
<tr>
<td>2.4</td>
<td>Estimated MZ values obtained using MZ fitting (blue) and LS (orange). The value obtained using LS is overestimated.</td>
<td></td>
</tr>
<tr>
<td>2.5</td>
<td>Representation of aspherical surfaces based on nominal shape coefficients.</td>
<td></td>
</tr>
<tr>
<td>2.6</td>
<td>Methodology of EPF and PDIP comparison.</td>
<td></td>
</tr>
<tr>
<td>2.7</td>
<td>Generated reference data for the configuration V.</td>
<td></td>
</tr>
<tr>
<td>2.8</td>
<td>Evolution of the difference between the obtained value of MZ and the reference value for PDIP and EPF in the case of configuration III.</td>
<td></td>
</tr>
<tr>
<td>2.9</td>
<td>A photograph of the AO775 asphere.</td>
<td></td>
</tr>
<tr>
<td>2.10</td>
<td>A simplified flowchart of the hybrid trust region algorithm.</td>
<td></td>
</tr>
<tr>
<td>2.11</td>
<td>Evolution of the execution time of EPF and HTR in function of the number of points in the data set for configurations II, III and IV.</td>
<td></td>
</tr>
<tr>
<td>2.12</td>
<td>AO775 optical aspherical lens: measured data using the ultra-high precision Nanomefes machine (31390 points).</td>
<td></td>
</tr>
<tr>
<td>2.13</td>
<td>Evolution of PV value for HTR and EPF algorithms applied on measured data.</td>
<td></td>
</tr>
<tr>
<td>2.14</td>
<td>Fitting residuals of the AO775 asphere (31390 points).</td>
<td></td>
</tr>
<tr>
<td>2.15</td>
<td>Comparison of estimated MZ value when considering the continuous manufacturing profile (blue) and the measured discretized profile (orange).</td>
<td></td>
</tr>
<tr>
<td>2.16</td>
<td>Comparison of estimated MZ based on exact measured points (orange) and measurement including errors (green).</td>
<td></td>
</tr>
<tr>
<td>2.17</td>
<td>Illustration of black-box testing methods.</td>
<td></td>
</tr>
<tr>
<td>2.18</td>
<td>Uncertainty analysis of MZ fitting algorithm.</td>
<td></td>
</tr>
<tr>
<td>2.19</td>
<td>Difference between obtained PV values and MZ\textsubscript{ref} values and their associated uncertainties.</td>
<td></td>
</tr>
<tr>
<td>3.1</td>
<td>Type F1 software measurement standard.</td>
<td></td>
</tr>
<tr>
<td>3.2</td>
<td>Type F2 software measurement standard.</td>
<td></td>
</tr>
<tr>
<td>3.3</td>
<td>A summarised flowchart for the generation of reference data for MZ fitting.</td>
<td></td>
</tr>
<tr>
<td>3.4</td>
<td>Representation of contacting points to the enclosing envelope.</td>
<td></td>
</tr>
<tr>
<td>3.5</td>
<td>Representation of a vertex solution for the case of a linear programming involving the two variables X1 and X2.</td>
<td></td>
</tr>
<tr>
<td>3.6</td>
<td>Representation of a vertex solution for the case of a nonlinear programming involving the two variables X1 and X2.</td>
<td></td>
</tr>
</tbody>
</table>
3.7 Generated reference data with a non-vertex solution  77
3.8 Data residuals after MZ fitting  78
3.9 Maximum inscribed circle problem with multiple associated features  79
3.10 Reference data with form deviation generated using: (a) fractional Brownian motion, (b) Gaussian noise  80
3.11 Illustration of uncertainty on MZ_ref  83
3.12 Difference between MZ_ref,t and the actual MZ_ref  84
3.13 Performance measure in function of degree of difficulty for a given algorithm  86
3.14 Survey questionee' profile  87
3.15 Performance measure in function of degree of difficulty for the "algorithm under test"  93
4.1 Artefacts for the calibration of tactile CMMs  99
4.2 Artefacts for the calibration of optical CMMs  99
4.3 Photograph of NPL freeform artefact  100
4.4 Modular Freeform Gauge configuration  100
4.5 The "Doppelsinusfläche" artefact  101
4.6 The "Shoe Model" artefact  101
4.7 Layout of the ‘artificial’ added form error to the aspherical shape  102
4.8 Construction of the thermo-invariant material standard for aspherical shape seen in the Y=0 plane  103
4.9 Design of the thermo-invariant material standard "Artefact I"  104
4.10 Design of the thermo-invariant material standard "Artefact II"  105
4.11 A photograph of the manufactured artefacts. The diameter of the two artefacts is 50 mm  105
4.12 Participating partners to the inter-laboratory comparison  106
4.13 The LNE’s high precision profilometer  107
4.14 Subaperture Stitching Interferometer (SSI)  108
4.15 Zygo NexView™ NX2  108
4.16 LuphoScan 260 HD measuring machine  109
4.17 MarForm MFU 200 Aspheric 3D measuring machine  110
4.18 VTT’s multi sensor optical profilometer  110
4.19 UA3P-4000 measuring machine ........................................ 111
4.20 ITO - NPMM200 measuring machine ................................. 112
4.21 Optical microscopy view of Artefact I. Left: before cleaning, right: after cleaning .................................................. 113
4.22 Measurement strategy .................................................... 114
4.23 Cropping out data illustrated on a measurement made by NMIJ .................................................. 114
4.24 Obtained MZ values and associated uncertainties for Artefact I .................................................. 119
4.25 Artefact II, manufacturer’s report ....................................... 122
4.26 Obtained MZ values for Artefact II .................................... 123
List of Tables

1.2 Typical boundary functions in freeform surface descriptions ............... 13
1.3 Overview of polar-defined freeform surface representations ................... 14
1.4 Overview of Cartesian-defined freeform surface representations ............... 15

2.1 Form tolerance specifications: symbols and standards ......................... 29
2.2 Form tolerance specifications: symbols and standards ......................... 32
2.3 An overview of the performances of the selected MZ fitting techniques .... 40
2.4 Nominal shape coefficients of generated reference data ....................... 43
2.5 Relative errors between obtained values of MZ and reference values in (%) . Blue: PDIP, orange: EPF ............................................ 47
2.6 Nominal shape parameters of the AO775 asphere ................................ 48
2.7 PV results obtained using EPF, PDIP and L-BFGS (LS) applied on Nanomefos measured data, $N = 1129$ ................................................. 49
2.8 Nominal coefficients of aspherical surfaces used for the validation of the HTR algorithm ................................................................. 53
2.9 Values of PV-MZ$_{ref}$ (mm) and execution time (s) for HTR and EPF (configuration II) ................................................................. 53
2.10 Values of PV-MZ$_{ref}$ (mm) and execution time (s) for HTR and EPF (configuration III) ................................................................. 54
2.11 Values of PV-MZ$_{ref}$ (mm) and execution time (s) for HTR and EPF (configuration IV) ................................................................. 54
2.12 Comparison of HTR and EPF on benchmark data (8100 points) .......... 55
2.13 Comparison of HTR and EPF on measured data (31390 points) .......... 56

3.1 Initial positions before fitting ....................................................... 81
3.2 Obtained PV values using EPF and PDIP without taking LS as initial alignment ................................................................. 82
### LIST OF TABLES

<table>
<thead>
<tr>
<th>Table</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>3.3</td>
<td>Obtained results using EPF and PDIP when taking LS as initial alignment</td>
<td>82</td>
</tr>
<tr>
<td>3.4</td>
<td>Estimation of $\alpha_1$, $\alpha_2$ and $\alpha_3$ based on the survey answers</td>
<td>89</td>
</tr>
<tr>
<td>3.5</td>
<td>Estimated values of $E_0$ and $T_0$ for $\lambda_{min}$, $\lambda_{max}$ and $\lambda_0$</td>
<td>90</td>
</tr>
<tr>
<td>4.1</td>
<td>Nominal shape parameters of the aspheric surface used in &quot;Artefact I&quot;</td>
<td>103</td>
</tr>
<tr>
<td>4.2</td>
<td>Nominal shape parameters of the aspheric surface used in &quot;Artefact I&quot;</td>
<td>105</td>
</tr>
<tr>
<td>4.3</td>
<td>Measurements and number of collected data points for the comparison</td>
<td>115</td>
</tr>
<tr>
<td>4.4</td>
<td>Obtained MZ values for Artefacts I and II</td>
<td>118</td>
</tr>
<tr>
<td>4.5</td>
<td>Layout of the residuals' maps</td>
<td>121</td>
</tr>
</tbody>
</table>
General introduction

The shapes of optical elements have been continuously changing from simple spherical surfaces to conical shapes (ellipsoid, paraboloid, hyperboloid, etc.) then aspherical surfaces up to freeforms nowadays [1]. The more the shapes are complicated, the more degrees of freedom they give to optics’ designers and the more the resulting optical system is efficient. Freeforms along with aspherical surfaces (which are a subset of freeform surfaces with a rotational symmetry) have increasing applications in many domains including imaging systems, lithography, automotive, medical devices, to name a few. The main advantages of these elements are their ability to improve imaging quality, reducing the number and weight of the imaging system, increasing the compactness of optics as well as their ability to correct optical aberrations [2]. In the past, the use and applications of freeform and aspherical surfaces were limited by the manufacturing costs and techniques. Their use becomes more popular with the advancements made in manufacturing techniques such as moulding, grinding and polishing methods.

Nowadays, modern optics manufacturers are able to remove material at the single nm level [3, 4]. However, the opportunities given by these modern manufacturing processes are not exploitable without the ability to measure the manufactured surfaces with sufficient accuracy. Hence, metrology is the limiting factor for the development of aspherical and freeform surfaces. A number of institutions are working on the metrology of such type of optical elements across the Europe and over the world since there is an urgent need to strengthen the metrology capabilities at the level on National Metrology Institutes (NMIs), Designated Institutes (DIs) as well as industry. In this scope, a European project, called EMRP Form-IND10 : Optical and tactile metrology for absolute form characterisation, was carried out between 2011 and 2014 under the coordination of Dr. Michael Schulz from the Physikalisch-Technische Bundesanstalt (PTB). Within this project, uncertainty below 100 nm was fulfilled for aspherical surfaces. Nevertheless, this achievement does not satisfy the need of research institutes and industry claiming an uncertainty at the nm level. This was strongly emphasised during the discussions
conducted at the High Level Expert Meetings and workshops of the Competence Centre of Ultra-Precise Surface Manufacturing (UPOB) and the European Optical Society (EOS) [3, 4].

The enhancement of metrology capabilities of aspherical and freeform surfaces requires the enhancement of each component in the metrology loop illustrated in figure 1. This includes the development of ultra-high precision measurement instruments, reference software and reference data (Softgauges) as well as thermo-invariant material standards.

---

The European project FreeFORM-15SIB01 coordinated by Prof. Hichem Nouira from LNE (France) and involving 19 international partners, of which this thesis is part, addresses these issues [5]. The main objectives of the FreeFORM project are listed below.

1. Develop robust reference Minimum Zone (MZ) algorithms for form error determination including the generation of reference data for the validation of developed
2. Design, manufacture and characterise innovative aspherical and freeform reference optical elements made of thermo-invariant materials in order to develop reference/primary calibration chain at the European NMIs and DIs and to facilitate the transfer of traceability between NMIs/DIs and stakeholders,

3. Develop advanced techniques for data analysis (stitching, data fusion, etc.),

4. Improve measurement capabilities of NMIs/DIs on aspherical and freeform standards,

5. Develop a strategy for the long-term operation of the capability developed including the take up of the technology and measurement infrastructure developed during the project.

Thus, this thesis addresses the first, the second and the last points. In fact, MZ fitting is still a major challenge for complex shapes such as aspheric and freeform surfaces. Robust and deterministic MZ fitting algorithms with nanometric accuracy are essential for form metrology. Some attempts have been made using Least Squares (LS) fitting instead of MZ but the resulting form error is overestimated. The development of fitting algorithms also requires the development of appropriate tools for the validation of the underlying output. Reference data, also called Softgauges, are used for this aim. Reference data must be generated with a defined logic, expected values of the geometrical and statistical quantities.

The development of innovative thermo-invariant reference asphere and freeform surfaces is also considered. These will be considered as traceable artefacts, which guarantee a transfer of the reference metrology chain and quality assurance. The developed artefacts should be calibrated using different ultra-high precision reference measurement instruments.

In Chapter 1, metrology of aspherical and freeform surfaces from a general perspective is discussed. First, the properties of the different mathematical representations of aspherical and freeform surfaces are outlined. Then, descriptions of some of the existing manufacturing processes as well as measuring instruments are presented.

In Chapter 2, form assessment of aspherical and freeform surfaces is discussed. The MZ problem is formulated and compared to LS. A comprehensive literature review is
conducted in order to select appropriate methods to solve the resulting minimisation problem. Two selected methods are compared using reference and measured data. An additional method with better performances is suggested and tested on reference, benchmark and measured data.

Chapter 3 deals with the validation of metrology algorithms. The existing methods for algorithms’ validation are listed. An emphasis is put on the one developed by Forbes et al. [6] and requirements on generated reference data are discussed. Definitions of two metrics namely the degree of difficulty and the performance measure are suggested. A validation procedure for reference and industrial metrology software, based on the two previously described metrics, is described.

In the last Chapter, a design of two thermo-invariant material standards is suggested (one for aspherics and the other for freeform surfaces). The two artefacts are manufactured and calibrated using measuring machines of the different project partners. This aims at characterising the manufactured artefacts and conducting an inter-laboratory comparison. The different results of the comparison are reported and discussed.
Chapter 1

Metrology of aspherical and freeform surfaces
1.1 Introduction

In the context of optical design, aspherical lenses are defined as rotationally symmetric surfaces whose radius of curvature varies gradually from the centre of the lens (figure 1.1). They were first introduced by Kepler in 1611 but their use begins in the twentieth century with the spread of Computer Numerical Control (CNC) machining processes [7]. Aspherical surfaces are mainly used because of their ability to eliminate spherical aberration. Spherical aberration is an optical phenomenon that causes indecent beams not to focus at the same point which results in unclear images and hence imaging systems with low performances [8]. Moreover, the use of one aspherical surface could replace a set of conventional spherical lenses, hence, resulting in a reduction of weights, cost, etc.

![Figure 1.1: A photograph of aspherical surfaces [9]](image)

Freeform surfaces could be regarded as surfaces with no degrees of invariance (figure 1.2) or surfaces with regular or irregular surface texture (figure 1.3) [10]. The added complexity compared to aspherical surfaces gives optical designers more degrees of freedom and hence an enlarged scope for innovation. Moreover, the resulting imaging systems have better performance regarding the aberration correction and field of view. Also, the use of freeform surfaces results in fewer optical elements, lower mass, lower cost and reduced stray-light [11]. However, regarding their shapes, the use of aspherical and freeform surfaces brought also some drawbacks especially in manufacturing and measurements which are more challenging than for conventional optics.

The main objective of this chapter is to give a general overview of the metrology
1.2. Mathematical descriptions of aspherical and freeform surfaces

In this section, the different mathematical representations of aspherical and freeform surfaces are presented and compared. Depending on the underlying shape (continuous or discontinuous) freeform surfaces could be described using a number of tools such as Computer-Aided Design (CAD) models, functional surfaces, etc. [11]. In the following, the term freeform refers to surfaces with no degrees of invariance that could be described...
using general equations. Surfaces with regular or irregular surface texture as well as those described using piecewise functions will not be taken in consideration. The mathematical descriptions of freeform and aspherical surfaces outlined in this section mainly follow the work of Broemel et al. in [14].

1.2.1 Mathematical descriptions of aspherical surfaces

1.2.1.1 The ISO 10110-12:2007 formulation

The mathematical description of aspherical surfaces given by the ISO 10110-12:2007 [15] is the most widely used. It was first introduced by Abbe in 1899 [16]. This formulation, also called monomial formulation, combines a basic shape: a conic section (considering its optical properties that are often exploited in optical design) to which additional terms, taken as monomials, are added to describe departure from the basic shape. This formulation is expressed using an explicit equation, i.e. one spatial coordinate (in this case the sag \( z \)) is expressed in term of the others. In cylindrical polar coordinates, the formulation is expressed as \( z = f(r, \theta) \). Since aspherics are rotationally symmetric, there is no angular dependence and thus \( z = f(r) \).

This formulation is completed by specification of the aperture size \( r_{\text{max}} \) for purpose of fabrication and testing. This value specifies the range of radius values \( 0 \leq r \leq r_{\text{max}} \) for which the formulation is valid. The mathematical expression is given in eq.(1.1).

\[
z(r) = \frac{r^2}{R(1 + \sqrt{1 - (1 + \kappa) \frac{r^2}{R^2}})} + \sum_{m=0}^{M} a_m r^{2m+4}
\]  \hspace{1cm} (1.1)

where:

\( z \) : sag of the surface parallel to the optical axis (see figure 1.4)
\( r \) : radial distance
\( R \) : radius of curvature
\( \kappa \) : conic constant

- \( \kappa > 0 \) oblate ellipse
- \( \kappa = 0 \) circle
- \( -1 < \kappa < 0 \) prolate ellipse
- \( \kappa = -1 \) parabola
- \( \kappa = -1 \) hyperbola
1.2. Mathematical descriptions of aspherical and freeform surfaces

\[ a_m : \text{monomial coefficients} \]

This representation could be used to approximate any symmetric shape with arbitrary accuracy while \( M \) is allowed to be large. However, this formulation lacks from several drawbacks. In [17], it was shown that even when using double precision arithmetic, the Gram matrix resulting from the LS fitting is seriously ill-conditioned and the fitting fails when the number of monomials is more than ten (\( M > 10 \)). In the case where less than ten monomials are considered, values of \( a_m \) alternate signs causing heavy cancellation between individual terms resulting in round-off problems. Numerical issues with this representation have risen in the early years of the 21\textsuperscript{st} century when commercial optical software give more aspheric terms than could be supported by 32-bit computing systems [18]. All these factors pushed towards the development of new formulations to address these issues.

![Aspherical surface](image)

**Figure 1.4: Aspherical surface**

\[ z(r) = \frac{r^2}{R(1 + \sqrt{1 - (1 + \kappa) \frac{r^2}{R^2}})} + D \left( \frac{r}{r_{\text{max}}} \right) \] (1.2)

### 1.2.1.2 Forbes’ formulations

In [17], G. Forbes came with a new representation of aspheric lenses that copes with monomial formulation drawbacks. This formulation makes use of a non-standard orthogonal basis in place of monomials. As a result, the ill-conditioning of the Gram matrix faced when using monomials could be avoided. This formulation is given in eq.(1.2) where \( D \) is defined according to the type of the aspheric (mild or strong).
1.2.1.2.1 Strong asphere

For strong aspheres, \( D \) is denoted as \( D_{\text{con}} \) and is given in eq.(1.3).

\[
D_{\text{con}}(u) = u^4 \sum_{m=0}^{M} a_m Q^\text{con}_m(u^2)
\]  

In this case, suppose we use Forbes model for strong asphere to fit a surface described by its explicit equation \( z = f(r) \). The first step consists of choosing a close-fitting conic and then finds \( a_m \) coefficients. In the LS fitting, we would like to minimise the root-mean-square sag error between the surface and the model. This could be expressed as minimising the quantity given in eq.(1.4).

\[
E^2(a_0, a_1, \ldots, a_M) = \int_0^1 (g(ur_{\text{max}}) - u^4 \sum_{m=0}^{M} a_m Q^\text{con}_m(u^2))^2 du
\]  

where \( u = \frac{r}{r_{\text{max}}} \) and \( g(u) \) is the difference between \( f \) and the best-fitting conic. To minimise \( E^2 \), its gradient with respect to \( (a_i)_{1 \leq i \leq M} \) is set to zero. This results in the linear system expressed in eq.(1.5).

\[
\sum_{n=0}^{M} G_{mn} a_n = b_m
\]  

\( b_m \) and \( G_{mn} \) are given by eq.(1.6) and eq.(1.7) respectively. \((G_{mn})_{m,n}\) coefficients define the elements of the Gram matrix of the LS fitting system.

\[
b_m = \int_0^1 g(ur_{\text{max}})Q^\text{con}_m(u^2)u^4 du
\]  

\[
G_{mn} = \int_0^1 Q^\text{con}_m(u)Q^\text{con}_n(u)u^4 du
\]  

An appropriate choice of \( Q^\text{con}_m \) polynomials gives a diagonal Gram matrix and hence make the linear system easily solved. In particular the one given in eq.(1.8)

\[
Q^\text{con}_m(x) = P_m^{(0,4)}(2x - 1)
\]  

where \( P_m^{(0,4)} \) are Jacobi polynomials of degree \( m \) [19]. Figure 1.5 shows the plot of the orthogonal basis elements for \( m = 1, \ldots, 5 \).
1.2. Mathematical descriptions of aspherical and freeform surfaces

1.2.1.2.2  Mild asphere

For mild aspheres, deviation from the best-fit sphere is constrained. This means that the transverse slope of the deviation between the surface and its best-fit sphere, when the deviation is measured along the normal direction to the surface, is limited [20]. For mild aspheres, the conic part is taken as the best-fit sphere. This could be defined in different ways. Here, the one that coincide with the surface at its axial point and around its perimeter is considered (figure 1.6). The radius of curvature is taken as:

\[ R = \left( r_{\text{max}}^2 + f(r_{\text{max}}) / (2f(r_{\text{max}})) \right) . \]

Since the conic section is taken as the best-fit sphere, \( \kappa = 0 \). Forbes formulation of mild asphere is given in eq.(1.9).

\[ z(r) = \frac{r^2}{R(1 + \sqrt{1 - r^2/R^2})} + D_{\text{bfs}}(\frac{r}{r_{\text{max}}}) \]  

(1.9)

where \( D_{\text{bfs}} \) is expressed in eq.(1.10).

\[ D_{\text{bfs}}(u) = \frac{u^2(1 - u^2)}{\sqrt{1 - \frac{u^2}{R^2}}u^2} \sum_{m=0}^{M} a_m Q_{\text{bfs}}^m(u^2) \]  

(1.10)

\( D_{\text{bfs}} \) is taken to be zero at \( u = 0 \) and \( u = 1 \). In this way, the coefficients values have no impact on the best-fit sphere. The cosine of the angle between the optical axis and the
local normal to the best-fit sphere expressed by $\sqrt{1 - \frac{r_{\text{max}}^2}{R^2} u^2}$ could be used to convert $D_{bfs}$ to a deviation measured along the surface normal (to first order). $Q_{bfs}^m$ are chosen in a way such that the weighted root-mean-square slope of the normal is just the sum of the squares of $(a_m)_{1 \leq i \leq M}$. After calculations, $Q_{bfs}^m$ are expressed as in eq. (1.11). These expressions were included in the amendment of the ISO 10110-12:2007 standard that was made in 2013 [15].

$$
\begin{align*}
Q_{0}^{bfs}(x) & = 1 \\
Q_{1}^{bfs}(x) & = \frac{13 - 16x}{\sqrt{19}} \\
Q_{2}^{bfs}(x) & = \frac{\sqrt{2}(29 - 4x(25 - 19x))}{\sqrt{95}} \\
Q_{3}^{bfs}(x) & = \frac{\sqrt{2}(207 - 4x(315 - x(577 - 320x)))}{\sqrt{2545}}
\end{align*}
$$

(1.11)

### 1.2.2 Mathematical descriptions of freeform surfaces in optics

Similar to aspherics, freeform shapes in optics are described by the mean of a basic shape (a conic or a biconic) to which high order terms are added. The second term,
1.2. Mathematical descriptions of aspherical and freeform surfaces

which contains the freeform contribution of the shape, could be expressed using a sum over a polynomial expansion set with a prefactor as shown in eq.(1.12) [14].

\[ z(x, y) = C(x, y) + P(x, y) \sum F(\bar{x}, \bar{y}) \] 

(1.12)

where \( C \) represents the basic shape, \( F \) is the freeform contribution and \( P \) is the prefactor. \( P \) is taken as in eq.(1.13) such that \( B \) is defined as the boundary function that controls the values of the deformation terms on the boundary line of the surface and \( H \) is a correction term to orient the additional sag correction along the local normal to the surface instead of the \( z \)-axis. \( \bar{x}, \bar{y} \) are the normalised coordinates defined as: \( \bar{x} = \frac{x}{x_{max}}, \bar{y} = \frac{y}{y_{max}} \) and \( \bar{r} = \sqrt{\bar{x}^2 + \bar{y}^2} \). Table 1.2 gives typical boundaries functions in freeform surface description.

\[ P(x, y) = \frac{B(\bar{x}, \bar{y})}{H(x, y)} \] 

(1.13)

<table>
<thead>
<tr>
<th>Boundary function</th>
<th>Description</th>
<th>Layout</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Uniform, no special constraints (unit circle/square)</td>
<td><img src="image" alt="Uniform, no special constraints" /></td>
</tr>
<tr>
<td>((1 - \bar{r}^2)\bar{r}^2)</td>
<td>Centre and boundary forced to be zero (unit circle)</td>
<td><img src="image" alt="Centre and boundary forced to be zero" /></td>
</tr>
<tr>
<td>((\bar{x}^2 + \bar{y}^2))</td>
<td>Centre forced to be zero (unit square)</td>
<td><img src="image" alt="Centre forced to be zero" /></td>
</tr>
</tbody>
</table>

Table 1.2: Typical boundary functions in freeform surface descriptions [14]

There exist a number of ways to describe the freeform contribution \( F \). The choice of the mathematical formulation depends on a number of parameters that are defined during the design stage such as the geometry of the supported area (circular or rectangular), \( z \)-orthogonal or slope orthogonal, Cartesian or polar coordinates, etc. However, orthogonality is the most desired characteristic when it comes to optimisation performances [21]. To each of these representations, a weighting function is added. This enables,
among others, to define the scalar product used for orthogonality determination.

Monomials is the simplest way to surface representation. They are expressed as Taylor expansion in $x$ and $y$ with no orthogonality. Chebyshev polynomials (in 2D) are taken as the Cartesian products of 1D Chebyshev polynomials [19]. There exist two variants of Chebyshev 2D: First kind and second kind [19]. They have similar recurrence definitions with a single difference in the definition of the second term. Chebyshev 2D polynomials are, however, spatially orthogonal on a unit square. Legendre 2D polynomials go in the same way as Chebyshev 2D by taking 1D Legendre polynomials instead of Chebyshev [22]. Zernike polynomials are defined using polar coordinates and are spatially orthogonal. This family of polynomials is widespread in optical design field since it is directly linked to wave-front errors and aberration [23]. Q-polynomials are defined based on the model developed by G. Forbes for mild aspherical surfaces. This formulation is gradient orthogonal and is intended to generate surfaces that are more adapted to manufacturing machines [24]. Table 1.3 (resp. Table 1.4) gives a summary of the different polar-defined (resp. Cartesian-defined) representations.

<table>
<thead>
<tr>
<th>Surface representation</th>
<th>Domain</th>
<th>Orthogonality</th>
<th>Boundary function</th>
<th>Graphical representation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Zernike Fringe</td>
<td>Unit circle</td>
<td>Spatial</td>
<td>1</td>
<td><img src="image" alt="Zernike Fringe Graphical Representation" /></td>
</tr>
<tr>
<td>Q-polynomials</td>
<td>Unit circle</td>
<td>Gradient</td>
<td>$(1 - r^2)^2$</td>
<td><img src="image" alt="Q-polynomials Graphical Representation" /></td>
</tr>
</tbody>
</table>

Table 1.3: Overview of polar-defined freeform surface representations [14]
### Table 1.4: Overview of Cartesian-defined freeform surface representations [14]

<table>
<thead>
<tr>
<th>Surface representation</th>
<th>Domain</th>
<th>Orthogonality</th>
<th>Boundary function</th>
<th>Graphical representation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Monomials</td>
<td>Arbitrary</td>
<td>None</td>
<td>None</td>
<td></td>
</tr>
<tr>
<td>Chebyshev 2D (first kind)</td>
<td>Unit square</td>
<td>Spatial</td>
<td>1</td>
<td><img src="image" alt="Graph" /></td>
</tr>
<tr>
<td>Chebyshev 2D (second kind)</td>
<td>Unit square</td>
<td>Spatial</td>
<td>1</td>
<td><img src="image" alt="Graph" /></td>
</tr>
<tr>
<td>Legendre</td>
<td>Unit square</td>
<td>Spatial</td>
<td>1</td>
<td><img src="image" alt="Graph" /></td>
</tr>
</tbody>
</table>
1.3 Manufacturing of aspherical and freeform surfaces

Regarding the wide field of applications of aspherical and freeform surfaces, different requirements on the quality and quantities of the artefact to produce come in role. Hence, the manufacturing process must be chosen accordingly. In figure 1.7, a classification of optical surfaces regarding the quality, the quantities to produce, the cost as well as the production methods is given.

![Figure 1.7: Classification of optical components regarding quality, quantity, price and production methods [25]](image)

A classical manufacturing process within the whole lifecycle of a freeform surface is depicted in figure 1.8. Various manufacturing techniques could be during the component manufacture stage.

In [11], Fang et al. give a comprehensive overview of the different manufacturing techniques for aspherical and freeform optics. The major techniques discussed in this section are: Single Point Diamond Turning (SPDT), Slow and Fast Servo Tools, diamond milling, fly Cutting, MRF, Precision Glass Molding (PGM) and Molded Polymer Asphere (MPA).

SPDT is mainly used for the manufacturing of rotationally symmetric surfaces made of non-ferrous metals or ceramics. It could be used for the manufacturing of freeform surfaces as well provided that an adaptation is applied [26]. This process was initiated by the need of large size lenses with high quality surfaces. A surface roughness below $10 \, nm$ could be achieved using this process [27].
1.3. Manufacturing of aspherical and freeform surfaces

The process of SPDT machining is performed in different steps. In the first ones, the manufacturing is done by the mean of CNC machines with increasing accuracy. The finishing of the machine is done by very accurately cutting away a thin chip or layer of the surface using a diamond tipped tool. The final workpiece does not need traditional polishing operation [28]. Figure 1.9 shows the PreciTech Optimum 2400 machine using SPDT process from Apollo Optical Systems.

Slow Tool Servo (STS) and Fast Tool Servo (FTS) are two improvements that could be brought to SPDT. In the STS, the contact of the tool tip with the manufactured part is intermittent since the z-axis oscillates during the manufacturing process. Very accurate asymmetrical surfaces with surface roughness less than 10 nm could be achieved using STS [29]. A wide range of optical parts could be manufactured using STS. This
includes, but not limited to, off-axis aspheric surfaces, prism array, micro lens arrays, etc.

FTS has the same configuration as STS with an additional actuator responsible of oscillating the tool tip. The tool tip can move in a range less than 1 \( \mu m \) and in some cases up tp 1 \( mm \) [30]. With machining speed higher than STS, FTS could achieve surface roughness less than 10 \( nm \) and could be used for the fabrication of lens arrays or diamond-turned surfaces with micro prisms structure [11].

Diamond milling is another material removal technique that could be used for the fabrication of aspherical and freeform surfaces. Here, the milling tool is rotated instead of the workpiece with speeds that might exceed 100000 RPM. The diamond milling process is slower than diamond turning but offers more flexibility in manufacturing. As for the presented processes, diamond milling could achieve surface quality with roughness less than 10 \( nm \) and could be used essentially for the fabrication of micro lens arrays or non-smooth surfaces where turning processes are not applicable [31].

In fly cutting process, the diamond is placed off-axis on a rotating tool. In this way, the diamond is not permanently in contact with the workpiece. Fly cutting is mainly used for the fabrication of micro structures and freeform surfaces with surface quality in the nanometre range [32]. Figure 1.10 illustrates the different manufacturing processes previously discussed.

PGM is an alternative to the classical manufacturing methods. The main advantage of this technique is its ability to produce large volume of lenses with minimal surface deviation. Also, the production time for an aspheric using PGM technique take eight to fifteen minutes which is shorter than the required time for the previous techniques [33]. Optical glass cores with listed moldability properties are heated to a predefined temperature and then pressed into an aspheric mold. A description of the PGM process is given in figure 1.11. The quality of the final workpiece depends on a number of parameters such as the mold material selection, the mold quality, the glass properties, etc [34]. The MPA is a process which is close to the PGM. Instead of using optical glass cores, a standard spherical lens is pressed into a thin layer of polymer placed inside an aspherical mold (figure 1.12).

An original approach using material removing for processing optical lenses with magnetic fluids called MRF is used to manufacture freeform optics. This technology was initiated by Willian Kordonski in 1990 [37]. The fundamental element of MRF is MR
1.3. Manufacturing of aspherical and freeform surfaces

Figure 1.10: Ultra-precision machining (UPM) methods. (a) Diamond turning, (b) STS/FTS, (c) Diamond milling, (d) Fly cutting [31]

Figure 1.11: PGM polishing fluid. It consists of a liquid composition that undergoes a change in mechanical properties in presence of magnetic fluid (figure 1.13). MR fluid contains very small ferromagnetic particles (0.1 μm) that are organised into chains of particles, forming then a spatial structure resulting in a change in mechanical properties. Without the magnetic field, the particles return progressively to a disorganised state and the initial condition of the overall material is restored.

In theory, MR polishing fluid contains four main constituents: magnetic particles, polishing abrasive, chemical additive and water. Nevertheless, water is almost used as a carrier fluid for polishing glasses and silicon substrates without any additional chemical agent.
Figures 1.14 and 1.15 illustrate the Q22 multiple axis-computer controlled MRF machine. The lens is fixed such as a converging gap can be formed between the lens and the rotating spherical wheel. MR polishing fluid is loaded into the closed-loop fluid delivery system, where fluid properties such as temperature and viscosity can be continually monitored.

From the conditioner, the fluid is driven in a thin ribbon in contact with the optical surface and then removed by a suction cup and fed back into the conditioner. A strong local electromagnetic field gradient is generated by the electromagnet located below the polishing wheel. The mechanical properties of the MR fluid change in the presence of...
1.3. Manufacturing of aspherical and freeform surfaces

Figure 1.14: Manufacturing of an aspherical lens using a MR process

Figure 1.15: Principle of the Q22 multiple-axis computer-controlled MRF machine

This magnetic field. It stiffens in milliseconds and then returns to its original fluid state at it leaves the field, again in milliseconds. The precisely controlled zone of the MR fluid that stiffened becomes the polishing tool. When the optical surface is placed into the fluid, the stiffened fluid ribbon is squeezed from its original thickness, which results in significant shear stress and subsequent polishing pressure over that section of the optical surface [38].
3D-printing is a newly used process for the fabrication of aspherical and freeform optics. In [39], a method based on modified ink-jet printing was used for manufacturing imaging-quality optics with a surface profile deviation of $\pm 500 \text{ nm}$ within a $12 - \text{mm}$ aperture diameter. The design of the surface to manufacture is converted to a printable CAD model and then sliced into a number of 2D layers. Following these layers, liquid polymer droplets are successively deposited on a substrate. A smooth surface is formed after the droplets are merged. This process is described in figure 1.16.

Figure 1.16: Modified ink-jet 3D printing process. (a) Printhead depositing polymer on the substrate, (b) layer-by-layer deposition, (c) forming of lens shape and (d) smooth lens surface after depositing extra droplets [40]

### 1.4 Measurement of aspherical and freeform surfaces

In this section, the major techniques for the metrology of aspherical and freeform surfaces are briefly discussed. Detailed descriptions of some measuring machines will be provided in Chapter 4. These techniques are namely ultra-high precision Coordinate Measuring Machines (uhp-CMMs), deflectometry and interferometry.

Uhp-CMMs are the most widely used in industry when inspecting the geometrical accuracy of freeform surfaces [41]. They exist in different configurations namely, bridge, cantilever, gantry and horizontal. A description of each type is given in the ISO 13060-1:2000 [42]. Coordinate Measuring Machines (CMMs), which could be considered as a displacement system that can move a probe (tactile or optical), are used to generate a set of data points lying on the surface to measure. The coordinates of the recorded data reflect the positions of the probe expressed in a frame associated to the machine. The
position of the probe along each axis of the machine is determined using sensors mounted on each axis. It could be tracked with micrometer accuracy for usual applications and nanometric accuracy for ultra-high precision measurements [43]. Tactile or optical probing systems could be mounted on a CMM. The scanning speed of CMMs depends on the embedded probing system. Tactile probing systems need more scanning time than optical ones. However, collected data are noisy using the latter. Figure 1.17 shows a measurement setup of a lens using a CMM. Profilometry which is similar to CMMs based metrology is destined to the measurement of roughness.

![Figure 1.17: Measurement of an optical lens using CMM [44]](image)

Deflectometry and interferometry are non-contact measurement techniques. They allow the measurement of the entire surface, with low uncertainty, in a single measurement. However, the measurement quality is highly affected by the environmental conditions and the interaction of the beam with the surface [11].

The idea behind deflectometry is to measure the amount of beam deflection induced by the interaction with the specimen under test [45]. It could be estimated by the determination of the deflection of a light ray with a known source. In this way, only the deflection at a given point will be known. This drawback was rectified with the usage of a technique called Phase Measuring Deflectometry (PMD). This methods consists of displaying different sinusoidal gratings using a spatial light moderator on the artefact. The reflected light is then recorded by a calibrated camera and reverse operations are applied in order to characterise the artefact. The principle of deflectometry is shown in
Chapter 1. Metrology of aspherical and freeform surfaces

Figure 1.18: Principle of deflectometry [46]

Interferometry is based on the superposition of two or more waves in order to create an output wave that differs from the input ones. This phenomenon is called interference. The power of the spatial shape of the resulting wave can be used for measurement [47]. A Michelson interferometer is one of the most widely used configurations of interferometers. It consists of splitting a beam of monochromatic light into two equal amplitude beams. One beam is reflected on a fixed mirror and the other on a moving one. The two reflected beams, having two different lengths are projected on a detector where the interference pattern is observed [48, 49, 50]. Figure 1.19 shows the working principle of the Michelson interferometer.

A number of techniques such as Moiré Interferometry, Sub-Aperture Interferometry, Computer Generated Hologram (CGH) interferometry, White Light Interferometry (WLI), etc. make use of the Michelson or other configurations of interferometers such as Fizeau interferometer, Twyman-Green interferometer, etc. We refer to [52] where the description of a number of other measurements techniques are described. In [41], Savio et al. provided an evaluation of some measuring techniques based on the dimensions of the part to measure, its complexity and other criteria.

In order to achieve measurements with nanometric uncertainty, a combination of
the previous techniques (CMMs including the probing system and interferometers) is necessary. The main principle of these machines is using laser interferometers in each axis in order to track the position of the probe while respecting the Abbe principle [53]. This was the case of a number of ultra-high precision measuring machines that were recently developed such as the ISARA 400 [54], the Nanomefos [7], the LNE’s profilometer [55], the METAS µCMM measuring machines [56], etc. [49].

1.5 Conclusion

This chapter gives a general overview on aspherical and freeform surfaces. The need and relevance of such surfaces in industry was demonstrated. Mathematical descriptions for aspherical and freeform surfaces was also discussed. A more attention was given to the ISO 10110-12:2007 and the Forbes formulations in the case of aspherical surfaces. A general approach based on a number of polynomial sets with different properties was presented in the case of freeform surfaces.

Manufacturing and measurement of aspherical and freeform surfaces was also the subject of this chapter. Some of the main manufacturing and metrology techniques were briefly described. This includes techniques such as SPDT, diamond turning, fly
Chapter 1. Metrology of aspherical and freeform surfaces

cutting, PGM and MPA. A short description of three of the main techniques used in metrology was given. This concerns mainly CMMs based techniques, deflectometry and interferometry.
Chapter 2

Fitting algorithms of aspherical and freeform surfaces
Chapter 2. Fitting algorithms of aspherical and freeform surfaces

2.1 Introduction

During the design process of parts, designers specify form error limits according to the functional requirements. In the quality control stage, conformance of manufactured parts to design tolerance specifications is verified. The aim is to assess the variability of the manufactured part compared to a perfect shape. If \( V(\Sigma) \geq 0 \) denotes a function that represents the variability of a given shape \( (\Sigma) \) and \( \epsilon \) is the design tolerance, the part will be accepted if \( V(\Sigma) \leq \epsilon \) and rejected otherwise [57, 58]. Intrinsic geometrical deviations from the desired form could be classified as form errors, waviness and roughness. These errors need to be separated in order to be assessed [58].

Despite advancements made in precision engineering, three essential issues still arise and slow CMM-based metrology approaches [59]:

- The development of suitable measurement techniques ensuring that collected data points represent the measured part with a very low uncertainty,
- A lack of correct interpretations of tolerance definitions as stated in the ISO Geometrical and Product Specification standards (GPS),
- The development of accurate and robust metrology algorithms conformant with the ISO GPS.

ISO Geometrical and Product Specification standards specify the general framework for form tolerance specifications [60]. These rules depend on the form of the inspected artefact. Table 2.1 gives a list of standards providing vocabulary and parameters for form and profile tolerance specifications. For the case of straightness (resp. flatness), form error is defined as the distance between two lines (resp. planes) containing all measured data points and having the least separation [61]. For roundness and cylindricity, form error is determined by the radial distance between two concentric elements containing all data points and having the least radial separation [62]. However, these definitions could not be transposed to the case of a line or a surface profile since the bounding shapes of a profile are not similar. This problem arises when inspecting assessment methods for elliptical forms. In [63], Murthy et.al demonstrated that the bounding shapes of an elliptic profile with tolerance \( t \) are not ellipses with semi major and minor diameters that are increased with \( t/2 \). Besides, the bounding shapes are not even ellipses.
2.1. Introduction

A conceptual definition of form tolerance in the case of line/surface profile could be given as the smallest value of the diameter of a circle/sphere whose centre is moving along the nominal shape such as all data points are included in the space covered by the moving circle/sphere (figure 2.1) [66].

Form error evaluation requires the fitting of measured data to a nominal shape. Fitting could be defined as the process of determining parameters of the geometric features that best describe the measured data according to a defined criterion. The obtained geometry is called the “associated feature” [67]. There exist a number of criteria that could be considered, including but not limited to LS (also called Gaussian or $L_2$ fitting) with or without external/internal constraints, minimax fitting (also called Chebyshev fitting) with or without external/internal constraints as well as one sided measures such as Minimum Circumscribed (MC) or Maximum Inscribed elements (MI). While the two latter criteria are suitable for circular or cylindrical shapes, the former ones could be applied for almost any shape. The choice of the criterion is based on the adopted variability function $\mathcal{V}$.

This chapter is organised around fitting algorithms for aspherical and freeform surfaces. In the next section, LS and MZ fitting criteria are presented and the resulting minimisation problems are compared. Section 2.3 gives a general overview of the existing methods for the resolution of the MZ fitting problem. Two methods were considered for a comparison in Section 2.4. The comparison was conducted based on reference data as well as measured data. In Section 2.5, a new fitting algorithm with better performances

---

<table>
<thead>
<tr>
<th>Tolerances</th>
<th>Symbol</th>
<th>Standard</th>
</tr>
</thead>
<tbody>
<tr>
<td>Straightness</td>
<td>-</td>
<td>ISO 12780-1:2011 [61]</td>
</tr>
<tr>
<td>Flatness</td>
<td>□</td>
<td>ISO 12781-1:2011 [64]</td>
</tr>
<tr>
<td>Roundness</td>
<td>○</td>
<td>ISO 12181-1:2011 [62]</td>
</tr>
<tr>
<td>Cylindricity</td>
<td>≪</td>
<td>ISO 12180-1:2011 [65]</td>
</tr>
<tr>
<td>Line profile</td>
<td>⸤</td>
<td>ISO 1660:2017 [66]</td>
</tr>
<tr>
<td>Surface profile</td>
<td>○</td>
<td>ISO 1660:2017</td>
</tr>
</tbody>
</table>

Table 2.1: Form tolerance specifications: symbols and standards

---

\footnote{The associated feature could be seen as a nominal shape that verifies the fitting criterion.}
is suggested and validated by the mean of reference data, benchmark data and measured data. In the last section, a method is described for the estimation of the uncertainty on the returned form error values.

![Figure 2.1: Form tolerance definition for: (a) line profile, (b) surface profile](image)

**2.2 Form error assessment**

The ISO GPS standards indicate that the specifications are applied to the indicated extracted integral feature (a geometrical feature\(^2\) consisting of a finite number of points)

\(^2\)According to the ISO 22432:2011 [67], a geometrical feature is a point, line, surface, volume or a set of these previous items. The non-ideal surface model is a particular geometrical feature, corresponding to the infinite set of points defining the interface between the workpiece and the surrounding.
or derived feature itself (a geometrical feature, which is a median, displaced, congruent or reflected feature resulting from a set of operations on an integral or filtered feature). It is also mentioned that if a filter is indicated, the fitting shall be to the filtered feature (a non-ideal feature which is the result of a filtration of a non-ideal feature). However, it is highly recommended to separate waviness roughness and form error before proceeding to the fitting. In the following, a short overview on these three components are given and then the difference between LS and MZ fitting criteria is discussed.

### 2.2.1 Waviness, roughness and form error

There exist numerous factors that result in irregularities in the manufactured surfaces such as the tool, the machined geometry, the environmental conditions, etc [68]. These irregularities have different frequencies and could be classified accordingly. The high frequencies or short wavelength components are referred to as roughness, the medium frequencies as waviness and low frequencies as form (figure 2.2) [69]. Each manufacturing process has a different wavelength regime. Reciprocally, identifying the different spectrum in the surface profile indicates about the manufacturing process which could be a useful tool for process control and diagnostics [70].

![Figure 2.2: Form error, waviness and roughness](image)

Despite being considered as separate wavelength regimes, the boundaries between roughness, waviness and form error are ambiguous. These boundaries could be defined according to the specific application. A number of standards exist for the metrological requirements and guidelines associated with each regime and corresponding to the
considered application. Few of them are listed in table 2.2 [72].

<table>
<thead>
<tr>
<th>Standard</th>
<th>Roughness</th>
<th>Waviness</th>
<th>Form Error</th>
</tr>
</thead>
<tbody>
<tr>
<td>ISO 3274:1998 [73]</td>
<td>X</td>
<td></td>
<td></td>
</tr>
<tr>
<td>ISO/NP 21920-2 [76]</td>
<td>X</td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>ISO 13565:1998 [77]</td>
<td>X</td>
<td></td>
<td></td>
</tr>
<tr>
<td>ISO 1101:2017 [60]</td>
<td>X</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 2.2: Form tolerance specifications: symbols and standards

In order to separate form error, waviness and roughness from a measurement, filtering methods such as Gaussian filtering or discrete modal decomposition methods are used. Three setting parameters should be previously known before proceeding to filtering namely the cutting frequency at which the three components are separated, the sharpness of the filter referring to how clean the filter separates two regimes and the distortion of the filter [75]. In the following, only form error components are considered and the given data sets are supposed already filtered. The following subsection gives the mathematical formulation of the two fitting problems namely MZ and LS.

2.2.2 Minimum Zone vs. Least Squares

LS and MZ are the two widely used criteria in data fitting. While the former criterion, which originates from the maximum likelihood theory, is more adapted when random measurement errors predominate, the latter is appropriate when measurement errors are small compared to manufacturing ones since the estimated form error using this method is highly sensitive to measuring errors [78].

Suppose given are a set of $N$ measured points $(p_i)_{1 \leq i \leq N}$ and let $S : F(\hat{X}, s) = 0$ be the generic equation describing the nominal shape of the measured artefact where $\hat{X}$ is the space vector $(\hat{x}, \hat{y}, \hat{z})$ in Cartesian coordinates and $s$ denotes the shape parameters. To each of the measured points $p_i$ we associate its orthogonal projection on the surface $S$, labelled $q_i$. We define also the form deviation $d_i = (-1)^{r_i}||p_i - q_i||$, where $||.||$ is the Euclidean norm, $r_i = 0$ if the point $p_i$ lies above the nominal shape and $r_i = 1$.
2.2. Form error assessment

if the point $p_i$ is below, $d_i$ is considered as the directed orthogonal distance between $p_i$ and the surface $S$ (figure 2.3). It is to be noted that $d_i$ depends on $x = (s, m) \in \mathbb{R}^n$ where $m$ are transformation parameters: rotation and translation applied to $(p_i)$.

![Figure 2.3: Definition of form deviation. Measured data are supposed to lie exactly on the manufactured shape](image)

To derive the LS formulation, we suppose that form deviations $(d_i)_{1 \leq i \leq N}$ are independent and identically distributed random variables and that each $d_i$ follows a Gaussian distribution $\mathcal{N}(0, \sigma^2)$ with mean 0 and variance $\sigma^2$ (this is why LS fitting is often called Gaussian fitting). The probability density function in this case is given as in eq.(2.1).

$$\phi(y) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{1}{2} \left(\frac{y}{\sigma}\right)^2\right)$$

(2.1)

The maximum likelihood function associated with the set of form deviations $(d_i)_{1 \leq i \leq N}$ is given in eq.(2.2) [79].

$$\mathcal{L}(d_1, ..., d_N, x) = \prod_{i=1}^{N} \phi(d_i) = \prod_{i=1}^{N} \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{1}{2} \left(\frac{d_i(x)}{\sigma}\right)^2\right)$$

(2.2)

Taking the logarithm of the maximum likelihood function results in eq.(2.3).

$$\log(\mathcal{L}(d_1, ..., d_N, x)) = N\log\left(\frac{1}{\sigma\sqrt{2\pi}}\right) - \frac{1}{2\sigma^2} \sum_{i=1}^{N} d_i^2(x)$$

(2.3)

From eq.(2.3) we conclude that maximising the maximum likelihood function is equivalent to minimising the second term in eq.(2.3) since the first term does not depend on $x$. Hence, we get the formulation of the problem to solve in the case of LS fitting presented in eq.(2.4).
Chapter 2. Fitting algorithms of aspherical and freeform surfaces

\[ \min_x \Phi_{LS} \quad \text{with} \quad \Phi_{LS} = \sum_{i=1}^{N} d_i^2(x) = ||d||^2_2 \quad \text{and} \quad d = (d_1, \ldots, d_N) \quad \text{(2.4)} \]

The same reasoning could be followed in order to derive the objective function for the case of MZ fitting by replacing the Gaussian distribution function with a uniform one with parameters \( U(d_{\text{min}}, d_{\text{max}}) \) and the resulting minimisation problem to solve is given in eq.(2.5).

\[ \min_x \Phi_{MZ} = \max_{1 \leq i \leq N} d_i(x) - \min_{1 \leq i \leq N} d_i(x) \quad \text{(2.5)} \]

If we consider symmetric tolerances only, the MZ given in eq.(2.5) is equivalent to minimising the maximum absolute value deviations; which is equivalent to minimising the infinite norm (also called the Chebyshev norm or the uniform norm) of form deviations vector \( d = (d_1, \ldots, d_N) \) [80]. This results in the Chebyshev fitting problem given in eq.(2.6).

\[ \min_x \Phi'_{MZ} = \max_{1 \leq i \leq N} |d_i(x)| = ||d||_\infty \quad \text{(2.6)} \]

Unlike LS, the objective function in MZ fitting is not smooth and so a number of derivative-based techniques could not be used to solve the problem in eq.(2.6). Moreover, the objective function depends only on few measured data so the returned value of MZ could be highly affected by outliers resulting from measurement errors. A pre-processing step consisting of cleaning the set of measured data is essential before performing MZ fitting. In the following, we suppose that the data to be fitted have been already cleaned.

The two objective functions could be used according to the adopted variability function. If the root of mean squares (RMS) is considered, the LS fitting is more appropriate. The MZ fitting is used when the least value of the peak-to-valley (PV) is sought. The PV is defined as the difference between maximum and minimum form deviations. In the case of form error estimation, the PV is considered as the variability function. Nevertheless, it happens that the LS fitting is used even so. The main reason behind this choice is the simplicity of solving the LS fitting problem compared to the MZ one as well as the uniqueness of the obtained solution. The drawback of this approach is that the calculated value of the true MZ is overestimated which results in the rejection of a number of conforming parts (figure 2.4). The MZ value found using a Chebyshev fitting has been found to be around 80% of the value obtained by LS [81].
Figure 2.4: Estimated MZ values obtained using MZ fitting (blue) and LS (orange). The value obtained using LS is overestimated.

2.3 Literature review

A number of methods were developed for MZ fitting. Here, a comprehensive state of the art was conducted for both canonical and freeform shapes.

2.3.1 Canonical forms

MZ evaluation for the case of straightness, flatness, roundness and cylindricity was the subject of an extensive research and a spectrum of methods were developed for this aim. In [82], the strengths and weaknesses of a wide range of techniques from the literature were discussed.

Computational geometry techniques are a set of methods that aim at selecting critical points (also called significant points) defining the MZ, based on geometrical considerations as fast as possible and ignoring the remaining points. In most cases, this is performed by targeting the convex hull of the measured points and the applying simple geometrical operations to find the critical points. These techniques were applied in [83] for the case of straightness where a concept called EigenPolyGon (EPG) is described and its relationship with straightness is derived. In [84] a method called the ‘convex-hull edge method’ (CONHEM) is proposed for the case of flatness. Another approach based on computational geometry through Voronoi diagrams for the determination of MZ in the case of circularity is presented in [85]. Computational Geometry techniques were

**MZ** : Exact value of minimum zone (obtained using minimum zone fitting)

**MZ\_LS** : Estimated value of minimum zone (using least squares fitting)

\[
MZ \leq MZ\_LS
\]
also used to determine sphericity in [86] based on the construction of the 3D inner and outer convex hull. Other methods, also using the convex hull approach, were adapted simultaneously to the cases of straightness and flatness as in [87]. In [88], a technique called Control Line/Plane Rotation Scheme (CLRS/CPRS) was presented. This technique was extended for roundness and cylindricity tolerances. Other methods are given in [89, 90, 91] for roundness, in [92] for sphericity or in [93] for cylindricity tolerance. The main advantage of computational geometry techniques is that the geometrical aspects of the method could be visualised much better than the mathematical aspects of numerical techniques [94]. This transparency makes the algorithms easily accepted by inspectors. They represent also the advantage of returning accurate results despite being computationally intensive. Nevertheless, these techniques are restricted to the case of simple geometries and could not be extended to freeform shapes.

In [81] three numerical methods named Monte Carlo method, simplex search technique and spiral search technique are suggested for the case straightness, flatness, circularity and cylindricity. In the first method, the MZ surface is assumed to lie close to the LS mean surface. A number of random surface parameters are generated in the neighbourhood of the LS mean surface and used to find a smaller MZ. This method represents poor performance compared to other existing methods since the chances of missing the exact MZ value are not ruled. Moreover, due to the curse of dimensionality, this method is not appropriate when the number of variables are high. The simplex search technique gets its name from the regular geometric figures used in the search process. It is a variant of the Nelder-Mead algorithm which is sequential gradient search designed to climb up and down mathematical hills based on punctual evaluations of the objective function and applying some transformations on the simplex such as reflection, expansion, contraction, etc. [95]. The spiral search technique is similar to Monte Carlo method. Instead of performing the straight search using random variables around the LS, the search is performed by following a spiral path around the LS solution. In [96], the simplex search is compared to a bracketing method which is a multidimensional optimisation technique that makes use of a one-dimensional search direction in the case of flatness. Results show that the downhill simplex method is remarkably advantageous from the viewpoint of the number of iterative calculations and calculating time. A similar approach to the simplex search for the case of straightness, flatness and roundness was suggested in [97] and for sphericity in [98] and [99].
Another class of methods transforms the minimisation problem into a constrained linear/nonlinear problem by introducing some approximations. Classical minimisation techniques are then used to solve the resulting problem \([96, 100]\) for the case of straightness and flatness respectively. A linear search method with quadratic interpolation called (QIM) as well as Golden Section Minimisation method (GSM) are suggested in \([101]\) for the case of straightness. In \([102]\), the straightness and the flatness problems are formulated as nonlinear optimisation problems with a linear objective function and nonlinear constraints. A developed linear search method is used to transform the nonlinear problem to linear programming with only two constraints. In \([103]\), MZ problem for roundness tolerance is formulated as a nonlinear optimisation problem. Necessary optimality conditions are stated and then an exchange method is applied to calculate the global optimal solution. The MZ fitting problem could also be treated as an unconstrained minimisation problem as in \([104]\) where the Hooke-Jeeve direct search method is used for the case of roundness.

Heuristic methods such as genetic algorithms were used in \([105, 106, 107, 108, 109, 110]\). These algorithms use techniques that are based on the principles of natural selection and genetics in order to calculate the optimal solution. They have the advantage of being derivative-free methods so they could be applied for minimising non-differentiable objective functions which is the case in MZ fitting. Implementation of genetic algorithms is simple and the convergence to a global solution is guaranteed. However, the computational efficiency is not compared to the methods already discussed. Moreover, the heuristic aspect, resulting in obtaining a different solution for the same input each time the algorithm is run, makes it a second choice for users.

Learning methods such as Support Vector Machine (SVM) were also used for the determination of MZ for straightness, cylindricity and sphericity. SVM is a supervised learning technique that could be used for both classification and regression problems. This method was adapted for the case of sphericity tolerance in \([111]\) and the results were compared to linear LS as well as linear and nonlinear optimisation techniques. Results show that the use of support vector regression methods is quite well justified. For the case of straightness and cylindricity refer to \([112, 113]\).
2.3.2 FreeForm and aspherical shapes

The developed methods for MZ fitting of aspherical and freeform surfaces is not as abundant as for canonical shapes and little could be found about this topic. Nevertheless, some attempts were made to address this problem. In [114, 115], the infinite norm is approximated via the $L_p$-norm by gradually increasing the value of $p$. This approach results in a smooth objective function and hence derivative based methods could be used to minimise it. However, when the value of $p$ increases, numerical instabilities rise and the system becomes highly-ill conditioned which limits the performance of these methods.

As for the case of canonical shapes, derivative-free based methods such as simplex search could be used for MZ determination but these methods prove to be inefficient due to the higher number of unknowns as well as the nonlinearities of these complex shapes. On the other hand, genetic algorithms were developed for MZ fitting of complex surfaces. In [116] a heuristic optimisation method called Differential Evolution (DE) was presented. For large cloud of points (>100,000) this approach is time consuming which explains the development of an additional procedure based on alpha-shape to select critical points. This enables to reduce calculation time by over 97%. Nevertheless, the selection of ball radius used during alpha-shape determination is a crucial issue and no straightforward method exists for selecting an appropriate value. This could result in rejecting points defining the MZ region and hence returning an underestimated value.

Other methods aim at transforming the non-differentiable optimisation problem into a nonlinear programming and then applying adapted methods for solving the resulting optimisation problem. In [117], this approach was applied for form error determination of surfaces described using NURBS. Obtained results show that the methods yields a global optimum with a symmetric deviation map. Moreover, the tolerance zone is greatly reduced compared to the ordinary LS fitting.

Smoothing techniques have also been adopted, for example in [118] for fitting parametric curves and surfaces by $L_\infty$-norm regression. The method consists of approximating the initial MZ optimisation problem by recursive linearisation and then using conventional Gaussian-Newton algorithms to solve the resulting problems. A method called Exponential Penalty Function was suggested in [119] for the fitting of aspherical surfaces. This method is based on the formulation of a surrogate function that
converges to the original objective function when the smoothing parameter is large. A Gauss-Newton type method for fitting implicitly defined curves and surfaces was developed in [120, 121]. This is a generalisation the usual Gauss-Newton technique for a LS approximation by replacing the $L_2$ norm by a norm-like function that could be an approximation of $L_1$ or $L_\infty$.

### 2.3.3 Summary

Based on the established literature review, the performance of each class of methods is given for straightness, flatness, roundness, cylindricity and line/profile tolerances (table 2.3). The performance is estimated in function of the ability of the method to return accurate results as well as the execution time. This works shows that two classes of methods could be good candidates for MZ fitting namely smoothing techniques and constrained nonlinear problems. In the following section, two selected methods, each one belonging to each class, are implemented and compared using reference data and measured data. The method belonging to smoothing techniques (resp. constrained nonlinear problems) is called Exponential Penalty Function (EPF) (resp. Primal-Dual Interior Point method (PDIP)).

<table>
<thead>
<tr>
<th>Method</th>
<th>Tolerance</th>
<th>Straightness</th>
<th>Flatness</th>
<th>Roundness</th>
<th>Cylindricity</th>
<th>Line/Profile</th>
</tr>
</thead>
<tbody>
<tr>
<td>Computational geometry</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>geometry techniques</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Cannot</td>
</tr>
<tr>
<td>Simplex search</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>techniques</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>be used</td>
</tr>
<tr>
<td>Constrained linear problems</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Cannot</td>
</tr>
<tr>
<td>Smoothing techniques</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>be used</td>
</tr>
</tbody>
</table>
2.4 Form assessment using EPF and PDIP

2.4.1 Exponential Penalty Function (EPF)

Exponential penalty function is a smoothing technique that consists of approximating the non-smooth objective function in the MZ fitting problem by a parameterized continuously differentiable one. The resulting function is then optimised using Newton-based methods [119, 122]. By replacing the form deviations vector \( \vec{d} = (d_1, ..., d_N) \) in eq.(2.6) with \( \vec{d} = (\tilde{d}_1, ..., \tilde{d}_N) = (d_{11}^2, ..., d_{N1}^2) \), we get the equivalent MZ fitting problem presented in eq.(2.7).
2.4. Form assessment using EPF and PDIP

\[ \min_x \Phi''_{MZ} = \max_{1 \leq i \leq N} |d_i^2(x)| = ||\bar{d}||_\infty = F(x) \] (2.7)

Let \( F_p \) be the continuously differentiable approximation function and \( p > 0 \) the smoothing parameter as in eq.(2.8). One can easily prove the inequality given in eq.(2.9).

\[ F_p(x) = \frac{1}{p} \text{Log}(\sum_{i=1}^{N} \exp(pd_i(x))) \] (2.8)

\[ \forall x \in \mathbb{R}^n, F(x) \leq F_p(x) \leq F(x) + \frac{\text{Log}N}{p} \] (2.9)

When the value of \( p \) goes to infinity, \( F_p \) converges to \( F \). For a given value of the smoothing parameter \( p \), minimisation of \( F_p \) is carried out using derivative-based methods. Once the optimum of \( F_p \) is found, \( p \) is increased by multiplying it by a user-defined coefficient, and a new approximation function is formulated. This process is repeated until the resulting approximation function \( F_p \) becomes sufficiently close to the original objective function \( F \).

The use of Newton-based methods to minimise \( F_p \) requires the calculation of the Hessian matrix. Since calculation time of the Hessian matrix depends on the number of points in the data set, only points belonging to a subset of points defined in eq.(2.10), also called \( \epsilon \)-active set, where \( \epsilon \) is a user defined parameter, could be considered for the sake of computational efficiency. This represents the set of points that are likely to be the points defining the MZ.

\[ \Omega_p = \{ j | F_p(x) - \bar{d}_j(x) \leq \epsilon, 1 \leq j \leq N \} \] (2.10)

It should be noted that in the implemented version of the algorithm, the descent direction was found using the Newton method and the chosen step-length verifies the Wolfe conditions [123].

### 2.4.2 Primal-Dual Interior Point method (PDIP)

In the Primal dual Interior Point method, the unconstrained non-smooth MZ fitting problem is transformed into the nonlinear programming presented in eq.(2.11) by introducing the variable \( e \geq 0 \) [124, 125]. Primal dual interior points methods were originally developed for the case of linear programming then for convex programming and finally
were adapted to the case of non-convex nonlinear smooth constrained problems [123].

\[
\min_{x,e} e
\]

subject to \(-e \leq d_i(x) \leq e \quad \forall i \in \{1,...,N\}\) \hspace{1cm} (2.11)

Once the nonlinear programming is constructed, the Lagrangian of the problem is formulated and \(e\) is decreased by iteratively minimising the Lagrangian function until the first order optimality conditions, also called the Karush-Kuhn-Tucker (KKT) conditions, are satisfied [123].

A primal dual interior point method needs an initial feasible solution which is an "Interior Point" of the feasible domain (a solution that satisfies all constraints). This could be obtained by the resolution of an unconstrained LS problem [126].

### 2.4.3 Comparison of EPF and PDIP using reference data

The comparison of EPF and PDIP was performed using reference data as well as measured data. Reference data, also called Softgauges, are sets of data points for which the exact value of MZ is known with an associated uncertainty. The method for reference data generation (see Chapter 3 for details) enables the control of a number of parameters in the generated data sets including but not limited to, the density and distribution of the points, the value of the form error and the initial position of the data set.

The two algorithms were applied for the case of aspherical surfaces. Five configurations of nominal shape coefficients, supposed known, were considered (table 2.4). These configurations were selected in a way to test different shapes of aspherical surfaces covering flat as well as steep shapes (figure 2.5). The values of maximum aperture \(R_{\text{max}}\) was fixed to 10 mm for all tests since data is scaled before performing the fitting.
2.4. Form assessment using EPF and PDIP

<table>
<thead>
<tr>
<th>Config</th>
<th>$R$ (mm)</th>
<th>$k$</th>
<th>$a_4$ (mm$^{-3}$)</th>
<th>$a_6$ (mm$^{-5}$)</th>
<th>$a_8$ (mm$^{-7}$)</th>
<th>$a_{10}$ (mm$^{-9}$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>I</td>
<td>381.83</td>
<td>-1</td>
<td>$-4.51 \times 10^{-13}$</td>
<td>$-2.25 \times 10^{-13}$</td>
<td>$-1.12 \times 10^{-13}$</td>
<td>$-5.63 \times 10^{-14}$</td>
</tr>
<tr>
<td>II</td>
<td>76.18</td>
<td>-1</td>
<td>$-2.26 \times 10^{-12}$</td>
<td>$-1.13 \times 10^{-12}$</td>
<td>$-5.65 \times 10^{-13}$</td>
<td>$-2.82 \times 10^{-13}$</td>
</tr>
<tr>
<td>III</td>
<td>37.799</td>
<td>-0.9</td>
<td>$-4.55 \times 10^{-12}$</td>
<td>$-2.27 \times 10^{-12}$</td>
<td>$-1.13 \times 10^{-12}$</td>
<td>$-5.69 \times 10^{-13}$</td>
</tr>
<tr>
<td>IV</td>
<td>14.293</td>
<td>-0.9</td>
<td>$-1.20 \times 10^{-11}$</td>
<td>$-6.02 \times 10^{-12}$</td>
<td>$-3.01 \times 10^{-12}$</td>
<td>$-1.50 \times 10^{-12}$</td>
</tr>
<tr>
<td>V</td>
<td>7.9430</td>
<td>-0.8</td>
<td>$-2.16 \times 10^{-11}$</td>
<td>$-1.08 \times 10^{-11}$</td>
<td>$-5.42 \times 10^{-12}$</td>
<td>$-2.71 \times 10^{-12}$</td>
</tr>
</tbody>
</table>

Table 2.4: Nominal shape coefficients of generated reference data

Figure 2.5: Representation of aspherical surfaces based on nominal shape coefficients

The two considered algorithms make use of LS fitting as initial approximate solution for MZ since it is straightforward to compute. LS fitting of aspherical and freeform surfaces could be performed using quasi-Newtonian methods such as the Broyden-Fletcher-Goldfarb-Shanno (BFGS) method which are effective in the case of unconstrained nonlinear problems with a large number of data points [123, 127, 128]. This method approximates the inverse Hessian of the function by accumulating information from previous iterations. Therefore, a sequence of matrices is constructed throughout. The low memory Broyden-Fletcher-Goldfarb-Shanno (L-BFGS) is an enhanced version of BFGS optimisation algorithms that consists in reducing memory usage when storage is critical [123]. This method is suitable for applications involving large volume of data and variables. At each iteration, the Hessian is approximated using $K$ previous iterations, where $K$ is
Chapter 2. Fitting algorithms of aspherical and freeform surfaces

a user defined parameter. When applied to the case of aspherical surfaces, the L-BFGS method shows a linear computational complexity with respect to the number of points, runs faster and provides similar or better results than the Levenberg-Marquardt (LM) method [129, 51].

Figure 2.6 presents the procedure followed for the comparison of EPF to PDIP. First, LS fitting is performed on the generated reference data using L-BFGS algorithm so as an initial solution for MZ fitting could be determined. Then, the resulting data is submitted simultaneously to both algorithms (EPF and PDIP) and the returned value of PV of each algorithm is compared to the MZ reference value (MZ_ref). Finally, a statement about the acceptance or rejection of the algorithm could be made based on this comparison. It is to be noted that for this comparison, the two algorithms were coded using Matlab© software.

For each of the configurations presented in table 2.4 and for each value of MZ reference value chosen from the set MZ_ref=\{10^{-3} \text{mm}, 10^{-4} \text{mm}, 10^{-5} \text{mm}\}, a number of reference data containing \(N = \{132, 224, 504, 1188\}\) points were generated. Figure 2.7 shows generated reference data for the case of configuration V and MZ reference value of \(10^{-4} \text{mm}\).
2.4. Form assessment using EPF and PDIP

Figure 2.7: Generated reference data for the configuration V, $N = 1188$ and $MZ_{\text{ref}} = 10^{-4} \text{mm}$

Figure 2.8 represents the evolution of the difference between the returned value of $MZ$ and the reference value for both algorithms (configuration III with $N = 1188$ and $MZ_{\text{ref}} = 10^{-4} \text{mm}$). This represents a typical behaviour of the two algorithms for most of the considered cases. In fact, the convergence of EPF is faster than PDIP. PDIP takes approximately five times the time required for EPF to converge. Moreover, the convergence time of PDIP rapidly increases with the number of points compared to EPF. This is essentially due to the number of data points in the case of PDIP. In fact, PDIP requires the resolution of linear systems with a size of the same order as the number of existing points in the data set which explains the increasing execution time. Moreover, the resolution becomes infeasible beyond a given number of data points because of memory saturation.

Table 2.5 shows the relative error between $MZ$ values and reference ones. In terms of accuracy, the two algorithms return accurate results (to the third order in average). In the worst case, the obtained values are accurate to the third order with a slight superiority of EPF over PDIP.
Chapter 2. Fitting algorithms of aspherical and freeform surfaces

Figure 2.8: Evolution of the difference between the obtained value of MZ and the reference value for PDIP and EPF in the case of configuration III with $N = 1188$ and $MZ_{\text{ref}} = 10^{-4} \, mm$.

<table>
<thead>
<tr>
<th>Configuration</th>
<th>$MZ_{\text{ref}} , (mm)$</th>
<th>132</th>
<th>224</th>
<th>504</th>
<th>1188</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$10^{-3}$</td>
<td>5.079</td>
<td>1.512</td>
<td>5.262</td>
<td>1.270</td>
</tr>
<tr>
<td></td>
<td></td>
<td>7.551</td>
<td>7.500</td>
<td>1.350</td>
<td>8.011</td>
</tr>
<tr>
<td></td>
<td>$10^{-4}$</td>
<td>2.929</td>
<td>6.082</td>
<td>4.854</td>
<td>1.174</td>
</tr>
<tr>
<td></td>
<td></td>
<td>6.223</td>
<td>2.001</td>
<td>2.046</td>
<td>1.002</td>
</tr>
<tr>
<td></td>
<td>$10^{-5}$</td>
<td>1.275</td>
<td>2.001</td>
<td>1.005</td>
<td>1.769</td>
</tr>
<tr>
<td></td>
<td></td>
<td>7.001</td>
<td>2.002</td>
<td>1.002</td>
<td>5.888</td>
</tr>
</tbody>
</table>

II

|              | $10^{-3}$                | 2.889 | 2.041 | 5.907 | 2.079 |
|              |                          | 5.433 | 2.319 | 1.676 | 2.424 |
|              | $10^{-4}$                | 1.026 | 3.256 | 3.601 | 1.384 |
|              |                          | 4.784 | 1.856 | 1.211 | 3.411 |
| $10^{-5}$    |                          | 6.143 | 6.368 | 4.006 | 3.751 |
2.4. Form assessment using EPF and PDIP

<table>
<thead>
<tr>
<th></th>
<th>10^{-3}</th>
<th>10^{-4}</th>
<th>10^{-5}</th>
</tr>
</thead>
<tbody>
<tr>
<td>III</td>
<td>2.331 \times 10^{-3}</td>
<td>1.610 \times 10^{-6}</td>
<td>6.039 \times 10^{-5}</td>
</tr>
<tr>
<td></td>
<td>1.049 \times 10^{-8}</td>
<td>1.304 \times 10^{-7}</td>
<td>4.002 \times 10^{-3}</td>
</tr>
<tr>
<td></td>
<td>1.610 \times 10^{-6}</td>
<td>1.021 \times 10^{-4}</td>
<td>2.708 \times 10^{-6}</td>
</tr>
<tr>
<td></td>
<td>5.179 \times 10^{-10}</td>
<td>9.465 \times 10^{-5}</td>
<td>8.947 \times 10^{-9}</td>
</tr>
<tr>
<td></td>
<td>7.429 \times 10^{-5}</td>
<td>1.303 \times 10^{-3}</td>
<td>1.722 \times 10^{-3}</td>
</tr>
<tr>
<td></td>
<td>7.266 \times 10^{-10}</td>
<td>1.222 \times 10^{-3}</td>
<td>1.811 \times 10^{-3}</td>
</tr>
<tr>
<td></td>
<td>3.285 \times 10^{-7}</td>
<td>3.591 \times 10^{-2}</td>
<td>1.773 \times 10^{-7}</td>
</tr>
<tr>
<td></td>
<td>6.115 \times 10^{-8}</td>
<td>8.535 \times 10^{-6}</td>
<td>1.722 \times 10^{-3}</td>
</tr>
</tbody>
</table>

Table 2.5: Relative errors between obtained values of MZ and reference values in (%).
Blue: PDIP, orange: EPF

2.4.4 Comparison of EPF and PDIP using measured data

The two algorithms were tested on the measured data of the AO775 lens used in the previous project FORM-IND10 [51]. This artefact was manufactured by Anteryon \(^{©}\) company from Germany using a Single Point Diamond Turning (SPDT) process and finished with a high precision polishing process and class coating (figure 2.9). It has
a rectangular base of dimension 11.1 mm by 19.2 mm and a height of 2.2 mm. The nominal shape parameters of the selected asphere are listed in table 2.6.

![Figure 2.9: A photograph of the AO775 asphere](image)

<table>
<thead>
<tr>
<th>$R$ (mm)</th>
<th>$10^{20}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$k$</td>
<td>-1</td>
</tr>
<tr>
<td>$a_2$ (mm$^{-3}$)</td>
<td>0.0223</td>
</tr>
<tr>
<td>$a_4$ (mm$^{-5}$)</td>
<td>$7.29310^{-6}$</td>
</tr>
<tr>
<td>$a_6$ (mm$^{-7}$)</td>
<td>$4.5210^{-9}$</td>
</tr>
<tr>
<td>$a_8$ (mm$^{-9}$)</td>
<td>$-1.06110^{-11}$</td>
</tr>
<tr>
<td>$a_{10}$ (mm$^{-9}$)</td>
<td>$9.88710^{-15}$</td>
</tr>
</tbody>
</table>

Table 2.6: Nominal shape parameters of the AO775 asphere

The asphere was measured at the Netherlands Organisation of Applied Scientific Research (TNO) using the ultra-high precision Nanomefios measurement machine [7, 49], designed specifically for non-contact measurement of aspherical and freeform optics. This machine has an original architecture and higher measurement speed compared to orthogonal setup machines.

A set of 1129 points was collected and analysed using the implemented EPF and PDIP fitting methods. Table 2.7 shows the obtained results. The PV values obtained by LS (729 nm) are considerably higher than those given by EPF (624 nm) and PDIP.
2.5. The Hybrid Trust Region algorithm (HTR)

(671 nm). The same behaviour as for reference data was noticed when comparing PV values given by EPF and PDIP as well as execution time. In fact, EPF returns a lower value than PDIP with an execution time which is approximately five times shorter. In the next section, a new fitting algorithm called the Hybrid Trust Region (HTR) is presented. The performance of the latter algorithms especially in terms of execution time is what motivates going beyond EPF.

<table>
<thead>
<tr>
<th></th>
<th>EPF</th>
<th>PDIP</th>
<th>LS</th>
</tr>
</thead>
<tbody>
<tr>
<td>PV (nm)</td>
<td>624</td>
<td>671</td>
<td>729</td>
</tr>
<tr>
<td>Execution time (s)</td>
<td>60.84</td>
<td>396.14</td>
<td>31.03</td>
</tr>
</tbody>
</table>

Table 2.7: PV results obtained using EPF, PDIP and L-BFGS (LS) applied on Nanomefos measured data, \( N = 1129 \)

2.5 The Hybrid Trust Region algorithm (HTR)

In this section, the hybrid trust region algorithm suggested by Wang et al. in [130, 131] is introduced in the domain of metrology for the first time and adapted to the case of MZ fitting problems. The algorithm is then compared to EPF since it represents better performances than PDIP as shown in Section 2.4. The main idea of HTR consists of performing either trust region step, line search step or curve search step according to the specific situation faced at each iteration. This enables to avoid the resolution of the trust region problem many times.

In the following, an equivalent MZ fitting problem to eq.(2.5) with a different notation is adopted eq.(2.12).

\[
\min_{\mathbf{x}} \Psi(\mathbf{x}) \quad \text{where} \quad \Psi(\mathbf{x}) = \max_{1 \leq i \leq N} f_i(\mathbf{x}) \quad \text{and} \quad \mathbf{x} = \{m, s\} \in \mathbb{R}^n \quad (2.12)
\]

\( f_i \) is the Euclidean distance between the measured point \( (P_i) \) and its corresponding projection \( (Q_i) \) on the surface \( S \). \( s \) and \( m \) are the same as in subsection 2.2.2.

At each iteration, a first step consists in constructing the approximate quadratic programming and obtaining a trust region step \( d_k \) by solving the quadratic problem given in eq.(2.13).
Chapter 2. Fitting algorithms of aspherical and freeform surfaces

\[ QP(x_k, B_k) : \begin{cases} 
\min_{(d, z) \in \mathbb{R}^{n+1}} \frac{1}{2} < d, B_k d > + z = M_k(d, z) \\
\text{s.t.} \quad < \nabla f_i(x_k), d > - z \leq \Psi(x_k) - f_i(x_k), \quad i = 1, ..., N \\
||d||_\infty \leq \Delta_k 
\end{cases} \tag{2.13} \]

where \( B_k \) is an \( n \) by \( n \) symmetric positive definite matrix, \( \Delta_k \) is the parameter defining the trust region domain, \( z \) is an introduced parameter depending on the first derivative of the objective function \( \Psi \), \( \nabla f_i \) is the gradient of the function \( f_i \) and \( < ., . > \) denotes the dot product.

The trust region domain is defined using \( L_\infty \) instead of \( L_2 \) norm so that \( QP \) becomes an easily-solved quadratic problem. It should be mentioned that the proposed \( QP \) in eq.(2.13) has always a solution since \((0, 0)\) lies inside the feasible domain. This problem could be solved using classical methods adapted to quadratic problems such as interior point methods [132].

If the resulting trust region trial step \( d_k \) could not be accepted, a corrected step \( d_k + \tilde{d}_k \) is determined by solving the problem in eq.(2.14).

\[ \tilde{QP}(x_k, B_k) : \begin{cases} 
\min_{(\tilde{d}, \tilde{z}) \in \mathbb{R}^{n+1}} \frac{1}{2} < \tilde{d}, B_k (d_k + \tilde{d}) > + \tilde{z} = \tilde{M}_k(\tilde{d}, \tilde{z}) \\
\text{s.t.} \quad < \nabla f_i(x_k), \tilde{d} > - \tilde{z} \leq \Psi(x_k + d_k) - f_i(x_k + d_k), \quad i = 1, ..., N \\
||d_k + \tilde{d}||_\infty \leq \Delta_k 
\end{cases} \tag{2.14} \]

If neither the initial trust region step \( d_k \) nor the corrected step \( d_k + \tilde{d}_k \) could be accepted in trust region scheme, a line search along \( d_k \) or a curve search is performed if \( d_k \) is a descent direction (the actual reduction \( r_k > 0 \) in eq.(2.16)). Otherwise \( (r_k \leq 0) \), a curve search is used to find a step length \( t_k \) that verifies eq.(2.15).

\[ \Psi(x_k + t_k d_k + t_k^2 \tilde{d}_k) \leq \Psi(x_k) - \alpha t_k < d_k, B_k d_k > \tag{2.15} \]

where \( \alpha \in [0, 1/2] \), \( d_k \) is the solution of eq.(2.13) and \( \tilde{d}_k \) is the solution of eq.(2.14). In the case \( ||d_k|| \leq ||\tilde{d}_k|| \), \( \tilde{d}_k \) should be set to 0. The implemented algorithm is presented
as follows:

**Step 1:** Given initial values: \( x_0 \in \mathbb{R}^n, \epsilon > 0, \Delta_{max} > 0, \Delta_0 \in [0, \Delta_{max}], 0 < \tau_1 < 1 < \tau_2, \alpha \in [0, \frac{1}{2}], \beta \in [0, \frac{1}{2}], 0 < \mu < 2\alpha, \eta \in [\frac{1}{2}, 1], B_0 = I, k := 0 \) and \( k_{max} \).

**Step 2:** Determine \((d_k, z_k)\) by solving the quadratic problem in eq.(2.13). If \(||d_k|| \leq \epsilon\) or \(k_{max}\) is attained, stop. Otherwise;

**Step 3:** Compute the ratio between the actual reduction and the predicted reduction.

\[
 r_k = \frac{\Psi(x_k) - \Psi(x_k + d_k)}{M_k(0, 0) - M_k(d_k, z_k)} 
\]  (2.16)

**Step 4:** (Update the iteration point)

- (4.1) if \( r_k > \mu \), set \( s_k = d_k \), \( x_{k+1} = x_k + s_k \), go to step 5. Otherwise;
- (4.2) compute the second-order correction step \( \tilde{d}_k \) by solving the problem in eq.(2.14). In the case \(||d_k|| \leq ||\tilde{d}_k||\), \( \tilde{d}_k \) is set to be 0.
- (4.3) compute corrected \( \tilde{r}_k \)

\[
 \tilde{r}_k = \frac{\Psi(x_k) - \Psi(x_k + d_k + \tilde{d}_k)}{M_k(0, 0) - M_k(d_k, z_k)} 
\]  (2.17)

- (4.4) if \( \tilde{r}_k > \mu \), set \( r_k = \tilde{r}_k \), \( s_k = d_k + \tilde{d}_k \), go to step 5. Otherwise,
- (4.5) if \( r_k > 0 \), set \( \tilde{d}_k = 0 \).
- (4.6) (perform curve search) Compute \( t_k \): the first number in the sequence of \( \{1, \beta, \beta^2, \ldots\} \) to verify eq.(2.15). Set \( s_k = t_k d_k + t_k^2 \tilde{d}_k \) and \( x_{k+1} = x_k + s_k \)

**Step 5:** (Update \( \Delta_k \))

- if \( r_k \leq \mu \), \( \Delta_{k+1} \in [||s_k||, \tau_1 \Delta_k] \);
- if \( r_k \geq \eta \), \( \Delta_{k+1} = min(\tau_2 \Delta_k, \Delta_{max}) \);
- otherwise, \( \Delta_{k+1} = \Delta_k \)

**Step 6:** (Update \( B_k \)). Update \( B_k \) to \( B_{k+1}; k := k + 1 \), go to step 1.

To update \( B_k \), the Powell modification of BFGS formula is used [133]. A simplified flowchart of the hybrid trust region algorithm is shown in figure 2.10.
Validation and comparison of the newly introduced algorithm to EPF was conducted using reference data, benchmark data as well as measured data.

### 2.5.1 Application on reference data

As for the comparison of EPF and PDIP, a number of reference data were used for the validation of HTR. Nominal coefficients of the considered configuration of the aspherical shapes, supposed known, are listed in table 2.8.

For each configuration, data with a predefined number of points $N=\{121, 1024, 10404, 100489\}$ and a previously known $M_{\text{ref}}=\{10^{-3} \text{ mm}, 10^{-4} \text{ mm}, 10^{-5} \text{ mm}\}$ are generated in order to assess the performance of HTR with regards to EPF, each generated data was submitted to both algorithms. The corresponding PV values as well as execution time for both algorithms are compared.
2.5. The Hybrid Trust Region algorithm (HTR)

Table 2.8: Nominal coefficients of aspherical surfaces used for the validation of the HTR algorithm

<table>
<thead>
<tr>
<th>Config.</th>
<th>R (mm)</th>
<th>k</th>
<th>$a_4$ (mm$^{-3}$)</th>
<th>$a_6$ (mm$^{-5}$)</th>
<th>$a_8$ (mm$^{-7}$)</th>
<th>$a_{10}$ (mm$^{-9}$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>I</td>
<td>101.58</td>
<td>-1</td>
<td>$-1.70 \times 10^{-13}$</td>
<td>$-8.51 \times 10^{-14}$</td>
<td>$-4.25 \times 10^{-14}$</td>
<td>$-2.12 \times 10^{-14}$</td>
</tr>
<tr>
<td>II</td>
<td>19.79</td>
<td>-0.9</td>
<td>$-1.51 \times 10^{-17}$</td>
<td>$-7.55 \times 10^{-18}$</td>
<td>$-3.77 \times 10^{-18}$</td>
<td>$1.88 \times 10^{-19}$</td>
</tr>
<tr>
<td>III</td>
<td>8.88</td>
<td>-0.8</td>
<td>$-1.94 \times 10^{-12}$</td>
<td>$-9.72 \times 10^{-13}$</td>
<td>$-4.86 \times 10^{-13}$</td>
<td>$-2.43 \times 10^{-13}$</td>
</tr>
<tr>
<td>IV</td>
<td>4.14</td>
<td>-0.9</td>
<td>$-4.17 \times 10^{-12}$</td>
<td>$-2.08 \times 10^{-12}$</td>
<td>$-1.04 \times 10^{-12}$</td>
<td>$-5.21 \times 10^{-13}$</td>
</tr>
<tr>
<td>V</td>
<td>0.77</td>
<td>-1</td>
<td>$-2.22 \times 10^{-11}$</td>
<td>$-1.11 \times 10^{-11}$</td>
<td>$-5.56 \times 10^{-12}$</td>
<td>$-2.78 \times 10^{-12}$</td>
</tr>
</tbody>
</table>

Tables 2.9, 2.10 and 2.11 present obtained results of PV–$M_{Z_{\text{ref}}}$ as well as the execution time for configurations II, III and IV and $M_{Z_{\text{ref}}}=10^{-4}$ mm. EPF and HTR return accurate values of MZ with superiority of HTR in average.

Regarding execution time, EPF is five times slower than HTR especially for data sets that exceed 1000 points. When the number of points contained in the data set are multiplied by 1000, HTR execution time ($T_{HTR}$) is multiplied by 50 while $T_{HTR}$ by 150 in their corresponding worst cases. Figure 2.11 shows the evolution of the execution time as a function of the number of points in the data set for the two considered algorithms.

Table 2.9: Values of PV–$M_{Z_{\text{ref}}}$ (mm) and execution time (s) for HTR and EPF (configuration II)

<table>
<thead>
<tr>
<th>N</th>
<th>$M_{Z_{\text{HTR}}}-M_{Z_{\text{ref}}}$ (mm)</th>
<th>$M_{Z_{\text{EPF}}}-M_{Z_{\text{ref}}}$ (mm)</th>
<th>$T_{HTR}$ (s)</th>
<th>$T_{EPF}$ (s)</th>
</tr>
</thead>
<tbody>
<tr>
<td>121</td>
<td>$4.06 \times 10^{-19}$</td>
<td>$1.45 \times 10^{-15}$</td>
<td>0.84</td>
<td>2.68</td>
</tr>
<tr>
<td>1024</td>
<td>$9.50 \times 10^{-16}$</td>
<td>$1.11 \times 10^{-16}$</td>
<td>2.07</td>
<td>9.15</td>
</tr>
<tr>
<td>10404</td>
<td>$1.78 \times 10^{-16}$</td>
<td>$4.51 \times 10^{-13}$</td>
<td>12.1</td>
<td>61.91</td>
</tr>
<tr>
<td>100489</td>
<td>$1.64 \times 10^{-16}$</td>
<td>$4.92 \times 10^{-15}$</td>
<td>41.51</td>
<td>226.96</td>
</tr>
</tbody>
</table>

EPF is a smoothing technique that consists of approximating the non-differentiable objective function using a smooth one at each iteration and then minimising it. For this aim, a Newton-based method could be used which requires the computation of the Hessian matrix. Since the Hessian matrix calculation time is proportional to the number of points in the data set, the execution time significantly increases. Moreover, the obtained descent direction is not always accurate since the obtained Hessian matrix
Chapter 2. Fitting algorithms of aspherical and freeform surfaces

<table>
<thead>
<tr>
<th>N</th>
<th>MZ_{HTR}−MZ_{ref} (mm)</th>
<th>MZ_{EPF}−MZ_{ref} (mm)</th>
<th>T_{HTR} (s)</th>
<th>T_{EPF} (s)</th>
</tr>
</thead>
<tbody>
<tr>
<td>121</td>
<td>4.24 × 10^{-15}</td>
<td>1.13 × 10^{-14}</td>
<td>2.18</td>
<td>2.48</td>
</tr>
<tr>
<td>1024</td>
<td>2.18 × 10^{-16}</td>
<td>2.61 × 10^{-16}</td>
<td>4.32</td>
<td>4.47</td>
</tr>
<tr>
<td>10404</td>
<td>1.41 × 10^{-15}</td>
<td>3.08 × 10^{-15}</td>
<td>4.27</td>
<td>36.50</td>
</tr>
<tr>
<td>100489</td>
<td>6.73 × 10^{-15}</td>
<td>6.78 × 10^{-15}</td>
<td>20.52</td>
<td>255.06</td>
</tr>
</tbody>
</table>

Table 2.10: Values of PV−MZ_{ref} (mm) and execution time (s) for HTR and EPF (configuration III)

<table>
<thead>
<tr>
<th>N</th>
<th>MZ_{HTR}−MZ_{ref} (mm)</th>
<th>MZ_{EPF}−MZ_{ref} (mm)</th>
<th>T_{HTR} (s)</th>
<th>T_{EPF} (s)</th>
</tr>
</thead>
<tbody>
<tr>
<td>121</td>
<td>4.03 × 10^{-15}</td>
<td>9.78 × 10^{-14}</td>
<td>1.14</td>
<td>1.79</td>
</tr>
<tr>
<td>1024</td>
<td>8.14 × 10^{-16}</td>
<td>8.42 × 10^{-15}</td>
<td>5.23</td>
<td>7.24</td>
</tr>
<tr>
<td>10404</td>
<td>1.15 × 10^{-16}</td>
<td>2.14 × 10^{-15}</td>
<td>35.88</td>
<td>200.16</td>
</tr>
<tr>
<td>100489</td>
<td>9.86 × 10^{-15}</td>
<td>1.05 × 10^{-14}</td>
<td>50.03</td>
<td>274.75</td>
</tr>
</tbody>
</table>

Table 2.11: Values of PV−MZ_{ref} (mm) and execution time (s) for HTR and EPF (configuration IV)

![Figure 2.11: Evolution of the execution time of EPF and HTR in function of the number of points in the data set for configurations II, III and IV](image)

54
is not always definite positive. Hence, corrections must be brought to the Hessian matrix whenever needed. On the other hand, when establishing the $QP$ for HTR algorithm, the matrix $B$ is chosen to be symmetric positive definite, the Powell modification of BFGS formula proves to be efficient for this purpose and there is no need to calculate the second order derivation terms, which considerably reduces execution time.

### 2.5.2 Application on benchmark data

Available data in literature were also used to assess the results obtained using the proposed algorithm. A heuristic method based on DE method was suggested in [116] and a set of 8100 points was used to verify the algorithm. The nominal coefficients of the aspheric surface are: $R = 520 \, mm$, $k = -0.7$, $a_4 = 5.2 \times 10^{-5} \, (mm^{-3})$, $a_6 = -6.5 \times 10^{-6} \, (mm^{-5})$, $a_8 = 3.11 \times 10^{-8} \, (mm^{-7})$ and $a_{10} = 3.222 \times 10^{-9} \, (mm^{-9})$. Fractal Brownian function was used to generate noise with amplitude $\sigma = 1 \, \mu m$ around the nominal surface. The same data points were adopted for a comparison between EPF and DE in [119].

Table 2.12 shows the obtained results by EPF and HTR. The PV value given by EPF is $3.2 \, \mu m$ while the newly proposed HTR algorithm returns a smaller value of $3.15 \, \mu m$ which is approximately $60 \, nm$ lower. A $60 \, nm$ difference on form errors may result in the rejection of a conforming lens. With respect to execution time, the values are approximately similar.

<table>
<thead>
<tr>
<th></th>
<th>HTR</th>
<th>EPF</th>
</tr>
</thead>
<tbody>
<tr>
<td>PV ((\mu m))</td>
<td>3.15</td>
<td>3.21</td>
</tr>
<tr>
<td>Execution time (s)</td>
<td>2.39</td>
<td>2.34</td>
</tr>
</tbody>
</table>

Table 2.12: Comparison of HTR and EPF on benchmark data (8100 points)

### 2.5.3 Experimental investigation

The experimental investigation of the HTR algorithm was carried out on a different measurement of the AO775 asphere (31390 points) obtained using the Nanomeforn machine (figure 2.12).
Table 2.13 shows the obtained results using EPF and HTR as well as PV returned by LS fitting. HTR gives a PV value 9 nm smaller than EPF. This value could be significant when targeting uncertainties at the nanometre level.

<table>
<thead>
<tr>
<th></th>
<th>LS</th>
<th>MZ</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>HTR</td>
<td>EPF</td>
</tr>
<tr>
<td>PV (nm)</td>
<td>536</td>
<td>470.36</td>
</tr>
<tr>
<td>Execution time (s)</td>
<td>10.23</td>
<td>107.69</td>
</tr>
</tbody>
</table>

Table 2.13: Comparison of HTR and EPF on measured data (31390 points)

Figure 2.13 presents the evolution of EPF and HTR in function of execution time. This shows the rapidity of HTR to converge compared to EPF. The final residuals obtained using HTR are presented in figure 2.14.
2.5. The Hybrid Trust Region algorithm (HTR)

Figure 2.13: Evolution of PV value for HTR and EPF algorithms applied on measured data

Figure 2.14: Fitting residuals of the AO775 asphere (31390 points)
2.6 Numerical uncertainties

With reference to the guide to the expression of uncertainty in measurement (GUM) [134], the uncertainty is defined as a "parameter associated with the results of a measurement that characterises the dispersion of the values that could be attributed to the measurand". The standard uncertainty (denoted $u$) is often taken as a dispersion measure.

When dealing with form error determination, there might always be an error on the estimated value even when the recorded data lie exactly on the manufactured part. This is due to the used sampling strategy and points positions when measuring the artefact [135, 136, 137]. Figure 2.15 is an illustration of this phenomenon in the case of flatness tolerance. In fact, the sampled points, even being exactly on the manufactured artefact, could not record all the peaks which results in an underestimation of the returned value.

![Figure 2.15: Comparison of estimated MZ value when considering the continuous manufacturing profile (blue) and the measured discretized profile (orange)](image)

In the following, this error component is not addressed since it could not be avoided. The dispersion of the returned MZ must be expressed in function of the dispersion of the recorded coordinates due to the different measurement errors (figure 2.16).

Unlike other fitting criteria, the returned value of the MZ depends on the position of each measured point. This implies that errors on one recorded point might drastically change the MZ value. A pre-processing step that consists of filtering and eliminating aberrant points (outliers) is necessary.

A conventional way to determine the uncertainty of the algorithms’ output is to use the propagation rules stated in the Guide to the Expression of Uncertainty in Measurement (GUM) [134]. This straightforward method could not be used in the case of complex algorithms and hence only black-box testing methods could be used. These
2.6. Numerical uncertainties

Figure 2.16: Comparison of estimated MZ based on exact measured points (orange) and measurement including errors (green)

methods aim at estimating the uncertainty of algorithms’ output without necessarily understanding how these algorithms work (figure 2.17).

![Black-box testing methods](image)

An attempt has been made in [138] to determine algorithm uncertainty based on matrix form of the algorithm. The determination of the matrix form of the algorithm is not usually an easy task. This work describes an identification method of its coefficients. However, due to the complexity of the MZ fitting algorithms, this method could not be extended to our case. In [139], an uncertainty analysis of geometric best fit is proposed by examining the uncertainty zones of the six motion parameters. The bootstrap method was used in [140] to determine the uncertainty of extreme fit. In [141], sensitivity analysis was employed to determine bound of the transformation parameters with respect to the measurement data in the case of LS fitting. Other statistical-based methods were used in [142] adapted to size tolerance. The common feature with all the stated methods is that the uncertainty is generally estimated on the best fit geometry.

The suggested method in this work focuses on directly determining the uncertainty on
Chapter 2. Fitting algorithms of aspherical and freeform surfaces

the MZ by combining reference data generations and Monte Carlo simulations [143]. A first step consists in estimating the uncertainty associated to the measuring instrument. This could be achieved by listing all the elements that contribute to the uncertainty using fishbone (Ishikawa) diagram [144]. The individual uncertainty of each component is combined to get a global uncertainty on each coordinate $X$, $Y$ and $Z$ ($\sigma_X$, $\sigma_Y$ and $\sigma_Z$). The next step is to generate a high number of reference data, for which we know the exact value of form error, to which simulated measuring errors are added. The resulting data points are successively submitted to the fitting algorithm that gives a different value of MZ corresponding to each inputted set. A statistical analysis could be performed in order to determine the uncertainty on the MZ value ($\sigma_{MZ}$). This procedure is summarised in figure 2.18.

![Figure 2.18: Uncertainty analysis of MZ fitting algorithm](image)

It is to be noted that the algorithms are supposed to return accurate results of MZ values in absence of measuring errors. This statement could be justified by the obtained results in Sections 2.5 and 2.4. Also, the generated data should reproduce the actual measuring conditions such as the number of measured points, the nominal coefficients values, the points density, etc.

As an illustration and given the uncertainties on the measured data, the described
method is used to determine the uncertainty on the returned MZ value obtained by the validated HTR algorithm. The nominal shape coefficients of the aspherical surface are: 

\[ R = 9.127 \times 10^{20} \text{mm}, \quad k = -1, \quad a_4 = 1.278 \times 10^{-9} \text{mm}^{-3}, \quad a_6 = 7.922 \times 10^{-16} \text{mm}^{-5}, \]
\[ a_8 = -1.859 \times 10^{-18} \text{mm}^{-7} \quad \text{and} \quad a_{10} = 1.733 \times 10^{-21} \text{mm}^{-9}. \]

The uncertainties on the measured data are estimated as: \( \sigma_X = 20 \text{nm}, \sigma_Y = 20 \text{nm} \) and \( \sigma_Z = 15 \text{nm} \). These values are supposed spatially uniform.

Fifty simulations were run for each value of form error in the set: \{10^{-3} \text{mm}, 10^{-4} \text{mm}, 10^{-5} \text{mm}\} and for each number of points in: \{144, 2704, 5929 and 10201\}. Figure 2.19 gives the difference between obtained PV and \( \text{MZ}_{\text{ref}} \) as well as the associated uncertainties.

![Figure 2.19: Difference between obtained PV values and \( \text{MZ}_{\text{ref}} \) values and their associated uncertainties](image)

Obtained results show the existence of systematic and random errors in the obtained results compared to exact ones. The magnitude of the systematic errors is increasing with the number of points contained in the data set. This result seems coherent since, as stated before, the value of MZ is highly sensitive to the measuring errors associated to each point in the data set. With an increasing number of points, the probability to
affect the resulting MZ value is higher. For the same reason, the standard deviation of random errors is decreasing.

These results remain illustrative of the general behaviour of uncertainty on the MZ as a function of the number of points and the estimated value of form error. The graph may slightly change when a different probability distribution is used to generate reference data.

2.7 Conclusion

In this chapter, the relevance of using MZ fitting for the assessment of form error of aspherical and freeform components was outlined and the mathematical formulation of the underlying optimisation problem was clearly stated. A comprehensive state of the art allowed the identification of candidate fitting algorithms. Two selected methods (EPF and PDIP) were thoroughly compared using reference and measured data. Obtained results show superiority of EPF over PDIP regarding accuracy of obtained results as well as execution time.

An existing method called the Hybrid Trust Region algorithm (HTR) was introduced in metrology and adapted to the case of MZ fitting of aspherical and freeform surfaces. According to the comparison of HTR to EPF using reference data, measured data and benchmark data, HTR outperforms EPF regarding returned form error values and execution time.

Since form error values obtained using MZ fitting are highly affected by errors on measured data, an estimation of the uncertainty of MZ is necessary. A new method combining the use of reference data and the Monte Carlo simulations was used. This method returns the estimated uncertainty on MZ in function of the number of points contained in the data set as well as the estimated uncertainty on measured data.

In the future, this work should be extended to the determination of shape parameters based on the same fitting criterion namely MZ. The huge difference in parameters' magnitude (notice the difference between $R$ and $|a_{10}|$ in the case of aspherical surfaces for instance), the existence of multiple local solutions as well as the poor number of methods to guess the initial solution are the major hurdles. Nevertheless, the determination of form error using MZ fitting and taking in consideration motion parameters is enough since the nominal shape coefficients are usually taken as input and only the derivation
of the obtained artefact from the desired one are to be quantified.

To conclude, despite its conformance to the ISO GPS, MZ fitting criterion must be used with caution since it is sensitive to outliers and no standard filtering (in the sense of outliers removal) methods exist. The development of more robust fitting criteria could also be considered. The work presented in this section was the subject of two journal publications in "Precision Engineering journal" and "Journal of the International Measurement Confederation" respectively titled: "Investigation of minimum zone assessment methods for aspheric shapes" [145] and "A novel hybrid trust region minimax fitting algorithm for accurate dimensional metrology of aspherical shapes" [146].
Chapter 3

Validation of fitting algorithms
3.1 Introduction

Development of reference fitting algorithms is essential in the metrology of aspherical and freeform surfaces but it is not enough since procedures for the validation of these algorithms must be set.

The aim of software validation is to make sure that the fitting algorithms return correct values. The definition of the term "correct" might be broad but there exist different quality metrics to assess the correctness of the values returned by an algorithm [147]. The need for software validation was initiated during a project led by the PTB and supported by the Community Bureau of Reference (BCR) in 1983 [148, 149]. This work has concluded that there is an urgent need to improve fitting software and suggest validation methods since a good number of available computational software give incoherent results.

It is to be noted that software validation procedures for reference algorithms is not the same as for operating (commercial) ones. In [150], key characteristics of reference software were outlined. It is stated that they must:

- perform well for representative data,
- work sensibly for unrepresentative data and be able to detect extreme cases,
- must perform efficiently in poor cases

Once again, the terms "perform well", "work sensibly" and "must perform efficiently" must be clearly defined. For this reason, a performance measure will be defined later on in this chapter.

Despite its importance, performance in terms of execution time is not the first characteristic sought for reference algorithms. However, the algorithm must be stable and robust. Stability means that the underlying numerical operations are numerically stable. In other words, small perturbations in input data must only result in small perturbations in output. Robustness refers to the software ability to handle extreme cases. Unlike reference software, performance is the key for commercial ones. The algorithm is supposed to perform well only for representative data.

In this chapter, the validation of metrology software using reference data is discussed. The next section presents the different works that have been conducted for the validation of metrology software. These works do not necessarily address the fitting problem but
the underlying logic could be extended to fitting algorithms. In Section 3.3, reference data generation methods for the case of LS and MZ fitting are detailed. Section 3.4 makes an extension of these methods to a special case known as reference data with a non-vertex solution applicable in the case of MZ fitting. Section 3.5 presents a detailed procedure for the validation of fitting algorithms based on two metrics namely, the degree of the difficulty and the performance measure. These parameters are explicitly defined for each fitting problem. A case study on the validation of a given algorithm concludes the section.

### 3.2 Literature review on algorithms’ validation

There exist mainly two methods for the validation of metrology software either using a reference pair or a reference software. These methods are called black-box testing methods which contrast with white box testing methods, where the source code is reviewed against its design [151]. Reference pairs are known in the field of surface texture as *Type F1 software measurement standards*, defined in the ISO 5436-2 2012 [152]. Even if this standard is common in surface texture domain, the principle of *Type F1* software standards could be extended to fitting algorithms. *Type F1* standards could be regarded as a numerical representation of the measured part to which we associate a reference measurand value known with a given uncertainty. For the evaluation, reference data are inputted to the software under test, the returned value is compared to the reference measurand and then a decision could be made whether the software is accepted or rejected (figure 3.1).

The second way is to evaluate metrology software by the mean of reference software, also defined in ISO 5436-2 2012 [152] as *Type F2 software measurement standards*. Reference software are traceable software against which the algorithm under test will be compared. A common set of data points is submitted to both software (reference software and software under test) and the two results are then compared in order to take an accept/reject decision (figure 3.2). Reference software do not exist for a wide range of applications in metrology. Moreover, their development is not always straightforward especially for MZ or LS fitting algorithms dedicated to aspherical and freeform surfaces.

At the level of NMIs a number of platforms were set for the generation of reference data corresponding to a specific application. As for surface texture, the PTB developed
a web-based reference software for roughness analysis [153]. The National Institute of Standards and Technology (NIST) also developed an algorithm testing system for the same purpose where the user can either download virtual data surface files and corresponding parameters from a database or use the embedded software to make calculations on its own surface topography [154]. In 2008, the National Physical Laboratory (NPL)
carried out a project named "SoftGauges" for the evaluation of areal surface texture parameters. The aim of the project is to develop consistent definitions, specifications and reference software for the evaluation of areal texture parameters [155]. A detailed comparison of the three platforms and improvement recommendations are given in [156].

Reference data generation for LS fitting was the subject of the ISO 10360-6:2001 [157]. This standard presents a reference data generation method that could be used in order to test software dedicated to the determination of associated Gaussian features to measured data. Among others, this standard specifies the required number of data sets, the file format and the detailed procedure for data generation. However, the presented method is only applicable to simple geometries such as lines, planes, circles, spheres, cylinders, cones, and torus. Moreover, the described method does not guarantee that the reference value of the measurand matches the generated data. PTB provided an online service called Traceability for Computationally-Intensive Metrology "TraCIM" that offers on-demand reference data for both LS and Chebyshev fitting. Generation of reference data and evaluation of obtained results conformed to some requirements specified by the TraCIM association [158]. In [6], A. Forbes et al. give practical methods in order to generate reference data for both LS and MZ fitting. An emphasis is put on these two methods in the next section.

3.3 Reference data generation

In this section, the method suggested by A. Forbes et al. in [6] for reference data generation was adapted to aspherical surfaces for the case of LS and MZ fitting. The main idea behind reference data generation is to state the optimality conditions for the considered fitting problem and then derive data sets that perfectly meet these conditions. For LS fitting, the method is called the null-space method since it is based on the determination of the null-space of the Jacobian matrix. For MZ fitting, KKT conditions are used for data generation.

3.3.1 Least Squares

We recall the fitting problem in the case of LS in eq.(3.1). \(d, x\) and \((p_i)_{1 \leq i \leq N}\) have the same definitions as in Chapter 2 and let \(x^*\) be the desired solution of the fitting problem.
Chapter 3. Validation of fitting algorithms

\[
\min_x \Phi_{LS} \quad \text{with} \quad \Phi_{LS} = \sum_{i=1}^{N} d_i^2(x, p_i) \quad \text{and} \quad d = (d_1, \ldots, d_N) \quad (3.1)
\]

For \(x^*\) to be a solution of the problem given in eq.(3.1), the first order optimality conditions imply that the gradient of \(\Phi_{LS}\) with respect to \(x = (x_1, \ldots, x_n)\) expressed at \(x^*\) must be equal to zero. This could be written in matrix form as in eq.(3.2).

\[
J^T d = 0 \quad (3.2)
\]

where \(J\) is the Jacobian matrix of \(d\) given in eq.(3.3).

\[
J^T = \begin{pmatrix}
\frac{\partial d_1}{\partial x_1} & \cdots & \frac{\partial d_N}{\partial x_1} \\
\vdots & \ddots & \vdots \\
\frac{\partial d_1}{\partial x_n} & \cdots & \frac{\partial d_N}{\partial x_n}
\end{pmatrix} \quad (3.3)
\]

In other words, the first order optimality conditions require that the deviation vector \(d\) must be in the null-space of \(J\). From eq.(3.1), we can see that the Jacobian matrix \(J\) is function of \((p_i)_{1\leq i\leq N}\) and \(x\). However, in [6], it was shown that if \(p_i = q_i + d_i n_i\), where \(q_i\) has the same definition as in 2 and \(n_i\) is the normal vector to the nominal surface in the point \(q_i\), the Jacobian matrix associated with \((q_i)_{1\leq i\leq N}\) is equal to \(J\).

The determination of the null-space could be derived using standard techniques such as QR or Singular Value Decomposition (SVD) [159]. The obtained factorisation using the former method is shown in eq.(3.4).

\[
J = QR = \begin{pmatrix} Q_1 & Q_2 \end{pmatrix} \begin{pmatrix} R_1 \\ 0 \end{pmatrix} \quad (3.4)
\]

with \(Q\) an orthogonal matrix \((QQ^T = I)\) and \(R\) is an upper-triangular matrix. A vector \(d\) satisfying eq.(3.2) could be chosen as \(d = Q_2 \hat{d}\) with \(\hat{d}\) is an arbitrary \(N - n\) vector. The reference data generation algorithm for the case of LS could be presented as follows:

**Step 1:** Given a nominal shape described using an implicit or explicit equation. Select \(N\) points lying on the surface denoted by \((q_i)_{1\leq i\leq N}\).
Step 2: Calculate the Jacobian matrix $J$.

Step 3: Perform the QR decomposition on $J$.

Step 4: From $Q$ select the last $N - n$ columns to form the $Q_2$ matrix.

Step 5: For any arbitrary $N - n$ vector $\hat{d}$, calculate the deviation vector $d = Q_2 \hat{d}$.

Step 6: Construct the reference data using $p_i = q_i + d_i n_i$.

The suggested method can also generate reference data with a predefined form deviation $d_0$. This could be achieved by selecting the closest deviation vector $d$ to $d_0$ that verifies eq.(3.2). This problem could be formulated as a quadratic programming as given in eq.(3.5).

$$\min_d ||d - d_0||$$
subject to $J^T d = 0$ (3.5)

The resolution of this problem is straightforward if a QR decomposition of the Jacobian matrix $J$ such in eq.(3.4) exists. In this case, the generation method follows the same previous scheme except in step 5 where the deviation vector is taken as $d = Q_2 Q_2^T d_0$. Similar approach could be used for the generation of pre-assigned spatially correlated form error.

3.3.2 Minimum Zone

The derivation of the optimality conditions in the case of MZ is more convenient when the fitting problem is formulated as a nonlinear programming. KKT conditions, which are a generalisation of Lagrange multipliers for equality constrained problems, give first and second order sufficient conditions [123]. Figure 3.3 summarises the method suggested by A. Forbes et.al for the generation of reference data for MZ fitting [6].
3.4 Extension to non-vertex solutions for aspherical surfaces

Reference data with non-vertex solutions are very important for testing metrology software since special techniques are required for algorithms to converge. Moreover, non-vertex solutions occur in practice as reported in [6]: "from 1000 randomly generated data sets involving four data points on each of four parallel circles on a cylinder, a non-vertex solution was detected in approximately 16% of the data sets".

Figure 3.3: A summarised flowchart for the generation of reference data for MZ fitting using the method developed in [6]. LICQ: Linear Independence Constraint Qualification [123]
3.4. Extension to non-vertex solutions for aspherical surfaces

3.4.1 Vertex vs. non-vertex solutions

In order to make the difference between vertex and non-vertex solutions, Chebyshev fitting is formulated as a nonlinear programming eq.(3.6).

$$\begin{align*}
\min_{x,e} & \quad e \\
\text{subject to} & \quad -e \leq d_i(x) \leq e \quad \forall i \in \{1,...,N\}
\end{align*}$$

where $e$, $x$ and $d_i$ are the same as in Chapter 2. The problem in eq.(3.6) could be formulated in the standard form given in eq.(3.7).

$$\begin{align*}
\min_y & \quad f(y) \\
\text{subject to} & \quad c_i^+(y) \geq 0 \quad \forall i \in \{1,...,N\} \\
& \quad c_i^-(y) \geq 0 \quad \forall i \in \{1,...,N\}
\end{align*}$$

where $y = (e,x) \in \mathbb{R}^{n+1}$, $c_i^+(y) = e - d_i(x)$, $c_i^-(y) = e + d_i(x)$ and $f(y) = e$.

A solution $y^*$ is said to be vertex if the number of active constraints is greater than or equal to $n + 1$. If the number of active constraints is strictly less than $n + 1$, $y^*$ is said to be a non-vertex solution. By definition, a constraint $c(y) \geq 0$ is active at $y^*$ if $c(y^*) = 0$. An active constraint could also be interpreted as a point belonging to the measured data for which the distance to the reference surface is equal to the form error (a contacting point to the enclosing envelope (figure 3.4)).

![Figure 3.4: Representation of contacting points to the enclosing envelope](image)

If the Chebyshev fitting could be formulated as a linear programming (as for the case of straightness tolerance) [160], the solution will be at some vertex of the feasible
domain (figure 3.5).

Figure 3.5: Representation of a vertex solution for the case of a linear programming involving the two variables X1 and X2

For more complex surfaces, the solution could be either at some vertex of the feasible domain (vertex solution), or at a face (or edge) (non-vertex solution (figure 3.6).

Figure 3.6: Representation of a vertex solution for the case of a nonlinear programming involving the two variables X1 and X2
3.4.2 General approach

In [6], an attempt has been made in order to generate reference data with a non-vertex solution for the case of cylindricality. The proposed method is a modified version of the one presented in Section 3.3.2 that makes use of the structure of the gradient vector of distances in the case of cylindricality. In this case, the determination of the Lagrangian multipliers that satisfy dual feasibility conditions is easier than for the case of complex shapes.

Here, we present an extension of this method to the case of aspherical surfaces when only motion parameters are sought. Since aspherical lenses are rotationally symmetric surfaces, only five motion parameters are unknown $\mathbf{x} = (T_X, T_Y, T_Z, \theta_X, \theta_Y)$ (translations in $X$, $Y$ and $Z$ directions as well as rotations around $X$ and $Y$ axis). Therefore, a non-vertex solution consists of five contacting points (active constraints) at most.

**Step 1:** five points $\{ \mathbf{q}_i = (x_{q,i}, y_{q,i}, z_{q,i}) \}_{1..5}$ are randomly generated on the nominal aspherical surface. These points represent the orthogonal projections of the contacting points $\{ \mathbf{p}_i \}_{1..5}$ onto the nominal shape.

**Step 2:** to each point in $\{ \mathbf{q}_i = (x_{q,i}, y_{q,i}, z_{q,i}) \}_{1..5}$, an index $\alpha_i \in \{0, 1\}$ is associated. A point with $\alpha_i = 1$ represents the orthogonal projection of a contacting point assigned to the lower surface ($S^-$) i.e. for which $d_i(\mathbf{p}_i, \mathbf{q}_i) = -e$, those with $\alpha_i = 0$ are assigned to the upper surface ($S^+$) i.e. for which $d_i(\mathbf{p}_i, \mathbf{q}_i) = e$, where $e$ is the desired form error and $d(\mathbf{A}, \mathbf{B})$ denotes the directed Euclidean distance between the two points $\mathbf{A}$ and $\mathbf{B}$.

**Step 3:** verify that the gradient vector of distances with respect to motion parameters $\nabla d_i(\mathbf{p}_i, \mathbf{q}_i)$ are linearly independent. Otherwise, go to **Step 1**. The gradient vectors could be easily calculated using eq.(3.8)-(3.12).

$$\frac{\partial d_i}{\partial T_X} = (-1)^{\alpha_i} n_{i,X}$$  \hspace{1cm} (3.8)

$$\frac{\partial d_i}{\partial T_Y} = (-1)^{\alpha_i} n_{i,Y}$$  \hspace{1cm} (3.9)

$$\frac{\partial d_i}{\partial T_Z} = (-1)^{\alpha_i} n_{i,Z}$$  \hspace{1cm} (3.10)
\[
\frac{\partial d_i}{\partial \theta} = (-1)^{n_i}(q_i \times n_i).e_X \tag{3.11}
\]

\[
\frac{\partial d_i}{\partial \theta} = (-1)^{n_i}(q_i \times n_i).e_Y \tag{3.12}
\]

with \( n_i = (n_{i,X}, n_{i,Y}, n_{i,Z}) \) is the normal vector to the nominal shape at the point \( q_i \), \( a \times b \) refers to the cross product of the two vectors \( a \) and \( b \). \( e_X \) (resp. \( e_Y \)) is the unit direction vector of \( X \) (resp. \( Y \) direction).

**Step 4:** determine the Lagrangian multipliers \( \lambda^* = (\lambda_1^*, ..., \lambda_5^*) \) by solving the quadratic programming given in eq.(3.13).

\[
\min_{\lambda} \|G\lambda - b\| \tag{3.13}
\]

subject to \( \lambda \geq 0 \)

with \( G \) is a 6 \times 5 matrix such that \( G = (g_1, ..., g_5) \), \( g_i^T = ((-1)^{\alpha_i} \nabla d_i^T, 1) \) and \( b^T = (0, 0, 0, 0, 0, 1) \). If \( \|G\lambda - b\| > \epsilon \), where \( \epsilon \) is a predefined parameter, go to **step 1**. This step represents the resolution of the equation resulting from KKT conditions of the problem given in eq.(3.6).

**Step 5:** determine a nonzero null space vector \( p \) such that \( G^Tp = 0 \).

**Step 6:** verify that: \( p_b^THp_b > 0 \) where \( p_b \) is the vector composed of the first five elements of \( p \) and \( H \) is the Hessian matrix of the Lagrangian given in eq.(3.14).

\[
H = \sum_{i=1}^{5} (-1)^{\alpha_i} \lambda_i^* \nabla^2 d_i \tag{3.14}
\]

if this condition is not satisfied, go to **step 1**. Otherwise, calculate control points coordinates by setting \( p_i = q_i + (-1)^{\alpha_i} e_n_i \) for \( i \in \{1, ..., 5\} \).

**Step 7:** generate additional random points \( \{p_i\}_{i=6,...,N} \) such that \( p_i = q_i + \theta_i n_i \) with \( \{\theta_i\}_{i=6,...,N} \) is a set of randomly selected numbers in the domain \([-e, e]\) and \( \{q_i\}_{i=6,...,N} \) are arbitrary points belonging to the nominal shape.
3.4.3 Case study

The proposed method for the generation of reference data with non-vertex solutions was applied in the case of aspherical surfaces described by the monomial formulation given in the ISO 10110-12:2007 [15]. A set of 7729 points was generated as shown in figure 3.7.

![Generated reference data with a non-vertex solution](image)

Figure 3.7: Generated reference data with a non-vertex solution

The generated data have five contacting points only and a value of form error (the difference between maximum and minimum form deviations) that equals $2 \times 10^{-4} \text{mm}$. A MZ fitting algorithm is applied to the data in order to assess the form error. Figure 3.8 shows data residuals after MZ fitting. The fitted data contains five contacting points, three of them (red squares) lay on the upper surface ($S^+$) and the two remaining (blue squares) lay on the lower surface ($S^-$).
Chapter 3. Validation of fitting algorithms

Figure 3.8: Data residuals after MZ fitting. Only five contacting points (red and blue squares) exist. Red squares indicate points for which residuals are equal to \( +\epsilon \), blue squares indicate points for which residuals are equal to \( -\epsilon \).

3.5 Beyond data generation

The construction of a data set with the corresponding reference measurand value must not be the only concern during the design of reference data. Before proceeding to data generation, the scope and the characteristics of the software under test must be clearly identified. In other words, abilities claimed by the software should be determined so that task-specific data points are generated on their basis and the algorithm to test is not "disfavoured". In the following, some of these characteristics are identified.

3.5.1 Requirements on generated reference data

**Data extraction:** recorded data using some CMMs with tactile probes have equal spacing in X and Y directions. The included software, especially those using computation geometry techniques, may rely on the spacing distance in order to perform some calculations. In the case of testing such kind of software, a set of
reference data with varying spacing should not be considered.

**Number of points:** due to memory issues or for performance concerns, the number of points contained in the generated reference data must not exceed the maximum number of points that could be handled by the software.

**Uniqueness of the solution:** generated reference data must have a unique associated reference solution. This issue was outlined in [161] for the problem of maximum inscribed circle (figure 3.9). For fitting problems, reference data with multiple "associated features" will be problematic especially when these geometries will be taken as datum. These situations must be avoided when generating reference data.

![Figure 3.9: Maximum inscribed circle problem with multiple associated features](image)

**Error-free:** validation of metrology algorithms must be performed in "the perfect operator" approximation. This means that measuring errors (or any related errors) must not be embedded in constructed data. Adding measuring errors to the reference data might induce some difficulties since it would be difficult to tell whether inaccuracy comes from measurement or processing.

**Reference measurand value:** In some cases, fitting algorithms could not be used if values of form deviations are beyond a given value. Generally this is expressed as a percentage of a characteristic dimension of the artefact. For the case of LS, calculation of the Hessian matrix are often based on neglecting second order derivatives of residuals. For the case where the form deviations value become high, this approximation is no longer valid and the resulting Hessian matrix is not accurate. Therefore, the returned results are not accurate.
Form deviations distribution: As previously stated, the generated reference data must mimic real measured data. In this optic, the generated data must have the same layout regarding form deviations distribution. Figure 3.10 shows two reference data sets with the same reference output but with different form deviation distributions.

![Figure 3.10: Reference data with form deviation generated using: (a) fractional Brownian motion, (b) Gaussian noise](image)

In the case where only MZ or RMS values are sought, the algorithm under test might not be sensitive to the deviations distribution and both cases shown in figure 3.10 will be considered alike (from a processing point of view). However, if other statistical parameters are required on the top of MZ or RMS, the deviation distribution must be taken into consideration.

Initial position: Initial position of measured data relatively to the reference shape
highly affects performances of fitting algorithms. Most fitting software underlie inner routines to perform a rough alignment in order to make measured data as close as possible to the reference model. Some algorithms could perform fitting only if the input data are well aligned with the model. Moreover, the calculation of points projections, which is a crucial step in fitting algorithms, is highly affected by the initial position. These issues were experienced while comparing EPF and PDIP algorithms (see Chapter 2). In fact, during the comparison, three theoretical initial positions before fitting were considered: \( C_{d1}, C_{d2} \) and \( C_{d3} \). The first position \( C_{d1} \) was generated with an offset of 1 \( mm \) (for a value of \( R_{max} \) which is equal to 10 \( mm \)) in \( X, Y \) and \( Z \) directions with respect to the nominal model while keeping rotations at zero. The second position \( C_{d2} \) was generated with rotation of 5\(^\circ\) around \( X \) and \( Y \) directions while keeping translation to zero. The third position \( C_{d3} \) was generated by combing the two previous transformations (table 3.1). The next step was to determine the influence of the three conditions on the estimated MZ value.

<table>
<thead>
<tr>
<th>Translations</th>
<th>Rotations</th>
</tr>
</thead>
<tbody>
<tr>
<td>( T_X ) (mm)</td>
<td>( T_Y ) (mm)</td>
</tr>
<tr>
<td>( C_{d1} )</td>
<td>1</td>
</tr>
<tr>
<td>( C_{d2} )</td>
<td>0</td>
</tr>
<tr>
<td>( C_{d3} )</td>
<td>1</td>
</tr>
</tbody>
</table>

Table 3.1: Initial positions before fitting

Tables 3.2 and 3.3 give the values of the PV-MZ\(_{ref}\) obtained using EPF and PDIP (respectively with and without taking LS alignment as initial solution). For \( C_{d1} \), taking LS solution as initial alignment does not impact the results and both ways give accurate results. However, PV values are far from MZ\(_{ref}\) for \( C_{d2} \) and \( C_{d3} \).
### Table 3.2: Obtained PV values using EPF and PDIP without taking LS as initial alignment

<table>
<thead>
<tr>
<th></th>
<th>$C_{d1}$</th>
<th>$C_{d2}$</th>
<th>$C_{d3}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>PV</td>
<td>$7.99 \times 10^{-15}$</td>
<td>$11.66$</td>
<td>$8.28$</td>
</tr>
<tr>
<td>PV</td>
<td>$3.03 \times 10^{-14}$</td>
<td>$11.67$</td>
<td>$8.40$</td>
</tr>
<tr>
<td>PV</td>
<td>$1.87 \times 10^{-7}$</td>
<td>$4.64 \times 10^{-1}$</td>
<td>$8.14$</td>
</tr>
<tr>
<td>PV</td>
<td>$10^{-3}$</td>
<td>$10^{-4}$</td>
<td>$10^{-5}$</td>
</tr>
</tbody>
</table>

### Table 3.3: Obtained results using EPF and PDIP when taking LS as initial alignment

<table>
<thead>
<tr>
<th>$C_{d1}$</th>
<th>$C_{d2}$</th>
<th>$C_{d3}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>PV EPF-MZ$_{ref}$</td>
<td>PV PDIP-MZ$_{ref}$</td>
<td>MZ$_{ref}$</td>
</tr>
<tr>
<td>(mm)</td>
<td>(mm)</td>
<td>(mm)</td>
</tr>
<tr>
<td>$7.48 \times 10^{-15}$</td>
<td>$3.28 \times 10^{-8}$</td>
<td>$10^{-3}$</td>
</tr>
<tr>
<td>$2.72 \times 10^{-14}$</td>
<td>$7.23 \times 10^{-7}$</td>
<td>$10^{-4}$</td>
</tr>
<tr>
<td>$1.17 \times 10^{-13}$</td>
<td>$7.27 \times 10^{-9}$</td>
<td>$10^{-5}$</td>
</tr>
<tr>
<td>$2.27 \times 10^{-15}$</td>
<td>$4.27 \times 10^{-7}$</td>
<td>$10^{-3}$</td>
</tr>
<tr>
<td>$1.81 \times 10^{-14}$</td>
<td>$1.31 \times 10^{-7}$</td>
<td>$10^{-4}$</td>
</tr>
</tbody>
</table>
3.5.2 Reference data uncertainty

Uncertainty must be associated to reference values due to numerical inaccuracies. This will allow to set permissible limits used during the comparison of the obtained results to reference ones. In figure 3.11, a scenario of MZ\textsubscript{ref}, obtained MZ values using different algorithms as well as uncertainty on MZ\textsubscript{ref} are illustrated.

![Figure 3.11: Illustration of uncertainty on MZ\textsubscript{ref}](image)

\begin{align*}
\text{U(MZ}_{\text{ref}}) & : \text{Uncertainty on the MZ reference value.} \\
\text{MZ}_1,\ldots,\text{MZ}_4 & : \text{Minimum Zone values obtained using different algorithms under test.}
\end{align*}

Even if the obtained results using algorithms 2 and 3 (MZ\textsubscript{2} and MZ\textsubscript{3}) differ from the reference value, they are still acceptable. However, this is not the case for algorithms 1 and 4 and a measure of closeness to the exact solution must be established in order to determine the most performant algorithm in terms of accuracy among the two.

In [151], Linares et al. give some theoretical foundations on the assessment of numerical accuracy of numerical artefacts and numerical standards. The main objective is to be able to distinguish between the numerical uncertainties associated with the software from those associated with the numerical standards. Here, we present a method to estimate numerical uncertainty on MZ\textsubscript{ref} obtained using the method presented in Section 3.3.2. We suppose that we are attempting to construct a set of reference data with a targeted form error of MZ\textsubscript{ref,1} and we suppose also that the Linear Independence Constraint Qualification (LICQ) and dual feasibility conditions are satisfied (figure 3.3). Figure 3.12 illustrates the difference between MZ\textsubscript{ref,1} and the actual MZ\textsubscript{ref}. In fact, this difference is essentially due to numerical errors when calculating projection positions of control points, normal directions to the nominal shape at these projections and the positions of the control points.
Let us suppose that all coordinates as well as \( \text{MZ}_{\text{ref},t} \) are expressed using double precision. The actual value of form error can be formulated as in eq. (3.15).

\[
\text{MZ}_{\text{ref}} = 2 \times \sqrt{(x_p - x_q)^2 + (y_p - y_q)^2 + (z_p - z_q)^2}
\]

(3.15)

with

\[
x_p = x_q + \frac{1}{2} \text{MZ}_{\text{ref},t} n_x
\]

(3.16)

\[
y_p = y_q + \frac{1}{2} \text{MZ}_{\text{ref},t} n_y
\]

(3.17)

\[
z_p = z_q + \frac{1}{2} \text{MZ}_{\text{ref},t} n_z
\]

(3.18)

The uncertainty on \( \text{MZ}_{\text{ref}} \) denoted \( U(\text{MZ}_{\text{ref}}) \) could be calculated using propagation of uncertainty rules. Uncertainties of \( x_q, y_q, z_q \) and \( \text{MZ}_{\text{ref},t} \) depend on the accuracy of the machine architecture while uncertainties of \( x_p, y_p, z_p, n_x, n_y \) and \( n_z \) could also be analytically obtained using propagation rules.

It is to be noted that:

- In figure 3.12, only one contacting point was illustrated. The uncertainty on \( \text{MZ}_{\text{ref}} \) depends on each contacting point.

- For the case of aspherical lenses, this uncertainty was estimated to \( 10^{-14} \text{ mm} \). In [151], it is mentioned that "if it is assumed that the best practice in dimensional metrology operates at a numerical uncertainty on one part in \( 10^n \), then a minimal requirement is that software should be accurate in one part of \( 10^{n+1} \). Consequently, in order to assess such software, the numerical standards should be accurate in one part in \( 10^{n+2} \)." Regarding the uncertainty sought in the context of this project,
the uncertainty on the generated reference data is more than enough.

- A similar procedure for the case of LS similar to MZ fitting could not be obtained straightforwardly. We refer to [162] where this problem is partially addressed.

### 3.5.3 Difficulty vs. performance diagram

The requirements on the generated data presented in paragraph 3.5.1 are not exhaustive. In addition, most of these requirements could not be quantified, take the form deviation distribution as an example. Constructing data points with characteristics that meet software requirements is important but it does not respond to a number of questions that arise when testing algorithms such as how many data sets one must generate for instance.

The degree of difficulty is a concept introduced in [147] that aims at defining a quantity associated to each set of data. This quantity indicates at which extent the generated reference data set "challenges" the algorithm under test. Most of the relevant requirements could then be summarised in a single metric. In this way, the methodology of testing metrology software consists of generating data sets with an increasing difficulty number. For each set, the output of the software is assessed using performance measures. We can then plot the diagram of the performance measure against the degree of difficulty for each algorithm and then infer conclusions about the fit of purpose of the algorithm under test. Figure 3.13 shows the performance measure as a function of the degree of difficulty for a given algorithm (black dots fitted using the black curve). The regions corresponding to reference algorithms, operating algorithms as well as rejected algorithms are also transposed on the same graph.

Following this scheme, three major issues arise:

- First, an expression of the degree of difficulty depending on each fitting problem must be formulated,

- Second, a clear definition of the performance measure must be established. This quantity is highly subjective and must be defined according to each testing situation,

- Third, the borders between each algorithm profile (reference, operating and rejected algorithms) must be defined.
Since addressing the two latter issues is a subjective matter, a survey involving a number of 45 colleagues from NMIs, industry, academic, etc. was conducted. The profile of the questionee is given in figure 3.14. The category "others" corresponds to one metrology software editor and one commercial role in university.

The survey includes the following questions:

**Q1:** In metrology what are the characteristics of a good metrology software?

- Gives accurate results,
- Runs in a short time,
- Easy to use,
- Returns deterministic results,
- Others.

**Q2:** How important "Gives accurate results" is for you (5-Very important, 1-Not important)?
3.5. Beyond data generation

Figure 3.14: Survey questionee’ profile

Q3: How important "Runs in short time" is for you (5-Very important, 1-Not important)?

Q4: How important "Returns a deterministic response" is for you (5-Very important, 1-Not important at all)?

Q5: How important "Easy to use" is for you (5-Very important, 1-Not important)?

Q6: For fitting software or similar, what is the maximum execution time you allow (for an indicative data set with 1 million points).

Q7: If you did not select "Easy to use", could you please explain why this choice is not relevant to you?

The performance measure could be quantified based on the answers given in the first question. Other characteristics of a good metrology software such as stability, algorithm transparency or coding standards were suggested. However, these characteristics could not be quantified and hence will not be included in the performance measure. The general form of the performance measure denoted $\eta$ is given in eq. (3.19).

$$\eta = f(E, T, \Delta)$$
Chapter 3. Validation of fitting algorithms

with $E$ is the difference between the returned and the reference values, $T$ is the execution time and $\Delta$ is a binary variable that indicates whether the algorithm under test is deterministic or heuristic.

If we suppose that there is no interaction between $E$, $T$ and $\Delta$, the function $f$ in eq.(3.19) could be expressed as the convex combination given in eq.(3.20).

$$\eta = f(E, T, \Delta, \lambda) = \alpha_1 f_1(E, \lambda) + \alpha_2 f_2(T, \lambda) + \alpha_3 f_3(\Delta, \lambda)$$  \hspace{1cm} (3.20)

where $\alpha_1$, $\alpha_2$ and $\alpha_3$ are respectively the positive weights associated to each of the parameters $E$, $T$ and $\Delta$ with $\alpha_1 + \alpha_2 + \alpha_3 = 1$ and $\lambda$ is the difficulty number that will be defined later. The values of $\alpha_i$ will be chosen according to the profile of the algorithm under test (reference or operating algorithm). The expression of the functions $f_1$, $f_2$ and $f_3$ are given respectively in equations (3.21), (3.22) and (3.23).

$$f_1(E, \lambda) = \begin{cases} \frac{E(\lambda)}{E}, & \text{if } E \geq E_0(\lambda) \\ 1, & \text{otherwise} \end{cases}$$ \hspace{1cm} (3.21)

$$f_2(T, \lambda) = \begin{cases} \frac{T(\lambda)}{T}, & \text{if } T \geq T_0(\lambda) \\ 1, & \text{otherwise} \end{cases}$$ \hspace{1cm} (3.22)

$$f_3(\Delta, \lambda) = \begin{cases} 0, & \text{if heuristic} \\ 1, & \text{if deterministic} \end{cases}$$ \hspace{1cm} (3.23)

$E_0$ and $T_0$ are characteristic numbers that are function of the difficulty number. They could be seen as the recommended error and execution time for each value of the difficulty number. The functions $f_1$, $f_2$ and $f_3$ are bounded in order to be in the range $[0, 1]$. This way, the performance measure will also lay in the domain $[0, 1]$. A summary of the validation process is given as follows:

**Step 1:** Define a difficulty number $\lambda$ according to the fitting problem to address.

**Step 2:** Define the weights $\alpha_1$, $\alpha_2$ and $\alpha_3$ in function of the algorithm to validate.

**Step 3:** For each value of the difficulty number $\lambda$, define $E_0$ and $T_0$.

**Step 4:** Generate a number of reference data sets with a varying difficulty number and calculate the performance measure $\eta$. 

88
Step 5: Infer conclusions about the acceptance/rejection of the algorithm according to the obtained results.

As previously stated, the weights $\alpha_1$, $\alpha_2$ and $\alpha_3$ depend on the profile of the algorithm to test. Their values were selected based on the survey answers. For each questionee’s profile, the answers to the questions Q2, Q3 and Q4 was averaged. $\alpha_1$, $\alpha_2$ and $\alpha_3$ were then taken as weighted averages of the obtained answers as shown in table 3.4.

<table>
<thead>
<tr>
<th></th>
<th>Associated weights</th>
<th>Answers averages</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Reference algorithms</td>
<td>Operating algorithms</td>
</tr>
<tr>
<td>NMI</td>
<td>0.7</td>
<td>0.25</td>
</tr>
<tr>
<td>Industrials</td>
<td>0.2</td>
<td>0.5</td>
</tr>
<tr>
<td>Researcher</td>
<td>0.1</td>
<td>0.25</td>
</tr>
<tr>
<td>Others</td>
<td>0</td>
<td>0.05</td>
</tr>
<tr>
<td>$\alpha_1$</td>
<td>0.41</td>
<td>0.40</td>
</tr>
<tr>
<td>$\alpha_2$</td>
<td>0.24</td>
<td>0.28</td>
</tr>
<tr>
<td>$\alpha_3$</td>
<td>0.35</td>
<td>0.32</td>
</tr>
</tbody>
</table>

Table 3.4: Estimation of $\alpha_1$, $\alpha_2$ and $\alpha_3$ based on the survey answers. For example, in order to calculate the $\alpha_1$ corresponding to reference algorithms, the coefficients in the column "Reference data" are multiplied one by one by the coefficients in Q2 and then summed up to give $\alpha_1$.

As an example, in order to calculate the value of $\alpha_1$ for the case of reference algorithms, the corresponding coefficient for each category (NMI, Industrials, etc.) is multiplied by the corresponding answer to the question Q2 and the results are summed to give $\alpha_1$.

According to the values obtained for $\alpha_i$, both reference and operating algorithms associate similar importance to the accuracy of the results. Operating algorithm are
more stringent regarding the execution time. The deterministic aspect is more important in the case of reference algorithms.

The determination of the characteristic numbers $E_0$ and $T_0$ was also made using the survey answers. Suppose for the moment that the difficulty number lies in the domain $[\lambda_{\text{min}}, \lambda_{\text{max}}]$ and let $\lambda_0 \in [\lambda_{\text{min}}, \lambda_{\text{max}}]$. $\lambda_{\text{min}}$ (resp. $\lambda_{\text{max}}$) represents the easiest (resp. the most challenging) data sets to fit and $\lambda_0$ is a usual difficulty number. Table 3.5 gives the estimated values of $E_0$ and $T_0$ for $\lambda_{\text{min}}, \lambda_{\text{max}}$ and $\lambda_0$.

<table>
<thead>
<tr>
<th></th>
<th>$\lambda_{\text{min}}$</th>
<th>$\lambda_0$</th>
<th>$\lambda_{\text{max}}$</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Reference algorithms</strong></td>
<td>$T_0$ (sec)</td>
<td>10</td>
<td>120</td>
</tr>
<tr>
<td></td>
<td>$E_0$ (mm)</td>
<td>$10^{-14}$</td>
<td>$E_{0,\text{ref}}$</td>
</tr>
<tr>
<td><strong>Operating algorithms</strong></td>
<td>$T_0$ (sec)</td>
<td>1</td>
<td>30</td>
</tr>
<tr>
<td></td>
<td>$E_0$ (mm)</td>
<td>$10^{-14}$</td>
<td>$E_{0,\text{ope}}$</td>
</tr>
</tbody>
</table>

Table 3.5: Estimated values of $E_0$ and $T_0$ for $\lambda_{\text{min}}, \lambda_{\text{max}}$ and $\lambda_0$

The expressions of $E_0$ and $T_0$ as functions of the difficulty number could be obtained by fitting the values given in table 3.5 to a mathematical model. Here, we considered a second order polynomial. The values in table 3.5 could be interpreted in the following way: $E_0$ is required to be $10^{-14} \text{mm}$ in the case of "easy" reference data (reference data with the smallest difficulty number). This value corresponds to the uncertainty on the generated reference data. In other ways, we allow only errors that originate from the generation of reference data. $E_{0,\text{ref}}, E_{0,\text{ope}}$ are set by the validation entity. These values depend on the intended application of the algorithm under test. For "more challenging" reference data, one can consider the same value of $E_0$ as for the case of $\lambda = \lambda_0$ multiplied by a coefficient $\beta$ taken as 10 or 100 for instance.

Coming now to the difficulty number $\lambda$. As previously stated, the definition of the difficulty number highly depends on the problem to address. It comes naturally that the definition for the case of LS fitting will not be similar to MZ. The degree of difficulty is more complicated to define than the performance measure especially for the case of MZ fitting. No straightforward method could tell that a given MZ fitting problem is more challenging than another. Difficulty could be regarded as the difficulty to converge to a global optimum or how many local optima exist and how the barriers between them are.
3.5. Beyond data generation

However, experience has shown that there are some features than could help predict the difficulty of a MZ fitting problem. First, the nature of the solution (vertex or non-vertex). Convergence to a local optimum could be very slow in the case of a non-vertex solution and few algorithms can deal with non-vertex solution [6]. Second, the number of points in the data set could affect both execution time and result accuracy especially for algorithms where linear systems whose dimensions are correlated with the number of points must be solved or when points projection should be determined. In the previous section, it has been proved that the initial position compared to the finale/solution position have an effect on the returned MZ value. In most of the cases, the algorithms fail to converge if the inputted data is too shifted from the optimal position. Similarly to the performance measure, we can define a difficulty number of the case of MZ fitting denoted $\lambda_{MZ}$ as the convex combination given in eq.(3.24)

$$
\lambda_{MZ} = \beta_1 V + \beta_2 \frac{N}{N_0} + \beta_3 \frac{\Theta}{\Theta_0}
$$

(3.24)

where $V$ is equal to 1 if the solution is non-vertex and 0 otherwise, $N_0$ is the maximum number of points that could be processed by the algorithm under test and $\Theta_0$ is an estimation of the maximum initial alignment permitted by the algorithm under test. The coefficients $\beta_1$, $\beta_2$ and $\beta_3$ are positive and verify $\beta_1 + \beta_2 + \beta_3 = 1$. The value of these coefficients could be assigned based on which difficulty aspect is important to the entity conducting the test. The estimation of the maximum initial alignment $\Theta_0$ could be calculated by taken the norm of the vector of maximum permissible translation and rotation between the measured data and the optimal position.

The same logic could be applied in order to define a difficulty number for the case of LS fitting. The vertex/non-vertex solution could be replaced by the condition number of the Gram matrix resulting from the LS fitting. In [20] G.Forbes et al. has outlined the relation between the condition number of the Gram matrix resulting from LS fitting and the difficulty to fit measured data. Moreover, this was his motivation to suggest a new description of aspherical lenses that gives a nearly diagonal Gram matrix and hence a more easily solved linear system. The number of points as well as the initial alignment of the measured data with respect to the optimal position must also be taken in consideration. In addition to that, a measure of how large residuals must be taken in consideration since in this case, some algorithms sometimes experience difficulty in
converging from a poor initial estimate of the solution [163]. Another aspect that must also be taken in consideration is the representativity of measured data. In other words, the ratio between the measured area and the aperture of the artefact. The difficulty number for the case of LS fitting could be taken as in eq.(3.25).

\[
\lambda_{LS} = \gamma_1 \frac{\kappa}{\kappa_0} + \gamma_2 \frac{N}{N_0} + \gamma_3 \frac{\Theta}{\Theta_0} + \gamma_4 \frac{\epsilon}{\epsilon_0} + \gamma_5 r
\]  

(3.25)

where \( \gamma_1, ..., \gamma_5 \) are positive weights such that \( \sum_{i=1}^{5} \gamma_i = 1 \), \( N, N_0, \Theta \) and \( \Theta_0 \) have the same meaning as in eq.(3.24). \( \kappa \) is the conditioning number of the Gram matrix that could be calculated according to [20] and \( \kappa_0 \) is the maximum permissible conditioning number taken as \( 10^{14} \). \( \epsilon \) is the maximum value of form deviation in the data set, \( \epsilon_0 \) is the maximum permissible form deviation that could be handled by the algorithm under test and \( r \) is the ratio between the measured area and the artefact surface.

### 3.5.4 Case Study

In this subsection, the previously described procedure for algorithms’ validation is applied for testing a candidate algorithm that we will call "algorithm under test". The considered algorithm is intended to be an operating algorithm for the fitting of aspherical surfaces. It is claimed that:

- the algorithm is destined for LS fitting,
- the algorithm is able to process data up to 50,000 points,
- the maximum allowed position of the measured data with respect to the optimal position is \( (\pm 1\,\text{mm}, \pm 1\,\text{mm}, \pm 1\,\text{mm}, \pm 10^\circ, \pm 10^\circ) \) (the translations in the three directions as well as the rotations around \( X \) and \( Y \) axis, no need for the rotation around \( Z \) axis since the aspherical shapes are rationally symmetric),
- the maximum permissible form deviation is \( 1\mu m \),
- the algorithm under test returns deterministic results.

In Step 1, a difficulty number must be identified. In this case we use the one given in eq.(3.25) with \( \gamma_1 = 0.3, \gamma_2 = 0.2, \gamma_3 = 0.2, \gamma_4 = 0.2 \) and \( \gamma_5 = 0.1 \). The coefficients \( \alpha_i \) to define in Step 2 will be chosen according to the values in table 3.4 corresponding
to operating algorithms: $\alpha_1 = 0.40$, $\alpha_2 = 0.28$ and $\alpha_3 = 0.32$. In Step 3, the values of $E_0$ and $T_0$ will be selected according to values in table 3.5 corresponding to operating algorithms. We take $E_{0,\text{ope}} = 10^{-8} \text{mm}$ and $\beta = 100$. It is to be noted that since the difficulty number is expressed as a convex combination of quantities in the range $[0,1]$, $\lambda_{\min} = 0$ and $\lambda_{\max} = 1$. Nine data sets with varying difficulty numbers were generated. For each set, the performance number was estimated. The difficulty vs. performance diagram is shown in figure 3.15.

Figure 3.15: Performance measure in function of degree of difficulty for the "algorithm under test". In blue: the fitted line of the obtained results. In orange: the line defining the accepted limits of the performance measure.

The diagram presented in figure 3.15 is slightly modified from the one presented in figure 3.13. In fact, since in eq.(3.21)-(3.23) the choice of $E_0$ and $T_0$ is given a function of the difficulty number $\lambda$, the borders between the regions defining reference algorithms and operating algorithms are horizontal. In other terms, the borders are adapted for each value of the difficulty number. In this way, we need only a single value to determine the limits between accepted and rejected regions. In the case of the "algorithm under test",...
a value of $\eta_{lim} = 0.8$ was selected. Taking in consideration all the obtained results, one can conclude that the "algorithm under test" could be accepted as an operating algorithm. A similar procedure could be used to check if the "algorithm under test" could be considered as a reference algorithm.

### 3.6 Conclusion

The validation of metrology software was deeply investigated for aspherical and freeform surfaces. In fact, the development of assessment methods for metrology software is as important as the development of the software itself. An urgent call for the establishment of validation rules was initiated during a project led by the PTB and supported by the BCR in 1983.

A number of existing research aiming at defining rules and developing validation methods were presented. A focus was put on the methods developed by A. Forbes et al. that were presented in details. This choice was motivated by the fact that unlike other methods, this work gives detailed methods with strong theoretical basis for the construction of reference data for both LS and MZ fitting.

An extension was made for the case of MZ fitting by suggesting a method for the development of reference data with non-vertex solution. This consists of a special case of MZ reference data whose solution presents some specific characteristics. Data sets with non-vertex solutions occur in practice and few algorithms can deal with; which makes them essential for the validation of algorithms.

The scope of the work presented in this chapter went beyond the development of explicit methods for the generation of reference data to the discussion of their requirements. The main objective was to establish a complete validation procedure that takes in consideration most of the aspects of data sets that could affect the behaviour of the algorithm under test. Two metrics namely the degree of difficulty and a performance measure were defined for the case of reference algorithms as well as industrial algorithms and also for both LS and MZ fitting. The interest of this validation procedure was demonstrated through a case study. Some of the work presented in this Chapter was the subject of two conference papers respectively titled "Reference data simulation for $L_{\infty}$ fitting of aspheres" [164] and "A reference data based method for the evaluation of aspherical and freeform fitting algorithms" presented during the The 15th CIRP Con-
ference on Computer Aided Tolerancing (CIRP-CAT) and the 8\textsuperscript{th} edition of the Asian Society for Precision Engineering and Nanotechnology (ASPEN) conference respectively.

Further research is still needed for the development of reference data with non-vertex solutions in the case of freeform shapes. Also, rigorous arguments must be brought to help in the choice of the different parameters used in the definition of both the degree of difficulty and the performance measure.
4.1 Introduction

In the two previous chapters, two components of the metrology loop illustrated in figure 1 were discussed. MZ fitting algorithms for aspherical and freeform surfaces were extensively studied and a new algorithm was introduced. This was validated using reference data, benchmark data and measured data. A structured method for the validation of fitting algorithms was also presented. This method is based on reference data and involves algorithm performance metrics as well as reference data difficulty measures. The current chapter deals with artefact’ design. Artefacts or material standards are essential for the characterisation of measuring machines. A number of artefacts exist for different purposes but a few of them are dedicated to freeform and aspherical surfaces. The main objective of this chapter is twofold:

- First, a design of two thermo invariant material standards is suggested. The first artefact will be considered as a material standard for the assessment of MZ for aspherical shapes while the second will be designed for the evaluation of MZ of freeform surfaces. The two artefacts are made of a thermo-invariant material standard and manufactured using a MRF technology.

- Second, the different measuring machines held by the different partners involved in the project, the developed and validated algorithms as well as the designed artefacts in this chapter will be put together in order to conduct an inter-laboratory comparison. The inter-laboratory comparison will aim at characterising the artefact as well as assessing the accuracy of the different measuring machines.

In the next section, the design of the two thermo-invariant material standards is presented. Section 3 presents the different measuring machines that were used for the inter-laboratory comparison. In section 4, the comparison methodology is presented. In the last section, the different results are reported and discussed.

4.2 Thermo-invariant material standards

4.2.1 Literature review on existing material standards

A material standard is a realisation of the definition of a given quantity with a stated value and an associated measurement uncertainty [165]. Material standards allow the
determination of metrological capabilities of the instrument being calibrated.

There exist a number of artefacts for the calibration and testing of measuring machines. The majority of these artefacts are destined to CMMs and are widely commercialised. This includes, in the case of tactile CMMs, gauge blocks, balls, internal and external cylinders, step gauges, etc. (figure 4.1), whereas the use of balls is the common technique for the calibration of optical 3D systems (figure 4.2) [166].

Figure 4.1: Artefacts for the calibration of tactile CMMs [167]

Figure 4.2: Artefacts for the calibration of optical CMMs [167]

Only few freeform artefacts exist in literature. In [168], NPL suggests the freeform artefact shown in figure 4.3. The main goal behind its design is to challenge optical-based 3D measurement systems. Convex and concave forms are embedded in the artefact in order to identify the limitations of the measuring machines.

In [169], Savio et al. suggest a concept called the Modular Freeform Gauge (MFG) where the freeform surface is approximated using items with regular geometry and well calibrated on their dimensions and forms. The resulting artefact simulate as close as possible the shape of interest (figure 4.4).

Further freeform artefacts were suggested by PTB. The "Doppelsinusfläche", translated in English as "the double sinusoidal surface", is intended for the calibration of tactile CMMs (figure 4.5). The analytical equation of the surface is given in eq.(4.1).

\[
f(x, y) = \frac{A}{\pi} \sin\left(\frac{\pi}{B} x\right) \sin\left(\frac{\pi}{C} y\right)
\]  

(4.1)
Chapter 4. Artefacts’ design and inter-laboratory comparison

The artefact has a changing curvature along the surface and could be easily manufactured using CNC processing machines thanks to its analytical equation.

The "Shoe Model" given in figure 4.6 was also suggested by PTB. It consists of a real workpiece made of light wood and calibrated on a tactile CMM. The artefact is mounted on a stable base plate. Four circle marks are placed in the corners of the base plate in order to define the workpiece coordinate system which is considered as a reference system.

4.2.2 Design of the material standards

In the context of the European project freeFORM-15SIB01, a number of material standards were developed and two artefacts were considered for the inter-laboratory comparison. The first, called "Artefact I", is designed for the assessment of MZ value of aspherical surfaces while the second, "Artefact II", is destined to the evaluation of
4.2. Thermo-invariant material standards

Figure 4.5: The "Doppelsinusfläche" artefact [170]

MZ of freeform shapes. These two artefacts were developed in collaboration with Fudan University and Thales Angénieux. The two material standards are made of Zerodur® which is a glass-ceramic with a low thermal expansion coefficient less than $0.005 \times 10^{-6} K^{-1}$ [171] and manufactured using the MRF process described in Chapter 1.

Figure 4.6: The "Shoe Model" artefact [170]
4.2.2.1 Artefact I

Artefact I is an aspherical surface described using the ISO 10110-12:2007 formulation to which 'artificial' form error were added along the normal directions to the asphere. The added form is asymmetric with respect of the revolution axis and has the shape of nine steps as shown in figure 4.7. The approach for the combination of the steps and the aspherical surface is summarised in figure 4.8.

![Figure 4.7: Layout of the ‘artificial’ added form error to the aspherical shape](image)

The main idea behind this design is to materialise the upper and lower surfaces defining the enclosing envelope. In this way, the location of the contacting points will be approximately \( a \ priori \) known before the process of MZ fitting. The difference between maximum and minimum form deviations is equal to \( 7 \mu m \) which will be the "previously known" MZ value associated to the artefact. This value is selected according to the manufacturing process since it is the smallest amplitude that could be realised using the MRF technology. Note that the considered manufacturing process is able to manufacture surfaces with form errors lower than \( 7 \mu m \). The final surface represents neither axis of symmetry nor degrees of invariance as illustrated in figure 4.9. The nominal shape parameters of the asphere are given in table 4.1.
4.2. Thermo-invariant material standards

Figure 4.8: Construction of the thermo-invariant material standard for aspherical shape seen in the Y=0 plane

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>$R\ (mm)$</td>
<td>$9.127 \times 10^{10}$</td>
</tr>
<tr>
<td>$k$</td>
<td>-1</td>
</tr>
<tr>
<td>$a_4\ (mm^{-3})$</td>
<td>$1.278 \times 10^{-9}$</td>
</tr>
<tr>
<td>$a_6\ (mm^{-5})$</td>
<td>$7.922 \times 10^{-16}$</td>
</tr>
<tr>
<td>$a_8\ (mm^{-7})$</td>
<td>$-1.859 \times 10^{-18}$</td>
</tr>
<tr>
<td>$a_{10}\ (mm^{-9})$</td>
<td>$1.733 \times 10^{-21}$</td>
</tr>
</tbody>
</table>

Table 4.1: Nominal shape parameters of the aspheric surface used in "Artefact I"
Artefact II

Artefact II is a freeform surface that has applications in industry. It is incorporated in transparent screens of the oxygen mask embedded in firefighters’ helmets which allows them to have real-time information through augmented reality during action. The shape is described using the explicit polynomial equation given in eq. (4.2) with no degrees of invariance. The nominal shape values were selected according to the constraints imposed by the manufacturing process in terms of amplitude and slope. The shape of Artefact II is illustrated in figure 4.10 and the nominal shape values are given in table 4.2. A photograph of the manufactured artefacts is given in 4.11.

\[
\begin{align*}
z &= a_1(x^3 + y^3) + a_2(xy^2 + x^2y) + a_3(x^5 + y^5) + a_4(xy^4 + x^4y) + \\
& a_5(x^2y^3 + x^3y^2) - a_6x - a_7y - a_8
\end{align*}
\]

(4.2)
4.2. Thermo-invariant material standards

Figure 4.10: Design of the thermo-invariant material standard "Artefact II"

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>$a_1 (\text{mm}^{-2})$</td>
<td>$9.792 \times 10^{-7}$</td>
</tr>
<tr>
<td>$a_2 (\text{mm}^{-2})$</td>
<td>$4.940 \times 10^{-7}$</td>
</tr>
<tr>
<td>$a_3 (\text{mm}^{-4})$</td>
<td>$-6.310 \times 10^{-10}$</td>
</tr>
<tr>
<td>$a_4 (\text{mm}^{-4})$</td>
<td>$-3.086 \times 10^{-10}$</td>
</tr>
<tr>
<td>$a_5 (\text{mm}^{-4})$</td>
<td>$2.551 \times 10^{-10}$</td>
</tr>
<tr>
<td>$a_6$</td>
<td>$3.087 \times 10^{-4}$</td>
</tr>
<tr>
<td>$a_7$</td>
<td>$3.087 \times 10^{-4}$</td>
</tr>
<tr>
<td>$a_8 (\text{mm})$</td>
<td>$-6.876 \times 10^{-10}$</td>
</tr>
</tbody>
</table>

Table 4.2: Nominal shape parameters of the aspheric surface used in "Artefact I"

Figure 4.11: A photograph of the manufactured artefacts. The diameter of the two artefacts is $50 \text{mm}$
4.3 Measurements

Measurements were conducted by the French National Metrology Institute (LNE), the University of Nottingham (UNOTT, UK), Thales Angénieux (Thales-Agx, France), the Japanese metrology institute (NMIJ), Technical Research Centre of Finland (VTT), Institute of Plasma Physics (IPP, Czech Republic) and the Institute of Applied Optics at University of Stuttgart (ITO, Germany). Traceable ultra-high precision measuring machines with different probing systems were used for this comparison.

![Participating partners to the inter-laboratory comparison](image)

**Figure 4.12: Participating partners to the inter-laboratory comparison**

4.3.1 LNE - ultra-high precision profilometer

The architecture design of LNE’s ultra-high precision profilometer perfectly respects the Abbe principle. The machine is capable of performing nanometric measurements with both tactile and chromatic confocal probes. Movements in the three directions $X$, $Y$ and $Z$ are ensured using three independent high precision guiding axes equipped with encoders (figure 4.13). The movements of the ZeroDur® table, on which the object is posed is tracked by two independent and accurate Renishaw laser interferometers to
a nanometric level. The measuring machine is positioned on a massive anti-vibration system inside a room where the temperature (T) and the humidity (H) are controlled (T=20° ± 0.2, H=50% ± 5) and the pressure (P) is continuously measured [55, 172, 173].

![Figure 4.13: The LNE’s high precision profilometer](image)

**Figure 4.13: The LNE’s high precision profilometer**

### 4.3.2 Thales-Agx - Subaperture Stitching Interferometer (SSI)

The Subaperture Stitching Interferometer (SSI) was developed by QED Technologies (figure 4.14). This technique combines interferometry, precision motion and control. The machine consists of a six-axis motion control platform engineered in collaboration with Schneider Opticmachines on which a commercial Fizeau interferometer is embedded. A special software is used for motion control and mathematical analysis for the stitching process. The measuring range of the SSI can reach diameters up to 200 mm in all cases and slopes up to 90° (concave and convex). Notice that the calibration is automatic and does not require operator or time to perform [174, 175].
Chapter 4. Artefacts' design and inter-laboratory comparison

4.3.3 UNOTT - Zygo NexView NX2

The instrument is a Coherence Scanning Interferometer (CSI) (figure 4.15). It uses a broadband and spatially extended light source with an interferometric objective to generate low-coherence interference fringes as the instrument scans along the optical axis of the system. The surface topography of a sample is then derived from a combination of the envelope and phase of these interference fringes. The vertical scan is $150 \mu m$ with precision Piezo drive and $20 mm$ with extended scan. The surface topography repeatability is $0.12 nm$. The maximum data scan speed is $96 \mu m/s$. The step height repeatability is $0.1\%$, height response linearity $\leq 30 nm$ and step height accuracy is $0.3\%$ [176].
4.3.4 IPP - LuphoScan 260 HD

The LuphoScan 260 HD uses multiple wavelength single point optical probe that performs a spiral scan over the entire surface of the object to measure and produces high density 3D data (figure 4.16). An air-bearing spindle rotates the object while scanning and the sensor is moved radially and axially using linear stages. A rotary stage keeps the sensor normal to the object surface. The layout of movement stages provides high flexibility, even for uncommon surface shapes including a form measurement accuracy of better than $\pm 50 \, nm$ (3$\sigma$) up to 90° object slope and high reproducibility of measurement results [177].

![Figure 4.16: LuphoScan 260 HD measuring machine](image)

4.3.5 IPP - MarForm MFU 200 Aspheric 3D

The MarForm MFU 200 Aspheric 3D uses length optical sensor based on multiple concentric polar profiles by rotating the spindle (figure 4.17). These measuring points are used to generate a topography. Prior to concentric measurement, two linear profiles offset by 90° are measured across the zenith of the lens in a single sequence in order to compensate for geometric deviations. Topography measurement accuracy is better than $\pm 50 \, nm$ (PV) [178].
4.3.6 VTT - Multi sensor optical profilometer

The instrument is based on the measurement of sub-images using scanning white light interferometry and stitching them together to a high resolution image (figure 4.18). The horizontal displacements and rotation of the sample between sub-images are tracked using heterodyne laser interferometers. Straight and accurately tracked movements of the sample allow for correcting only the height difference of sub-images mathematically. In the profilometer there is also a chromatic confocal sensor for quick coarse scans. Here, two different interferometric objectives are used, both with 0.55× secondary lens, a 20× Mirau-objective which gives a pixel size of 0.881 μm and a slope limit of 18.9° and a 2.5× Michelson-objective which gives a pixel size of 7.10 μm and a slope limit of 3°.
4.3.7 NMIJ - UA3P-4000

NMIJ used a UA3P-4000 machine to measure the two artefacts. This machine is manufactured by Panasonic corporation and has a measurement range of 100 mm, 100 mm, 35 mm in X, Y and Z. The measurement performed using a diamond stylus with a radius of 5 µm. The measurement setup is shown in Figure 4.19. The artefacts were placed on a cylinder block. Three spheres were also placed on the cylinder block to touch the side of the artefact. The origin of the work coordinate system was determined from each sphere centre coordinate. The artefacts were measured in multiple lines along the x-axis of the workpiece coordinate system. The measurement area was a circle with a radius of 24 mm, the line step in the y-axis direction was 0.5 mm, the scanning pitch was 0.01 mm and the scanning speed was 3.0 mm/s.

Figure 4.19: UA3P-4000 measuring machine

4.3.8 ITO - NPMM200

ITO used the Nanopositionning and nanomeasuring machine NPMM-200 to measure the material standard (figure 4.20). The machine has a measurement range of 200 mm, 200 mm, 25 mm in X, Y and Z. The optical focus sensor is mounted on a metrological
frame made of Zerodur®. It was used in null mode meaning that the machine controlled the Z-position of the sample holder such that the sample surface was kept in focus. The metrological frame has six interferometers that determine the relative position of the sample holder. The artefact was moved at $6 \text{mm/s}$ in a meander type scan-path [179].

![Figure 4.20: ITO - NPMM200 measuring machine](image)

### 4.4 Comparison methodology

The two thermo-invariant material standards were carefully cleaned inside the LNE’s clean-room before proceeding to measurement. An appropriate cleaning process was used in order to eliminate contamination. Contamination is a phenomenon that causes surfaces to be soiled with contamination substances. These sub-micron substances must be controlled or eliminated in order to reduce measurement uncertainty. In fact, the presence of these undesirable substances on the surface of the artefacts causes the obtained MZ value to heavily deviate from the actual one (in absence of the particles). An iterative process was used for the cleaning of the artefacts. This consists of the following steps:
4.4. Comparison methodology

1. Triton and Foam,
2. Ultrasonic bath Acetone during 10 min,
3. Ethanol ultrasonic bath during 10 min,
4. Rinse with milliQ water,
5. Compressed air dry,
6. Control of the surface using an accurate optical microscope.

Figure 4.21 shows optical microscopy view of Artefact I before and after cleaning.

![Optical microscopy view of Artefact I](image)

A document recalling the cleaning and the handling procedure, the coordinate system associated to the artefacts, the calibration procedure, the area to scan, the format and the units of the recorded data and some other additional information was prepared and distributed to all partners involved in the comparison. Figure 4.22 shows the suggested measurement strategy for the two artefacts.

The recorded data were submitted directly to the HTR algorithm. Outliers were detected among recorded data and were removed manually prior to fitting. Since the measured data result in different areas, the recorded data were cropped out in order to have the same radius as shown in figure 4.23.
Chapter 4. Artefacts’ design and inter-laboratory comparison

Figure 4.22: Measurement strategy

Figure 4.23: Cropping out data illustrated on a measurement made by NMIJ
A summary of the used ultra-high precision measurement instruments as well as the number of collected data points are summarised in table 4.3.

<table>
<thead>
<tr>
<th>Partner</th>
<th>Measuring machine</th>
<th>Number of recorded data</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Artefact I</td>
</tr>
<tr>
<td>LNE</td>
<td>LNE ultra-high precision profilometer</td>
<td>247590</td>
</tr>
<tr>
<td>Thales Angx</td>
<td>Subaperture Stitching Interferometry (SSI) - QEDTechnologies</td>
<td>192771</td>
</tr>
<tr>
<td>UNOTT</td>
<td>ZYGO Nexview NX2 coherence scanning interferometer</td>
<td>360291</td>
</tr>
<tr>
<td>IPP</td>
<td>LuphoScan 260 HD MarForm MFU 200 Aspheric 3D</td>
<td>90646</td>
</tr>
<tr>
<td>VTT</td>
<td>VTT’s multi sensor optical profilometer</td>
<td>260830</td>
</tr>
<tr>
<td>NMIJ</td>
<td>UA3P-4000</td>
<td>160504</td>
</tr>
<tr>
<td>ITO</td>
<td>NPMM200</td>
<td>237151</td>
</tr>
</tbody>
</table>

Table 4.3: Measurements and number of collected data points for the comparison

In order to estimate the uncertainties associated to each obtained MZ value, the method described in Chapter 3 was used in the case where only one measurement is performed. In the case of multiple measurements, Type-A uncertainty was evaluated according to the GUM [134]. In order to characterize the MZ associated to the artefacts taking in consideration all measurements, the Key Comparison Reference Value (KCRV) denoted $MZ_{KCRV}$ is calculated using the weighted mean given in eq.(4.3).
Chapter 4. Artefacts’ design and inter-laboratory comparison

\[ MZ_{KCRV} = \sum_i w_i \cdot MZ_i \]  
\[ (4.3) \]

\( MZ_i \) is the mean MZ value obtained by each partner and \( w_i \) is given in eq.\( (4.4) \).

\[ w_i = C \frac{1}{[u(MZ_i)]^2} \]  
\[ (4.4) \]

\( u(MZ_i) \) is the uncertainty associated to \( MZ_i \) and \( C \) is given in eq.\( (4.5) \).

\[ C = \frac{1}{\sum_i [u(MZ_i)]^2} \]  
\[ (4.5) \]

The uncertainty of the weighted mean is calculated using eq.\( (4.6) \). A coverage factor \( k = 2 \) is used for the calculation of the expanded uncertainty.

\[ u(MZ_{KCRV}) = \sqrt{\frac{1}{\sum_i [u(MZ_i)]^2}} = \sqrt{C} \]  
\[ (4.6) \]

4.5 Results and discussion

In this section, the different results obtained from fitting measured data to the nominal shapes using the developed HTR algorithm are presented. Table 4.4 gives MZ results for the two artefacts.

<table>
<thead>
<tr>
<th>Partner</th>
<th>Measuring machine</th>
<th>Obtained values</th>
<th>Artefact I</th>
<th>Artefact II</th>
</tr>
</thead>
<tbody>
<tr>
<td>LNE</td>
<td>LNE ultra-high precision profilometer</td>
<td>Mean (( \mu m ))</td>
<td>6.305</td>
<td>0.782</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Expanded uncertainty (( \mu m ))</td>
<td>0.005</td>
<td>0.024</td>
</tr>
<tr>
<td>UNOTT</td>
<td>ZYGO Nexview NX2 coherence scanning interferometer</td>
<td>Mean (( \mu m ))</td>
<td>6.303</td>
<td>0.801</td>
</tr>
</tbody>
</table>
4.5. Results and discussion

<table>
<thead>
<tr>
<th>Method</th>
<th>Subaperture Stitching Interferometry (SSI) - QED Technologies</th>
<th>Mean ($\mu m$)</th>
<th>Expanded uncertainty ($\mu m$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Thales Angx</td>
<td></td>
<td>6.302</td>
<td>0.737</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>VTT</td>
<td>VTT’s multi sensor optical profilometer</td>
<td>Mean ($\mu m$)</td>
<td>Expanded uncertainty ($\mu m$)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>6.302</td>
<td>0.737</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>IPP</td>
<td>LuphoScan 260 HD</td>
<td>Mean ($\mu m$)</td>
<td>Expanded uncertainty ($\mu m$)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>6.268</td>
<td>0.773</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>IPP</td>
<td>MarForm MFU 200 Aspheric 3D</td>
<td>Mean ($\mu m$)</td>
<td>Expanded uncertainty ($\mu m$)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>6.350</td>
<td>0.753</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>NMIJ</td>
<td>UA3P-4000</td>
<td>Mean ($\mu m$)</td>
<td>Expanded uncertainty ($\mu m$)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>6.259</td>
<td>0.702</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ITO</td>
<td>NPMM 200</td>
<td>Mean ($\mu m$)</td>
<td>Expanded uncertainty ($\mu m$)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>6.342</td>
<td>0.809</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>KCRV ($\mu m$)</td>
<td></td>
<td></td>
<td>6.303</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>0.768</td>
</tr>
</tbody>
</table>
Table 4.4: Obtained MZ values for Artefacts I and II

4.5.1 Artefact I

For Artefact I, the obtained value of KCRV is equal to 6.303 $\mu m$ with an associated expanded uncertainty of 2 $nm$. Looking at each measurement separately, measurements from LNE result in a mean value of MZ which is equal to 6.305 $\mu m$ with an expanded uncertainty of 5 $nm$. Measurements from UNOTT result in a mean value of 6.303 $\mu m$ with an expanded uncertainty of 2 $nm$. These two sets of measurements could be considered as the most accurate. The measurements of Artefact I made by all partners present a good agreement as shown in figure 4.24. These results prove the capabilities of all participants to carry out measurement of aspherical surfaces with high accuracy.

Taking into consideration the obtained value of KCRV, a deviation from the theoretical value of MZ associated to Artefact I (7 $\mu m$) by 697 $nm$ could be seen. This deviation is essentially due to the manufacturing process. A small error in the estimation of the wear rate of the tool used in the MR process may lead to a significant form error. Because of its complex shape, the manufacturing of Artefact I took approximately nine hours while a normal MR cycle takes fifteen to forty-five minutes, which may explain this deviation.

The residual maps of the different measurements are given in table 4.5. They are represented in the frame associated to the nominal shape. The data points in the case of measurements made by IPP or ITO are rotated around the $Z$ axis compared to the design. This is due to the first positioning of the part before measurements. Nevertheless, this does not affect the final value of MZ since the nominal shape is rotationally symmetric.
4.5. Results and discussion

Figure 4.24: Obtained MZ values and associated uncertainties for Artefact I
Chapter 4. Artefacts’ design and inter-laboratory comparison

- UNOTT - measurement 1
- UNOTT - measurement 2
- UNOTT - measurement 3
- IPP - LUPHOS
- IPP - MFU
- VTT

![Heatmaps and images representing different measurements and comparisons between UNOTT and IPP in the context of artefacts’ design and inter-laboratory comparison.](image-url)
Table 4.5: Layout of the residuals’ maps
4.5.2 Artefact II

For Artefact II, a manufacturing defect was detected at the centre of the artefact in all measurements (figure 4.25). The area corresponding to this defect was cropped out and the comparison was performed on the region for which the radius is between 3 and 14 mm.

![Figure 4.25: Artefact II, manufacturer’ report](image)

The obtained values of MZ in the case of Artefact II are more dispersed. The obtained value of KCRV for Artefact II was estimated in the same way as Artefact I. The obtained value is equal to $0.768 \mu m$ with an associated expanded uncertainty of $16 nm$. The different MZ values as well as the obtained expanded uncertainties for Artefact II are plotted in figure 4.26. The measurements performed by LNE has the lowest expanded uncertainty ($24 nm$).

The different results prove the interest of the design of thermo-invariant material standards dedicated for MZ fitting. Results with good agreement were obtained using both tactile and optical probes. Nevertheless, the obtained results are still highly sensitive to the presence of outliers in the collected data. Hence, a clear preprocessing for
4.6 Conclusion

In this chapter, two thermo-invariant material standards were designed and manufactured. They were used for the performance’ assessment of measuring machines in the context of the project freeFORM-15SIB01. These two artefacts were selected among other ones that were developed within the same project.

The two developed thermo-invariant material standards are destined to aspherical and freeform surfaces. The main idea behind the suggested design of the first one is to materialise the upper and lower surfaces defining the MZ. In this way, the exact value of MZ could be known prior to measurement and thus inter-comparison could be conducted. The second artefact represents a sample freeform used in industry. The two artefacts were manufactured using a MRF process and made out of Zerodur® regarding its interesting thermal properties.

The artefacts were measured by the different partners involved in the project, the collected data were submitted to the developed HTR algorithm in order to estimate the corresponding MZ values and their associated uncertainties. Fitting was applied after manually removing outliers and cropping out the data sets to have the same aperture. Coherence between the different measurements was observed especially for Artefact I.
It was found that taking in consideration all measurements, the KCRV deviates by 697 nm from the theoretical one. The results obtained for Artefact II are more dispersed but represent good agreement with a KCRV of 0.768 µm and an associated expanded uncertainty of 16 nm.

The difference between the obtained mean value of MZ and the nominal one could be explained by:

- Uncertainties in the measurements. Since not all the measurements were performed in the same conditions (temperature, pressure, etc.). However, regrading the material of the artefacts, this could not justify the disparity in the results.

- The manufacturing process. According to the manufacturer report provided by Thales Angénieux, the process took approximately nine hours while a normal MR process takes fifteen to forty-five minutes. In addition, a small error in the estimation of the wear rate of the tool used in MR process may lead to significant form errors.

- The presence of outliers. In fact, the MZ fitting is sensitive to the presence of outliers even when using the most performant algorithms. The presence of outliers was noticed when inspecting different data sets.

Improvement of the overall comparison methodology could be made following the previous points. In fact, no guidelines exist for outliers removal. Some techniques imply using LS fitting prior to MZ and then remove the points with residual values beyond $k$ standard deviations from the mean value. Nevertheless, this approach requires LS fitting for freeform surfaces that could not be always available. Further work is still needed by the ISO committees to produce standards giving clear guidance for filtering measured data points collected on aspherical or freeform surfaces. The content of this chapter was the subject of a Journal paper, involving the different partners, titled: "Inter-laboratory comparison of minimum zone aspheric and freeform surface measurements" that was submitted to Precision Engineering journal.
Aspherical and freeform surfaces are a superior class of optical elements that are widely used in different domains regarding the advantages they allow compared to conventional surfaces. The increasing applications of these elements must be aligned with the available metrology capabilities which include ultra-high precision CMMs, fitting algorithms and thermo-invariant material standards. Enhancing metrology capabilities of aspherical and freeform surfaces at NMIs and DIIs was the aim of the European project FreeFORM-15SIB01 from which this work is part.

Each of these components was addressed during the present thesis. The design, the manufacturing and the measurement of aspherical surfaces are highly dependent on the used mathematical formulation for surface description. The objective of the first chapter was a comprehensive presentation of the different existing mathematical tools for this aim. Two additional topics were also outlined within the same chapter namely, the different manufacturing and measurement techniques of aspherical and freeform surfaces.

In the second chapter, the different issues related to the mathematical handling of measured data points were clearly identified. This mainly concerns data fitting in order to determine the form error expressed using Minimum Zone (MZ) value. The two widely used criteria for data fitting were extensively studied and compared. Also, the underlying minimization problems were formulated. A comprehensive literature review of existing methods for solving the MZ problem in the case of canonical shapes and freeform surfaces was conducted. As a result, two major classes of methods were considered for a comparison: smoothing techniques and constrained non-linear programming. EPF and PDIP are respectively two selected methods from the previous classes. They were compared using reference and measured data. For the sake of performance, an additional method, named HTR, was adapted from literature for the first time to the case of aspherical and freeform surfaces. The comparison of HTR to EPF was conducted using reference data, measured data and benchmark data. The obtained results show
the superiority of the newly introduced method in calculation time and results accuracy. A method for the estimation of the uncertainty of the returned value of MZ was established. This method combines the use of reference data and Monte Carlo simulation in order to derive the uncertainty on MZ values as a function of the uncertainties of the measuring machine.

Fitting algorithms are worthless without established validation procedures. Few research were conducted in literature giving clear guidance for the validation of metrology software. Therefore, a validation procedure for MZ and LS fitting was suggested and detailed. This method is based, besides to reference data, on the definition of two metrics, the performance measure and the difficulty number associated to the generated reference data. The method was applied to a case study for illustration purposes.

The previously developed tools were put together in the fourth chapter. Here, a design of two thermo-invariant material standards was suggested and manufactured. The first one is destined to aspherical surfaces while the second to freeforms. The two manufactured artefacts were used in order to conduct an inter-laboratory comparison between the partners involved in the project. The analysis of the collected data was performed using the validated HTR algorithms. The obtained results reveal coherence among the measurements and validate the capabilities of some partners to calibrate aspherical and freeform surfaces especially with small amplitudes (< 10 µm) with an uncertainty of few tens of nanometres.

During this work, MZ fitting of data points considers motion parameters only. In the future, investigating the determination of shape parameters might prove important. The existence of multiple solutions as well as the lack of methods giving good estimates of the initial solution of the minimisation problem remain the biggest hurdles. It is also a question of future research to investigate more robust fitting criteria than MZ. In fact, MZ criterion must be used with caution since it is highly affected by outliers and no standard outlier removal method exists.

Future research could examine the development of reference data with a non-vertex solution in the case of freeform surfaces. Also, rigorous arguments must be brought in order to refine the definition of the different metrics suggested in chapter 3. The obtained results during the inter-laboratory comparison proved to be coherent but the overall comparison could be improved. One improvement axis, that was previously discussed, is outlier removal methods. Some existing techniques make use of LS fitting
prior to MZ and then points with residuals beyond $k$ standard deviations from the mean value are removed. These techniques are limited by the availability of LS fitting even though the result could not be guaranteed to be exact. The possibility of establishing clear guidance for filtering measured data points collected on aspherical and freeform surfaces warrants further investigation.

The suggested methodology for the estimation of uncertainties on MZ values as a function of measuring machine errors was developed taken in consideration residuals with a Gaussian distribution. Further studies should investigate the behaviour of the uncertainty when using a different of distributions.


Titre : Algorithmes de références robustes pour la métrologie dimensionnelle des surfaces asphériques et complexes en optique

Mots clés : Métrologie dimensionnelle, Algorithmes, Optimisation, Surfaces complexes

Résumé : Les surfaces asphériques et complexes représentent une classe spécifique d'éléments optiques. Leur application a considérablement augmenté au cours des dernières années dans les systèmes d'Imagerie, l'astronomie, la lithographie, etc. La métrologie de ces surfaces reste un défi, en raison de la grande variété des données acquises et la traçabilité à l'unité SI-mètre. Cette métrologie devrait faire usage de la norme infinie; (Méthode d'association de type zone minimum ou Min-Max). Dans ce sens, il conviendrait d'associer une forme parfaite associée aux points traités sur la base d'un critère d'optimisation en minimisant l'écart maxi entre le nuage de points et la forme théorique cible. La complexité de cette méthode s'amplifie en fonction du nombre de points et de plus, les algorithmes souvent utilisés sont non-déterministes. Bien que le critère de type zone minimum soit bien adapté à des géométries simples (lignes, plans, cercles, cylindres, cônes et sphères), il reste un défi majeur pour des géométries complexes (surfaces asphériques et complexes). Cependant, l'objectif principal de la thèse est focalisé sur le développement des algorithmes de référence, destinés à des problématiques d’association de type Min-Max sur des surfaces asphériques et complexes. En conséquence, ces algorithmes doivent être robustes. Cela implique une étape supplémentaire de validation sur plusieurs données de référence. Ces algorithmes sont utilisés pour traiter des données mesurées en utilisant des instruments de mesure de très haute exactitude sur des étalons thermo-invariants, traçables, développés dans le cadre de ce travail.

Title : Robust Reference Algorithms for form metrology: Application to aspherical and freeform optics

Keywords : Dimensional metrology, Algorithms, Optimisation, Freeform

Abstract : Aspheres and freeform surfaces are a very challenging class of optical elements. Their application has grown considerably in the last few years in imaging systems, astronomy, lithography, etc. The metrology for aspheres is very challenging, because of the high dynamic range of the acquired information and the traceability to the SI unit meter. Metrology should make use of the infinite norm; (Minimum Zone or Min-Max method) to calculate the envelope enclosing the points in the dataset by minimizing the difference between the maximum deviation and the minimum deviation between the surface and the dataset. This method grows in complexity as the number of points in the dataset increases. Also, the involved algorithms are non-deterministic. Despite the fact that this method works for simple geometries (lines, planes, circles, cylinders, cones and spheres) it is still a major challenge when used on complex geometries (asphere and freeform surfaces). Therefore, the main objective is to address this key challenge about the development of Min-Max fitting algorithms for both aspherical and freeform surfaces as well as least squares fitting, in order to provide robust reference algorithms. The reference algorithms to be developed should be evaluated and validated on several reference data (softgauges) that will be generated using reference data generators. The developed algorithms were used for precessing data resulting from the measurement of traceable thermo-invariant material standards that was developed in the context of this work.