High magnification moiré interferometric measurement of crack tip deformation fields in stainless steels

This item was submitted to Loughborough University’s Institutional Repository by the/an author.

Additional Information:

- A Doctoral Thesis. Submitted in partial fulfilment of the requirements for the award of Doctor of Philosophy of Loughborough University.

Metadata Record: https://dspace.lboro.ac.uk/2134/11143

Publisher: © Filomena A. La Porta

Please cite the published version.
This item was submitted to Loughborough University as a PhD thesis by the author and is made available in the Institutional Repository (https://dspace.lboro.ac.uk/) under the following Creative Commons Licence conditions.

For the full text of this licence, please go to: http://creativecommons.org/licenses/by-nc-nd/2.5/
High Magnification Moiré Interferometric Measurement of
Crack Tip Deformation Fields in Stainless Steels

By
Filomena A. La Porta

A Doctoral Thesis

Submitted in partial fulfillment of the requirements
for the award of
Doctor of Philosophy of Loughborough University

September 1999

© by Filomena A. La Porta (1999)
To Mike
   and
   to my parents
Abstract—Improvements of fatigue and fracture models require an accurate experimental description of deformation fields in the small region ahead of the crack tip, covering a few material grains, in which the mechanisms driving the fracture process take place. For this purpose, a high-magnification moiré interferometer has been constructed. The set-up includes phase-stepping optics and a high resolution CCD camera (1.4 million pixels). By switching from laser light to white light illumination, it is possible to superimpose the deformation fields in exact registration with the underlying specimen microstructure. Displacement and strain fields are obtained by automated fringe analysis with respect to an undeformed reference state, over a sub-millimetre field of view. This technique has been applied here, for the first time to the best of our knowledge, to measure near tip surface deformation with underlying microstructure in cracked austenitic and duplex stainless steels subjected to single load and during fatigue. The fields obtained for monotonically loaded cracks were compared with existing theoretical models in a region of about 0.5×0.5 mm² ahead of the crack tip. In the experimental condition employed here, these models do not reproduce satisfactorily the experimental data. Influence of the microstructure on the strain distribution was observed. Elastic-plastic crack tip fields were measured during fatigue at the tip of a crack enabling a possible qualitative interpretation of the material response to applied stress. For an austenitic stainless steel, the dislocation distribution at the crack tip was also studied qualitatively by transmission electron microscopy. Evidence of an unzipping crack propagation model was found. In conclusion it has been demonstrated that the technique employed here is a powerful tool for a quantitative strain analysis over the region that is believed to play a crucial role in fracture and fatigue mechanisms.
Acknowledgements

The work presented in this dissertation was carried out in the Structural Integrity research group at the Department of Mechanical Engineering, Loughborough University, between October 1996 and September 1999 under the supervision of Professor J.M. Huntley, Dr. T.E. Chung and Professor R.G. Faulkner.

I wish to thank my supervisors for their invaluable help, in particular Professor Huntley and Dr. Chung, whose support went well beyond the research project. I am grateful to the University of Loughborough and the Departments of Mechanical Engineering and IPTME for the financial support, to the Avesta Sheffield Ltd. for providing the stainless steels specimens and some technical support.

I wish also to acknowledge Mr. S. Palmer for his help in my first steps into the grating replication technique, Mr. D. Britton and everybody in the Mechanical Engineering workshop for their constant help. I am indebted towards Mr. T. Atkinson for his availability and assistance in preparing the various etching potions, Mr. J. Bates for his guidance in the TEM technique, Mr. J. Guy for teaching me how to use the vacuum evaporator and Mr. V. Scothern for developing the photographs.

In particular I am grateful to Mr. R. Coggrave and Dr. R. Wildman, Martin and Beth for their continuous help, moral support and friendship.
# Table of Contents

Abstract ......................................................................................................................... i
Acknowledgements ........................................................................................................ ii
Table of Figures ............................................................................................................... vi
Introduction ...................................................................................................................... 1

## Chapter 1:

**Literature review I: Fatigue and fracture** ................................................................. 4

1.1 Macroscopic approach .............................................................................................. 4
1.2 Microscopic approach .............................................................................................. 9
   1.2.1 Dislocations in crystals .................................................................................... 10
   1.2.2 A crack as a dislocations distribution .............................................................. 12
1.3 Fatigue Crack Propagation models ........................................................................... 12
   1.3.1 Phenomenological models .............................................................................. 13
   1.3.2 COD related models ....................................................................................... 16
   1.3.3 Dislocation-based models .............................................................................. 17
   1.3.4 LCF-FCP models ........................................................................................... 21
1.4 Microscopic description of FCP ............................................................................. 22
   1.4.1 Crack initiation ............................................................................................... 22
   1.4.2 Crack propagation .......................................................................................... 26
   1.4.3 Submicroscopic cleavage and reverse shear model ......................................... 32
1.5 Analysis of traditional models ............................................................................... 34
1.6 Crack tip elastic-plastic strain field ......................................................................... 35

## Chapter 2:

**Literature review II: Experimental techniques** ......................................................... 39

2.1 Introduction .............................................................................................................. 39
2.2 SEM based techniques ............................................................................................ 40
2.3 Speckle metrology .................................................................................................. 41
Chapter 3:

Literature review III: Fringe analysis

3.1 Introduction................................................................. 56
3.2 Fourier Transform Method........................................................ 58
3.3 Phase-Stepping Method......................................................... 60
  3.3.1 Three-frame technique.................................................... 61
  3.3.2 Carré technique............................................................ 62
  3.3.3 Four-frame technique....................................................... 63
  3.3.4 Five-frame technique....................................................... 63
  3.3.5 "2+1" technique............................................................ 66
  3.3.6 General algorithms......................................................... 66
3.4 Phase unwrapping.............................................................. 68

Chapter 4:

Experimental set-up........................................................................ 81

4.1 Interferometer set-up............................................................. 81
4.2 Grating replication technique.................................................. 87
  4.2.1 Sub-master preparation...................................................... 87
  4.2.2 Specimen grating............................................................ 87
4.3 Interferometer alignment......................................................... 91
4.4 PZT calibration................................................................. 93
4.5 The rotation test.............................................................. 95
4.6 The Brazilian test.............................................................. 99
4.7 Stainless steel specimen preparation........................................ 103
Chapter 5:

Results I: Monotonically loaded crack tip fields ............................................... 109

5.1 Introduction ................................................................................ 110
  5.1.1 Materials ....................................................................... 110
  5.1.2 Experimental procedure ............................................... 113

5.2 316 stainless steel ...................................................................... 115

5.3 316 L stainless steel .................................................................... 124

5.4 2205 duplex stainless steel ........................................................ 128

5.5 HRR field ................................................................................. 133

5.6 Logarithmic type of singularity ........................................................ 137

5.7 Strain gradient effects ................................................................... 140

5.8 Conclusions .............................................................................. 146

Chapter 6:

Results II: Evolution of strain fields during fatigue .............................. 148

6.1. Introduction ............................................................................ 148
  6.1.1 Experimental procedure .................................................... 149

6.2 316 stainless steel .................................................................... 150

6.3 316 L stainless steel .................................................................... 162

6.4 2205 stainless steel .................................................................... 164

6.5 Comparison of the strain fields in the two steels ............................ 176

6.6 Conclusions ............................................................................. 176

Chapter 7: Conclusions and future work

7.1. Conclusions............................................................................ 178

7.2. Future work............................................................................ 181

References .......................................................................................... 183

Appendix

I. List of Publication ............................................................................ 192
Figure caption

Figure 1.1.1: System of co-ordinates used to describe the stress field at a crack tip................................................................. 5
Figure 1.1.2: Evaluation of the plastic zone size................................. 5
Figure 1.1.3: Dugdale model of a crack subjected to tensile loading.............. 7

Figure 1.2.1: Defect in a lattice represented by an extra half plane of atoms A. 
\(b\) is the Burgers vector and \(\tau\) the shear stress........................................ 10

Figure 1.2.2: Screw dislocation in a lattice, represented by the line AB........ 11

Figure 1.3.1: Crack length \(a\) as a function of the number of load cycles 
\(N\) for two different stress levels \(\sigma_1 > \sigma_2\).............................................. 13

Figure 1.3.2: Behaviour of the cyclic crack growth rate as a function of the 
alternating stress intensity \(\Delta K\). 
The dotted line represents the Paris-Erdogan law................................... 14

Figure 1.4.1: Wood’s model for fatigue crack initiation............................. 23

Figure 1.4.2: Shear stress against shear strain showing the behaviour of 
PSB........................................................................................................ 24

Figure 1.4.3: Formation of extrusions in the PSB model............................ 25

Figure 1.4.4: Plastic blunting model.......................................................... 26

Figure 1.4.5: Shear slip process explaining striations................................. 27

Figure 1.4.6: Tomkins model of double-shear band.................................... 28

Figure 1.4.7: Neumann model with parallel slip bands............................... 29

Figure 1.4.8: Unzipping model.................................................................. 30

Figure 1.4.9: Slip bands decohesion model. The darker lines represent 
tense localised slip bands................................................................. 31

Figure 1.4.10: Submicroscopic cleavage model........................................ 33

Figure 1.4.11: Reverse shear model............................................................ 33

Figure 2.3.1: Scattering of light from a rough surface .................................. 42

Figure 2.4.1: Schematic representation of Raman scattering........................ 43

Figure 2.5.1: Schematic principle of moiré interferometry............................ 44

Figure 2.5.2: Diffraction orders................................................................. 44

Figure 2.5.3: An arrangement to produce two incident beams onto 
the specimen.................................................................................. 48
Figure 2.5.4: Effects of a strain variation ........................................................ 55

Figure 3.2.1: Schematic of Fourier spectra of equation [3.2.4] ............................ 60

Figure 3.3.1: Wrapped phase map obtained with five-frame technique
The boxed region indicates the crack ................................................................. 65

Figure 3.4.1: Branch cut map obtained using the modified nearest
neighbour algorithm to unwrap the phase map of Figure 3.3.1.
The region of bad data is shown magnified .................................................... 72

Figure 3.4.2: Branch cut map obtained using the minimum
cut-length algorithm to unwrap the phase map of Figure 3.3.1.
The region of bad data is shown magnified .................................................... 73

Figure 3.4.3: Unwrapped phase map obtained with modified
nearest neighbour algorithm ............................................................................ 74

Figure 3.4.4: Unwrapped phase map obtained with minimum
cut-length algorithm ....................................................................................... 74

Figure 3.4.5: Striations in the wrapped phase difference map ................................. 76

Figure 3.4.6: Reference $y$ fringe pattern ........................................................ 76

Figure 3.4.7: Unwrapped phase difference. The line is the
cut placed to stop unwrapping across the crack ............................................ 78

Figure 3.4.8: Result of subtracting the fit of the reference
phase map to the unwrapped phase of Figure 3.3.1 ......................................... 78

Figure 4.1.1: Schematic diagram of the high magnification
moiré interferometer .......................................................................................... 82

Figure 4.1.2: Photograph of the four-axis mounts holding
the fibres $O F_1$ and $O F_2$ ................................................................................ 83

Figure 4.1.3: Photograph of the microscope head ............................................. 85

Figure 4.1.4: Photograph of the experimental set-up mounted in front
of the Instron. The black box is the Kodak camera.
Note the diffraction colours from the specimen ............................................ 86

Figure 4.2.1: Preparation of specimen grating:
a) the epoxy is transferred onto the specimen surface and
b) squeezed between specimen and sub-master. After curing
c) the sub-master is removed and d) a gold layer is evaporated
onto the specimen grating ............................................................................. 89

Figure 4.3.1: Fringe pattern obtained from the test grating
after the alignment procedure ........................................................................ 92
Figure 4.4.1: Second order polynomial fit of $S(V)$. The dots represent the experimental data and the line is the fit.

Figure 4.4.2: Relationship between phase shift $\varphi$ and applied voltage $V$.

Figure 4.5.1: Wrapped phase map corresponding to the rotated fringe pattern in Figure 4.3.1. The white colour corresponds to $\pi$ phase values, the black colour to $-\pi$.

Figure 4.5.2: Result of the unwrapping process applied to the phase map in Figure 4.5.1.

Figure 4.5.3: $u_x$ component of the displacement in function of $y$ obtained by rotating the test grating by 0.4° (diamonds) and 0.9° (circles). The continuous lines are fits.

Figure 4.6.1: The disc in compression test.

Figure 4.6.2: Stress-strain curve for the Perspex disc.

Figure 4.7.1: CT specimen geometry.

Figure 4.7.2: Pin loading grip.

Figure 4.7.3: $u_x$ and $u_y/\angle$ for several values of the gold thickness $t_g$.

Figure 5.1.1: Microstructure of the austenitic 316 stainless steels (field of view 250×150 μm²).

Figure 5.1.2: Duplex microstructure. Note the band structure, in which the ferrite (black grains) and the austenite (white grains) phase can be identified (field of view 250×150 μm²).

Figure 5.2.1: Wrapped phase corresponding to horizontal displacement component $u_x$ in the 316 steel. Black and white represent phase values of $-\pi$ and $+\pi$, respectively.

Figure 5.2.2: Fringe pattern corresponding to vertical displacement component $u_y$ in the 316 steel.

Figure 5.2.3: $u_x$ contour map for 316 stainless steel loaded at $K=35.8$ MPa$^{-1}$m. $u_x$ values are obtained from the wrapped phase map shown in Figure 5. The contour interval is 1 μm. Regions marked A and B show a change in direction of the contour lines at the grain boundaries. The circle marks the crack tip.

Figure 5.2.4: $u_y$ contour map for 316 stainless steel loaded at $K=35.8$ MPa$^{-1}$m. The contour interval is 1 μm. Regions marked A and B show a change in direction of contour lines at the grain boundaries. The circle marks the crack tip.

Figure 5.2.5: $S_{yy}$ contour map superimposed on the microstructure for 316 stainless steel. The contour interval is 0.01.
Figure 5.2.6: System of co-ordinates with origin at the crack tip..........................121

Figure 5.2.7: Schematic diagram to evaluate the influence of non
collimated beams. The dotted lines represent the collimated beams...122

Figure 5.3.1: Wrapped phase corresponding to vertical displacement
component $u_y$ in the 316L steel. Black and white represent
phase values of $-\pi$ and $+\pi$, respectively
(field of view $486\times370$ $\mu$m$^2$). ....................................................124

Figure 5.3.2: $\varepsilon_{yy}$ contour map superimposed on the microstructure
for 316 L stainless steel. The contour interval is 0.01 .......................126

Figure 5.4.1: Wrapped phase map relative to the 2205 steel loaded
at $K=35.8$ MPa$\sqrt{m}$. White colour corresponds to
$\pi$ phase value, black to $-\pi$.............................................................128

Figure 5.4.2: $\varepsilon_{yy}$ at the tip of a crack in the 2205 duplex steel, which
is characterised by a symmetrical distribution. As for
the previous contour maps, the contour interval is 0.01 .....................129

Figure 5.4.3: Reconstruction of the precrack in the duplex.........................132

Figure 5.5.1: Bi-logarithmic plot of $\varepsilon_{yy}$ against the radial distance
$r$ from the crack tip in the 316 steel, obtained for $\theta=30^\circ$ (circles),
45$^\circ$ (crosses), 60$^\circ$ (diamonds). The continuous and dotted line represents
the HRR fields, corresponding respectively to $\theta=45^\circ$ and $\theta=60^\circ$...134

Figure 5.5.2: Bi-logarithmic plot of $\varepsilon_{yy}$ against the radial distance $r$
from the crack tip in the 316 L steel, obtained for $\theta=30^\circ$ (circles), 45$^\circ$ (squares), 60$^\circ$ (diamonds).
The continuous and dotted line represents the HRR fields,
corresponding respectively to $\theta=30^\circ$ and $\theta=60^\circ$.....................134

Figure 5.5.3: Bi-logarithmic plot of $\varepsilon_{yy}$ against the radial distance $r$ from
the crack tip in the 2205 steel, obtained for $\theta=30^\circ$ (circles),
45$^\circ$ (squares), 60$^\circ$ (diamonds). The continuous and dotted line
represents the HRR fields, corresponding respectively to $\theta=45^\circ$ and $\theta=60^\circ$...135

Figure 5.6.1: Semi-logarithmic plot of $\varepsilon_{yy}$ against $r$ for $\theta=30^\circ$ (circles),
45$^\circ$ (crosses) and 60$^\circ$ (diamonds) for 316 steel ...............................138

Figure 5.6.2: Semi-logarithmic plot of $\varepsilon_{yy}$ against $r$ for $\theta=30^\circ$ (circles),
45$^\circ$ (squares) and 60$^\circ$ (diamonds) for 316 L steel ..........................138

Figure 5.6.3: Semi-logarithmic plot of $\varepsilon_{yy}$ against $r$ for $\theta=30^\circ$ (circles),
45$^\circ$ (squares) and 60$^\circ$ (diamonds) for 2205 steel. The continuous
line is the theoretical power law for $\theta=45^\circ$.................................139

Figure 5.7.1: Fit of $\varepsilon_{yy}$ for the strain-gradient-effect model to the contour
Figure 5.7.2: Fit of $\varepsilon_{yy}$ for the strain-gradient-effect model to the contour of Figure 5.4.2 (2205 steel) ...................................................... 144

Figure 5.7.3: Fit of $\varepsilon_{yy}$ for the strain-gradient-effect model to the contour of Figure 5.3.2 (316 L steel) ..................................................... 145

Figure 6.2.1: Elastic-plastic $\varepsilon_{yy}$ strain fields at the crack tip of austenitic 316 stainless steel for a) $N=0$, $K=35.8$ MPa$\sqrt{m}$, b) $N=2\times10^3$, $K=0$, c) $N=2\times10^3$, $K=35.8$ MPa$\sqrt{m}$. The strain values corresponding to the colour code are indicated in each colour bar. .................................................. 151

Figure 6.2.2: 316 microstructure underlying the strain fields of Figure 6.2.1c............. 153

Figure 6.2.3: Reconstruction of the crack growth. The different stages of fatigue are marked by arrows. The portions of circle indicate the regions C and D where TEM samples where taken. Regions marked A and B show intragranular crack growth ..................................................... 155

Figure 6.2.4: Dislocation distribution in region C. Instrument magnification 20k (field of view 3.5x2.5 $\mu$m$^2$). a) Grid array, b) loop patches.......................................................... 157

Figure 6.2.5: Dislocation lines changing direction at grain boundary. Instrument magnification 20k ................................................................. 158

Figure 6.2.6: Dislocation distribution in region D:
   a) disordered distribution, instrument magnification 20k (field of view $3.5\times2.5$ $\mu$m$^2$).
   b) only region of order observed in the form of grid array.
   Instrument magnification 33k (field of view $2.1\times1.5$ $\mu$m$^2$).............................. 159

Figure 6.2.7: 316 fracture surface macrograph. The precrack region is marked P and fatigue region is marked F. The shear lip is indicated by the letter S. The arrow indicates the direction of crack growth......................................................... 161

Figure 6.3.1: Microstructure of 316L showing branches B$_1$ and B$_2$ in the crack. (Field of view 100x76 $\mu$m$^2$).................................................. 163

Figure 6.3.1: Fringe pattern showing sectors of nearly parallel fringes. (Field of view 486x370 $\mu$m$^2$)............................................................. 163

Figure 6.4.1: Elastic-plastic $\varepsilon_{yy}$ evolution at the loaded crack tip of the 2205 duplex steel. $K=35.8$ MPa$\sqrt{m}$. a) $N=0$ b) $N=2\times10^3$, c) $N=5\times10^3$, d) $N=8\times10^3$, e) $N=11\times10^3$. The colour code shown in a) is the same for each map .................................................. 164
Figure 6.4.2: Elastic-plastic $\varepsilon_{yy}$ evolution at the loaded crack tip of the unloaded crack tip of the 2205 duplex steel. a) $N=0$ b) $N=2 \times 10^3$, c) $N=5 \times 10^3$, d) $N=8 \times 10^3$, e) $N=11 \times 10^3$. The colour code shown in a) is the same for each map. 168

Figure 6.4.3: $\varepsilon_{yy}$ at the unloaded crack tip of the duplex steel in function of $y$ for $x=50 \mu m$. 171

Figure 6.4.4: $\varepsilon_{yy}$ for duplex steel in function of $N$ at three different points in the field of view. 171

Figure 6.4.5: Schematic representation of the strain field evolution during fatigue at the crack tip of the duplex steel. 174

Figure 6.4.6: Macrograph of the 2205 fracture surface. The crack propagation direction is indicated by the arrow. 175
**Introduction**

Fatigue fracture occurs by initiation and propagation of a crack across a material. Although extensive studies have been carried out, the analysis of fatigue crack propagation (FCP) and fracture mechanisms in metals still constitutes an important area of investigations since the life of structures and machinery is FCP limited. Understanding this phenomenon will greatly benefit engineering design, both in terms of the development of new fatigue resistant materials and in terms of fatigue life prediction.

The subject has been mainly dealt within the two traditional approaches: macroscopic and microscopic. Within the macroscopic approach, the material is generally considered as a homogeneous and isotropic continuum. The effect of applied stresses on a cracked material or structure is then studied in terms of the macroscopic stress distribution around the crack tip. However, these methods are deduced on purely empirical basis and are thus unable to explain the behaviour of the cracks or predict materials response to the stress field.

On the other hand, the microscopic approach examines in detail the relationship between microscopic features within material and crack development. Cracking is described in term of dislocation movements, and little consideration is given to the influence of the structure surrounding the grains under consideration.

In spite of the amount of investigations carried out, the use of the macroscopic and microscopic approaches alone is not sufficient to completely understand and characterise the mechanisms causing a crack to propagate during fatigue and fracture.

Crack propagation processes, in fact, take place in a small volume of material at and near the crack tip where the stress concentration is highest, identified as the process zone. This region, covering a few grains of metal, will undergo severe elastic and plastic deformations when the material is subjected to external stresses. Plastic deformations and crack growth are caused by dislocation movements. On the other hand, the extent of these deformations is also dependent on the resistance offered by the material to the dislocation movements, as well as on the effects of second phase particles and grain boundaries. These mechanisms are rather complicated but they
need to be identified if the understanding of the basic processes that cause a crack to propagate is to be put on a rational rather than empirical basis.

The difficulties in modelling these mechanisms are due to the lack of accurate experimental data over this process zone. Obtaining a quantitative description of the near-tip deformation fields is therefore a crucial issue for advances in this topic.

Many experimental techniques have been developed with this respect, among which moiré interferometry has proved to be a powerful tool, due to its ability of providing whole field surface displacement, and therefore strain field, with high sensitivity and high spatial resolution.

In this work a high magnification moiré interferometer with a built-in microscope head for in-plane displacement measurement has been constructed. In particular, the moiré microscope described here features a high-resolution camera (1.4 million pixels) and enables displacement measurement by automated fringe analysis over a sub-millimetre field of view, which covers a few tens of material grains. The set-up is mounted on an X-Y-Z translational stage in front of a tensile testing machine, which allows measurements to be taken at different stages of fatigue life. Moreover, by switching from laser light to white light illumination, the displacement field can be imaged in exact registration with the underlying specimen microstructure.

This moiré microscope has been employed to determine the strain fields at the crack tip of austenitic and duplex stainless steels, and therefore a technique to prepare the surface of these particular specimens has also been developed.

The strain fields observed ahead of monotonically loaded cracks have been compared to existing theoretical models (Hutchinson, Rice and Rosengren (HRR), logarithmic-type of singularity and a recent model that takes into account strain gradient effects).

The development of strain fields during fatigue has also been studied and the results obtained by moiré microscopy have been supplemented by information obtained by transmission electron microscopy and fracture macrographs.

The work is structured as follows: Chapter 1 is a review of the basic concepts in Fracture Mechanics employed to model fatigue crack propagation mechanisms. The aim is to address the need of obtaining accurate experimental near-tip fields. Some macroscopic and microscopic FCP models and theoretical models for deformation fields ahead of monotonically loaded cracks are described.

Chapter 2 is a review of the most employed experimental techniques used for
deformation measurements at the crack tips including moiré interferometry. The aim is to explain the choice of moiré interferometry as experimental technique.

Chapter 3 is a brief review of fringe analysis methods, and its main purpose is to describe the phase measurement technique employed in the experiments subsequently described.

Chapter 4 is a description of the high-magnification moiré interferometer constructed. It contains also details of specimen preparation and the results of preliminary tests such as the rotation test and the disc in compression test.

Chapter 5 describes the experimental results obtained ahead of monotonically loaded cracks in stainless steels, while in Chapter 6 the evolution of elastic-plastic field during fatigue at the crack tip is shown and discussed.

Chapter 7 contains conclusions and observations on possible future work.
Chapter 1

Literature review I: Fatigue and fracture

The purpose of this chapter is to introduce basic concepts of Fracture Mechanics that are employed to model Fatigue Crack Propagation (FCP). Some traditional macroscopic and microscopic FCP models will then be described. This review does not attempt an exhaustive and complete survey of the existing literature on the subject, as the vast amount of literature in this field precludes this. The goals are rather to establish the foundations that will explain the necessity of an accurate knowledge of the near crack tip deformation fields and to provide the theoretical framework for the interpretation and discussion of the experimental data collected during this project.

1.1 Macroscopic approach

The macroscopic approach relates cracks in an elastic or elastic-plastic material or structure to the applied stress field. It is able to predict fracture stresses of cracked structures or the rate of propagation of a cyclically loaded structure. It is a basic assumption of this approach that the material is an isotropic continuum. This can be a limitation on their applicability to real materials and structures.

In the linear elastic fracture mechanics approximation (LEFM), the stress field at the tip of a crack is described by the expression\(^1\):

\[
\sigma_{ij} = \frac{K}{\sqrt{2\pi r}} f_{ij}(\theta) \quad i, j = x, y
\]

[1.1.1]

where \(f_{ij}(\theta)\) is an explicit function of the polar angle \(\theta\) (Figure 1.1.1) and \(r\) is the distance of an elemental area of material from the crack tip. The quantity \(K\) is defined
as the stress intensity factor and its magnitude depends on the applied stress, loading mode, specimen geometry and crack length.

Figure 1.1.1: System of co-ordinates used to describe the stress field at a crack tip.

According to equation [1.1.1] the stress components at the crack tip \( r = 0 \) approach infinity for any value of applied stress \( \sigma \). In practice, this does not happen as real materials, especially metals, yield (i.e. deform plastically) when the components exceed critical values, resulting in the presence of a plastic zone at the crack tip. The plastic deformation within this zone restricts the stress components to finite values.

Figure 1.1.2: Evaluation of the plastic zone size.
An approximate estimate of the size of the crack tip plastic zone \( r_p^* \) assumes that plastic deformation occurs when the vertical normal stress component \( \sigma \) exceeds the uniaxial yield stress \( \sigma_y \) (Figure 1.1.2). Substitution of \( \sigma_y \) in equation [1.1.1], yields the plastic zone radius \( r_p^* \) in the plane \( \theta = 0 \) for Mode I loading:

\[
r_p^* = \frac{K_i^2}{2\pi\sigma_y^2}
\]

This operative definition of plastic zone size, being dependent on \( K \), is valid only in the hypothesis of LEFM, when the plastic zone is small compared to the crack size. A better estimate for \( r_p \) (the Irwin correction\(^1\)) is based on the assumption that the plastic region surrounding the crack tip causes the crack to behave as if it were longer than its physical size. The displacements then are larger and the stiffness is lower than in the elastic case.

An effective longer crack is also the basis for the Dugdale model for a crack in tension\(^2\). A crack in a semi-infinite plate of thickness \( t \) is represented by a slit of initial length \( 2c \). The nominal stress \( \sigma \) opens the crack and the slit extends to a length \( 2a \). A uniformly distributed internal tension \( S \), acting only on parts of the slit \( r_p = (a-c) \), partially constrains the extending and opening of the crack. If \( S \) is equated with \( \sigma_y \), the internal tension closely simulates the local support to the wedge forces \( \sigma_y \). The plastic zone where the material is not really cracked and can still bear the yield stress is therefore defined by:

\[
\frac{r_p}{a} = 2\sin^2 \frac{\beta}{2}
\]

where \( \beta = \pi\alpha\sqrt{2} \sigma_y \).
To obtain a complete description of the plastic zone shape its dependence on the polar angle $\theta$ should also be considered and a yield criterion should be imposed. A number of yield criteria have been proposed for multiaxial stressing.

The Tresca criterion states that yielding occurs if

$$\tau_{max} > \sigma_{y}/2$$  \[1.1.4\]

where $\tau_{max}$ is the maximum shear stress.

The Von Mises criterion is expressed in terms of the principal stresses and is given by:

$$\left(\sigma_1 - \sigma_2\right)^2 + \left(\sigma_2 - \sigma_3\right)^2 + \left(\sigma_3 - \sigma_1\right)^2 = \sigma_{y}^2$$  \[1.1.5\]
Many studies have been devoted to a more accurate prediction of the plastic zone shape. The experimental verification of these models is difficult, since elastic and plastic strains cannot be easily separated. Furthermore, the experimental observations are usually restricted to the specimen surface, where the stress state is plane stress in character. In thick structures the fracture behaviour may be more appropriately described as plane strain.

Hahn and Rosenfield (1965) studied the plastic zone in silicon steel using an etching technique to delineate the area of plastic yielding. They found that none of the theoretical descriptions predicted the experimentally observed plastic zone accurately. However, the Dugdale model gave a useful description of plastic zone shape and size under plane stress conditions. It was also observed that when the Dugdale description was appropriate, it was applicable to an unloaded crack as well as to a crack in tension. Moreover, they observed that internal tensions acting in the plastic zone could be determined if the strain field distribution at the crack tip was known.

For a description of crack behaviour, an energy balance approach can be considered as an alternative to the stress intensity factor approach. The energy criterion for crack growth states that growth can occur if the energy required for the crack to advance by an infinitesimal increment $da$ can be delivered by the system to the crack tip. If $U$ is the potential energy contained in the specimen, the energy release rate $G$ per unit thickness during the crack extension in a linear elastic material is:

$$G = \frac{dU}{da} = \frac{\pi \sigma^2 a}{E}$$  \hspace{1cm} [1.1.6]

where $E$ is the Young's modulus. In the presence of appreciable crack tip plasticity, $G$ cannot be determined by assumption of a linear elastic stress field. In this case the use of the $J$-Integral provides a means to determine such a quantity:

$$J = \int_r \left( Wdy - T \cdot \frac{\partial u}{\partial x} \right) ds$$  \hspace{1cm} [1.1.7]
where $\Gamma$ is an arbitrary path around the crack tip, $J$ being the path independent integral. $W$ is the strain energy per unit volume and using the summation convention is:

$$W = \int_0^L \sigma_y d\varepsilon_y$$  \[1.1.8\]

$T$ is the traction vector perpendicular to $\Gamma$ with components $\sigma_{ij}n_j$, where $n_j$ are the components of the unit vector normal to $\Gamma$ and using the summation convention. $u$ is the displacement vector and $ds$ is an element of $\Gamma$. Thus $J$ represents the change in potential energy per unit thickness for a crack of extension $da$:

$$J = -\frac{dU}{da}$$  \[1.1.9\]

The major limitation to the use of the $J$ integral is imposed by the assumption that the material is non-linear elastic. This precludes its use in situations where unloading occurs in real materials, as in fatigue. A second limitation is that the plastic region around the crack tip must be small with respect to the size of the region in which $J$ controls the stress field.

### 1.2 Microscopic approach

To study the problems connected with the presence of a crack in a structure, the approach adopted in the material science field has been essentially of microscopic nature, involving mainly the analysis of the role of dislocations during fatigue. Dislocations, being responsible for plastic deformation in crystals, play a vital part in the development of the plastic zone that is present at the tips of propagating cracks.
1.2.1 Dislocations in crystals

The basis of the theory of plastic deformation in crystalline solids lies in the observation that the high theoretical strength necessary for one atomic plane to slide over another, or 'slip', is rarely attained in practice. To allow slip at much lower stress levels, the existence of lattice defects has been postulated\textsuperscript{11-13}. If an extra half plane of atoms (indicated with A in Figure 1.2.1) is introduced in the lattice, the breakage of atomic bonds necessary for slip can be restricted to the immediate vicinity of the bottom edge of this half plane of atoms, called the dislocation line\textsuperscript{14}. As the dislocation line moves through the crystal, the atomic bonds break consecutively along the slip plane, thus avoiding the necessity to break all the bonds across the slip plane simultaneously.

![Figure 1.2.1: Defect in a lattice represented by an extra half plane of atoms A. b is the Burgers vector and \( \tau \) the shear stress.](image)

The edge of the extra half plane of atoms defines this type of defect, hence its name, 'edge dislocation'. A series of atom-to-atom steps along lattice vectors that generates a closed loop about any location in the lattice is called the Burgers circuit. In a perfect
lattice, the Burgers circuit beginning at a point and advancing an equal and opposite number of lattice vectors in the horizontal and vertical directions will return to its starting point. If the lattice contains an edge dislocation the circuit does not close, and the vector needed to close it is called Burgers vector $b$ (Figure 1.2.1). For an edge dislocation, $b$ is oriented normal to the line defect and plastic flow is ordinarily restricted to the plane defined by the dislocation line and its Burgers vector.

A second variety of dislocation which can result in the displacement of one part of the crystal relative to another is the 'screw dislocation'. This is defined by the line $AB$ in Figure 1.2.2, with the Burgers circuit about the defect assuming a helix shape. The screw dislocation is parallel to its Burgers vector $b$ and so is the slip direction, but the direction of movement of the screw dislocation is perpendicular to $b$. The movement of this type of dislocation during plastic deformation, or cross-slip, occurs in the same direction but on different planes if for example an obstacle such as a precipitate is encountered.

Figure 1.2.2: Screw dislocation in a lattice, represented by the line $AB$. 
1.2.2 A crack as a dislocations distribution

The role of dislocation theory in describing fracture processes is twofold.

Firstly, the presence of dislocation in crystal enables to model fracture initiation
and plastic relaxation phenomena associated with the presence of cracks and
microscopic inhomogeneities.

Secondly, the use of dislocation concepts in macroscopic fracture mechanics
enables cracks in a continuum to be represented and mathematically analysed as a
dislocation distribution.

In the latter category, Eshelby 15.16 (1956, 1960) visualised a crack as a region
of the continuum $R$ in which an amount of material is initially inserted in the form of a
continuous dislocation distribution $D$, producing a stress $\sigma$. If an external stress $\sigma_E$ is
applied and $D$ adjusted so that the resulting stress over $R$, $\sigma + \sigma_E$, is zero, then $R$ is
free from traction. Consequently, the medium can be cut over $R$ and the inserted
material can be removed without disturbing the surrounding field. The dislocation
distribution $D$ thus represents the shape of a thin crack or cavity in the region $R$. By
using the mathematical description of dislocation, Bilby, Cottrell and Swinden 17 (1963)
were able to describe a model for a crack able to extend analogous to the Dugdale
model (equation [1.1.3]).

1.3 Fatigue Crack Propagation models

Ideally, it would be desirable to establish a law that could predict the fatigue
crack propagation (FCP) growth rate. The macroscopic approach which describes FCP
and identifies the parameters controlling this process is based purely on empirical
observations. When cyclic loads are applied to a specimen with a sharp crack 14, the
resulting change in crack length $a$ is monitored and recorded as a function of the
number of load cycles $N$. A typical plot representing the crack length behaviour is
shown in Figure 1.3.1:
Figure 1.3.1: Crack length $a$ as a function of the number of load cycles $N$ for two different stress amplitudes $\Delta \sigma_1 > \Delta \sigma_2$.

The crack length is seen to increase at an increasing rate as the number of loading cycles $N$ increases. Other factors affecting crack propagation rates are the stress amplitude, specimen geometry, microstructure and environment. A predictive law that takes into account all those effects would be desirable. Up until now, FCP models are mostly phenomenological and not predictive, being built to fit the trend of experimental data. Dislocation-based models (which are based on more fundamental mechanisms) have also been proposed. Some of these models will be described in the following sections.

1.3.1 Phenomenological models

The rate of propagation of a crack during fatigue has been related to the amplitude of stress intensity variation, $\Delta K$, by the Paris-Erdogan law:

$$\frac{da}{dN} = C(\Delta K)^m$$  \[1.3.1\]

where $C$ and $m$ are experimentally determined constants, which include material variables and environmental factors. The plot of $\log(da/dN)$ as a function of $\log(\Delta K)$ is sigmoidal in shape (Figure 1.3.2), but Paris-Erdogan law provides a sufficiently good
correlation for the middle range of stress intensity amplitudes. Deviations from Paris-Erdogan law are seen at the extremes of higher and lower $\Delta K$. $da/dN$ is underestimated at higher rates, when the crack growth approaches infinity when the maximum stress intensity equals the critical fracture stress intensity value $K_c$. For lower rates, Paris-Erdogan law overestimates the growth rate. Below a threshold value for $\Delta K$, $\Delta K_{th}$, the crack does not propagate with stress fluctuations.

Figure 1.3.2: Behaviour of the cyclic crack growth rate as a function of the alternating stress intensity $\Delta K$. The dotted line represents the Paris-Erdogan law.

A very large number of empirical relations exist in the literature considering the deviation from the power law behaviour described in equation [1.3.1]. Forman\textsuperscript{21} (1967) proposed a modified power law that contains not only $\Delta K$ as a characteristic parameter, but also takes into account the effects of load ratio $R$ (defined as the ratio of minimum and maximum stress intensity factors) on crack propagation rate. This model takes into consideration the instability of crack growth when the value of the maximum stress intensity factor approaches the fracture toughness of the material $K_c$. According to Forman, the crack growth rate is given by:
\[ \frac{da}{dN} = \frac{C(\Delta K)^m}{(1-R)K_c - \Delta K} \]  

[1.3.2]

where the exponent \( m \) was found to be nearly three from different sets of data. The dependence of FCP rate on \( R \) was also suggested by Saxena\textsuperscript{22} (1979), who proposed an equation consisting of three terms associated to the three regimes shown in Figure 1.3.2:

\[ \frac{1}{da/dN} = \frac{A_1(R)}{(\Delta K)^{m_1}} + A_2(R)\left[ \frac{1}{(\Delta K)^{m_2}} - \frac{1}{[K_c(1-R)]^{m_3}} \right] \]  

[1.3.3]

\( A_1(R) \) and \( A_2(R) \) are functions controlling the load ratio dependencies in the regions \( A \) and \( B \) of Figure 1.3.2. \( m_1 \), \( m_2 \) and \( K_c \) are constant values obtained from experimental data.

Another equation derived by Roberts and Erdogan\textsuperscript{23} (1967) by considering the crack growth primarily as a function of \( K_{\text{max}} \) or \( \Delta K \), and taking into account the plastic zone length \( \lambda \) is:

\[ \frac{da}{dN} = B(1+\beta)^{2\alpha_1}(\lambda \Delta K)^{2(\alpha_1+\alpha_2)} \]  

[1.3.4]

where \( \beta = K_{\text{max}}/\Delta K \), and \( \alpha_1 \), \( \alpha_2 \) and \( B \) are determined from experimental data.

In the literature most of the studies on fatigue crack propagation have been conducted mainly in the linear region in Figure 1.3.2, and the most representative models proposed for determining crack propagation rate in this linear region are the Crack Opening Displacement (COD) related models, the dislocation-based models and the LCF related models. These approaches will be briefly described in the following sections.
1.3.2 COD related models

The concept behind COD models is that the magnitude of the observed cyclic crack tip blunting and shapening is directly related to the crack growth increment and to the striations present on the fracture surface\(^2\). In the system of co-ordinates defined in Figure 1.1.1 and in the elastic case, the COD is defined as:

\[
COD = \frac{4\sigma}{E} \sqrt{a^2 - x^2}
\]

where \(a\) is the crack length and \(E\) the Young’s modulus. Including plastic corrections, the COD can be expressed as:

\[
COD = \frac{4\sigma}{E} \sqrt{(a + r_p^*)^2 - x^2}
\]

At the crack tip \(x = a\) and since \(r_p^* \ll a\), the crack tip opening displacement (CTOD) is written as:

\[
CTOD = \frac{4\sigma}{E} \sqrt{2ar_p^*}
\]

By using the expression for \(r_p^*\) given in equation [1.1.2] and observing that for a crack in an infinite plate \(K_\tau = \sigma \sqrt{\pi a}\), it is obtained that in LEFM the CTOD is related to the stress intensity factor by the equation\(^1\):
CTOD = \frac{4K^2}{\pi E\sigma_{ys}} \quad [1.3.6]

Hence from equation [1.3.6], CTOD is proportional to the stress intensity factor. Since \( \frac{da}{dN} \) in the Paris regime appears to be also proportional to the stress intensity factor, it is deduced that it may be also proportional to the CTOD. Considering the cyclic opening displacement, it follows that for static or slow growing cracks, the fatigue crack propagation rate can be expressed by the equation:

\[
\frac{da}{dN} \propto CTOD \propto \frac{\Delta K^2}{E\sigma_{ys}}. \quad [1.3.7]
\]

1.3.3 Dislocation-based models

This last category of models can be also classified as critical values models, since they assume damage ahead of the crack tip. By means of energy balance approach, a Paris type of crack growth equation is derived\(^2\). Such models essentially consider the plastic work done per unit distance of crack growth \( U \), which can be shown to be equal to

\[
U = D_c \sigma_{ys} \quad [1.3.8]
\]

\( D_c \) is the critical value of the accumulated displacement that must be exceeded for the crack to advance by an incremental length \( \Delta a \). An accumulated work criterion is hence equivalent to an accumulated displacement criterion and the FCP rate is given by:
\[
\frac{da}{dN} = \frac{\Delta K^4}{8\pi\mu\sigma_y^2 U} \tag{1.3.9}
\]

where \(\mu\) is the shear modulus.

The plastic displacement can be derived from the BSC-Dugdale\textsuperscript{2,17} model of the plastic zone based on the description of a crack as a continuous dislocation distribution. Assuming that the plastic displacement \(D(x)\) at a point located at a distance \(x\) in front of the crack tip depends linearly on \(x\):\textsuperscript{25}

\[
D(x) = \left(\frac{2K^2}{\pi \sigma_y \mu}\right) \left[1 - \left(\frac{\pi \sigma_y x}{K^2}\right)\right] \tag{1.3.10}
\]

the accumulated displacement \(D_c\) is given by the equation:

\[
D_c = \sum_n |D_n| \tag{1.3.11}
\]

where \(D_n\) is the displacement increment that occurs during the \(n^{th}\) cycle. If equation \[1.3.10\] is substituted in equation \[1.3.11\] with the assumption that \(\Delta a\) is smaller than the plastic zone size \(r_p\), so that the summation in equation \[1.3.11\] can be replaced with an integration, the following equation is obtained for a push-pull loading at the point \(x_0\):

\[
D_c = \left(\frac{2K^2}{\pi \sigma_y \mu}\right) \left[2x_0 - \left(\frac{\pi \sigma_y x_0}{K^2}\right)x_0 r_p\right] \tag{1.3.12}
\]
Using equations [1.3.8] and [1.3.12], the FCP rate in equation [1.3.9] can then be rewritten as:

\[
\frac{da}{dN} = \frac{\Delta K^4}{8\pi^2 \mu \sigma_y^3 D_c} \tag{1.3.13}
\]

This equation is valid for a crack in a perfect plastic solid in which the centre of the plastic zone is located at \( x_0 = r_p / 2 \). When the material is not perfectly plastic but is characterised by a strain hardening exponent \( n \neq 0 \), the centre of the plastic zone is at:

\[
x_0 = r_p \frac{(1-n)}{2 (1+n)} \tag{1.3.14}
\]

The crack growth rate is then given by:

\[
\frac{da}{dN} = \frac{\Delta K^4 (1-n)}{8 (1+n) \pi^2 \mu \sigma_y^3 D_c} \tag{1.3.15}
\]

Another dislocation-based model was developed by taking into consideration the instability of dislocation emission from the crack tip (intended as dislocation movement away from the crack tip as responsible for the crack propagation), dislocation group dynamics and macroscopic fracture mechanics concepts. The crack opening displacement \( CTOD \) was calculated as:

\[
CTOD = 2n_d b \tag{1.3.16}
\]
where \( b \) is the Burgers vector and \( n_d \) is the number of dislocations emitted from the crack tip. Assuming that FCP rate is equal to half the \( CTOD \), the FCP rate will be given by:

\[
d\alpha/dN = n_d b \tag{1.3.17}
\]

The number \( n_d \) is calculated according to dislocation group dynamics. For large values, \( n_d \) is given by:

\[
n_d = \gamma(p) \left( \frac{4fb}{v_0} \right)^{-(p+1)/p+2} \left( \frac{\tau_0}{\mu} \right)^{-(p(p+1)/p+2)} \left( \frac{\tau_0}{\mu} \right)^{(p+1)!/p+2} \tag{1.3.18}
\]

where \( f \) is the loading frequency, \( p \) is the exponent on the stress in the equation for dislocation velocity \( \nu \) defined as:

\[
\nu = v_0 \left( \frac{\tau_0}{\tau_0^*} \right)^p \tag{1.3.19}
\]

in which \( v_0 = 1 \text{ cm/s} \). \( \tau_0 \) is the applied shear stress, \( \tau_0^* \) the stress required to give \( \nu = v_0 \), and \( \gamma(p) \) is a dimensionless parameter depending on \( p \).

Assuming that the applied stress near the crack tip is averaged over a distance \( s \) and the applied shear stress is \( \tau_0 = \Delta K/\nu_s \), the FCP rate for a perfect plastic material is given by:

\[
\frac{d\alpha}{dN} = b \gamma(p) \left( \frac{4fb}{v_0} \right)^{-(p+1)/p+2} \left( \frac{\tau_0}{\mu} \right)^{-(p(p+1)/p+2)} \left( \frac{\Delta K}{\mu \sqrt{s}} \right)^{(p+1)!/p+2} \tag{1.3.20}
\]
For a strain-hardening material, the applied stress $\tau_a$ is replaced by:

$$\sigma = f(n')\sigma_{ey}\left(\frac{\Delta K}{\sigma_{ey}\sqrt{S}}\right)^{2n'/n'}$$

[1.3.21]

where $\sigma_{ey}$ is the initial cyclic yield stress, $n'$ the cyclic strain-hardening exponent and $f(n')$ a function of $n'$.

### 1.3.4 LCF-FCP models

This variety of models is designed to describe FCP in low-cycle fatigue (LCF), in which plastic deformations become significant and in which fatigue crack growth comprises a major portion of the fatigue lifetime. The LCF-FCP models allow one to derive a functional dependence on stress intensity of a power factor other than two, which is obtained from COD related models. The physical concept behind LCF models is that FCP results from damage accumulation in small elements ahead of the crack tip, which are subjected to reverse yielding. The crack will grow a distance $l$ when sufficient damage has been accumulated. The distance $l$ indicates a 'process zone' in which the micro LCF process is presumed to operate. The distance $l$ has been identified with the unit crack advance in $\Delta N_i$ cycles, where $\Delta N_i$ represents the number of cycles to crack initiation at an average plastic strain range $\Delta \varepsilon_p$. The life of the element is calculated from the Coffin-Manson criteria:

$$\left(\Delta N_i\right)^{\theta} \Delta \varepsilon_p = C_0 \varepsilon_f'$$

[1.3.22]

where $\varepsilon_f'$ represent fatigue ductility and $C_0$ is a constant. The plastic strain is related to the stress intensity parameter and the equation for FCP results:
\[
\frac{da}{dN} = \frac{C}{\left(\sigma_{yc}^e e_f E\right)^{1/\gamma}} \cdot \frac{1}{1^{1/\beta-1}} \cdot \Delta K^{2/\beta} \tag{1.3.23}
\]

\(E\) is the Young's modulus, \(C\) a constant and \(\sigma_{yc}^e\) is the cyclic yield strength, counterpart of the monotonic yield strength \(\sigma_y\). This distinction is due to the hardening or softening of the material during cycling, which changes the yield strength.

1.4 Microscopic description of FCP

To explain the mechanisms of fatigue crack growth, many models have been proposed from microscopic observations of fracture surfaces and dislocations distribution at the crack tip. These attempts to explain FCP qualitatively do not easily yield quantitative information on the mechanisms involved. Although the primary interest here is FCP, two "static" crack initiation models will be described, since they could also be applied to FCP.

The process of fatigue failure can be briefly described as a crack that nucleates at a surface of a loaded element and gradually grows to a critical length above which the crack is no longer stable and propagates unstably through the material. Fatigue failure is usually divided into three stages\(^2\). At first the crack grows in a shear mode termed as Stage I, strongly influenced by microstructural features and surface roughness. After growing through a few grains, the macroscopic crack plane changes to become normal to the maximum principal stress (Stage II) and the crack operates in the opening mode. The third propagation period, Stage III, arises towards the end of life and involves rapid crack growth.

1.4.1 Crack initiation

A crack can be initiated under the action of cyclic load as a result of plastic deformation. In the Wood's model\(^3\) slip occurs on favourably oriented slip planes during the rising-load part of the cycle (Figure 1.4.1). In the falling-load part, slip takes place in the reverse direction on parallel slip planes. This process gives rise to
intrusions or extrusions, and an intrusion can grow into a crack in the structure during subsequent cycles.

Figure 1.4.1: Wood's model for fatigue crack initiation.

The formation of intrusions and extrusions was considered in the model for crack initiation proposed by Mughrabi (1983). In this model, the preferential sites for nucleation of crack are located in persistent slip bands (PSB) embedded in the crystal matrix. Persistent slip bands may be regarded as a form of matrix damage developed as the result of cycling prior to the appearance of a crack, since they are soft zone which facilitate dislocations motion. During cyclic loading PSB can form on several different slip planes, which in structural materials are often associated with stress raisers such as notches or pores.

Plotting the shear stress $\tau$ as a function of the plastic shear strain amplitude $\gamma_{pl}$ (Figure 1.4.2) it is observed that in the region marked A the cyclic strain hardening occurs more or less homogeneously.
Figure 1.4.2: Shear stress against shear strain showing the behaviour of PSB.

The onset of range B is marked by the first development of PSBs in thin lamellae parallel to the primary slip plane. The PSB then possess a dislocation pattern significantly different from that of the matrix. The deformation becomes inhomogenous and strongly localized in the PSBs. Increasing the shear strain amplitude, the saturation stress remains substantially unaffected until the crystal is completely filled with PSBs (represented by the plateau). Deformation at high $\gamma_{pl}$ in the region C leads once again to a new structure and to an increase of $\tau_s$. $\gamma_{pl,M}$ and $\gamma_{pl,PSB}$ are identified with the plastic shear strain amplitudes that can be accommodated by the matrix and the PSB respectively. During the process of glide and annihilation of edge dislocation, layers of dislocations are deposited at the PSB-matrix interface. Under the action of the applied stress, the interface dislocations tend to glide out of the crystal, leading to the formation of extrusions and intrusions, and the crack initiates at the PSB-matrix interface (Figure 1.4.3).

This model can be extended to polycrystalline materials, taking into account the intergranular or transgranular nature of crack development and it can be applied to explain slow crack propagation. The model assumes two pile-ups of interface dislocations occurring during cyclic deformation, which are connected with stress concentration and under some circumstances lead to cracking in a grain boundary inclined to the dislocation slip plane. The motion of interface dislocations is assumed
not to be influenced by other material defects. To take into account of the interaction between the two pile-ups and of the influence of the elastic anisotropy of neighbouring grains, a model has been proposed\textsuperscript{32}. This model assumes that the shear modulus depends on the grain orientation. Adjacent grains are not perfectly bonded in every case and this may give rise to a tangential slip that modifies the crack geometry and the critical stress level necessary to start the dislocations motion.

Figure 1.4.3: Formation of extrusions in the PSB model.
1.4.2 Crack propagation

A generally accepted characteristic of the fatigue crack growth process, except at very low growth rates in the near-threshold region, is that a fatigue crack advances an increment $\Delta a$ per cycle.$^{33}$

In the plastic blunting model$^{34}$, plastic deformation is limited to a small region near the tip of the crack and the remainder of the structure is elastic. A sharp crack in a specimen, subjected to a tensile load, causes a large stress concentration at its tip where slip can occur easily. The material above the crack may slip along favourably-oriented planes in direction of maximum shear stress and the crack opens extending in length by an amount $\Delta a$ (Figure 1.4.4). This extension is limited by work-hardening of the active slip planes.

![Figure 1.4.4: Plastic blunting model.](image)

Slip can now occur in another plane. As the load is reduced, the crack tip closes down, resharpenes, rewelds and may even fold on itself, causing the presence of striations. To explain this latter phenomenon a shear slip process model$^{35}$ was introduced, as schematised in Figure 1.4.5.
The crack tip is initially characterised by a double notch (Figure 1.4.5a and b). A notch is formed by the forward shear slip in one shear band during the loading half of a fatigue cycle and by the reverse shear slip in a neighbouring shear band during the unloading half of the cycle. This process also forms fatigue striations. The crack tip is blunted by the plastic deformation within the entire crack tip plastic zone and the blunting contributes to crack growth. The double notch configuration at the crack tip can also be explained by the double-shear band model proposed by Tomkins (1968) and illustrated in Figure 1.4.6.
Figure 1.4.6: Tomkins model of double-shear band.

The plastic deformation is still identified with shear slip, and the entire crack tip region contributes to crack growth. However, in this model the deformation takes place only along two shear bands.

The models described overestimate the fatigue crack growth, since in reality only a small part of the crack tip plastic deformation contributes to the crack growth. Neumann\textsuperscript{37} observed experimentally in a single crystal, in fact, that shear slip is concentrated in parallel slip bands separated by elastic slabs with little or no plastic deformation. Neumann thus proposed the following model: the slip near the crack tip is almost perfectly confined to two slip systems that emanate from the sides of the crack (Figure 1.4.7). On each side of the crack there is mainly one active slip, emanating from the tip. In front of the crack an elastic, almost slip-free triangular area is formed by the two slips. During a fatigue cycle, a number of shear decohesions take place alternately on two sets of intersecting conjugate slip planes and the crack tip moves along each active slip band alternately. As a result, the open crack tip is V shaped with a constant angle at the tip. During each increment of shear decohesion, the crack tip moves ahead by an increment $\Delta a$. The process of blunting by plastic deformation consists in a widening of the V. This shape will develop when the dislocations emerging from the tip can move freely away from the tip. This process is facilitated when the slip lines can spread to the back surface of a tension specimen. If the dislocations are relatively closer to the tip, blunting will be favoured over V-tip formation.
The Neumann model describes the mechanism of crack propagation in a single crystal, which has anisotropic properties. For a polycrystal, the stress-strain relationships averaged over a large number of randomly orientated grains can be considered isotropic. The shear deformation in polycrystalline materials is seen to take place with a decohesion mechanism along the planes of maximum shear stresses, and is localised to narrow bands along these slip planes, defined here as shear bands in contrast to slip bands in single crystals. A shear band is the result of extensive shear slip in many slip bands in the numerous grains present along the shear bands. The material between the neighbouring shear bands is deformed plastically but at much lower level. The plastic deformation in small scale yielding is therefore constrained and the crack propagation results in an 'unzipping' mechanism as illustrated in Figure 1.4.8.
Figure 1.4.8: Unzipping model.

Figure 1.4.8a shows the intersecting shear bands ahead of the crack tip before decohesion take place. As the applied load increases, the shear bands are activated, initiating the shear bands decohesion along $\alpha$ and the region above $\alpha$ moves upward. During a second increment, schematised in Figure 1.4.8b, the decohesion process is shifted along $\beta$ and so on (Figure 1.4.8c and Figure 1.4.8d). After the crack tip moves ahead of $\alpha$, the shear decohesion on $\alpha$ will blunt the crack tip (Figure 1.4.8e), but it will not contribute to crack growth, which will be caused by the process at shear bands at the very tip.

The FCP models just described are suitable for a Stage II crack. For Stage I cracks microstructural effects such as the individual grain orientation is believed to become important, since during the growth of small cracks multiple slips system can be active and the crack front adopts different orientations in adjacent grains. The overall crack path reflects the relative strength of the contributions to mixed mode loading from grain to grain and the constraints that arise from maintaining a continuous crack front between different grains.

Recently, a model for FCP unifying fracture event for small and long cracks has been described$^{39}$, also based on a slip band decohesion mechanism. Observation of local direction of crack growth in a fatigue cracked aluminium alloy shows that concentrated slip on multiple slip planes was involved in the fracture process.
Furthermore, the crack front was not contained at the intersection of an active slip plane and its conjugate. Consequently, striations were not formed and an alternating sliding process did not appear to describe the process over the entire crack path. It was then postulated that intense, highly localised plastic deformation occurring at the crack tip during cycling load is concentrated in intersecting slip bands, similar to the persistent slip bands previously described (Figure 1.4.9). These bands are visualised as forming a three-dimensional network immediately ahead of the crack tip. The result is a segmentation of the material into relatively undeformed cells, separated by planar interfaces that have undergone intense shear.

![Figure 1.4.9: Slip bands decohesion model. The darker lines represent intense localised slip bands.](image)

Crack growth occurs by separation of these slip band interfaces. The unit of incremental crack growth is given by the unit size of the slip band network. This model is viewed as being particularly relevant to crack growth when plastic deformation is sufficiently limited to be influenced by local grain orientation and by change of the matrix orientation at individual grain boundaries. Various crack paths can be accommodated within the model, through variation in the number of the active slip planes. The constraint due to adjacent grains and crack closure can be also accommodated. This model is consistent with the proposition that crack growth occurs by damage accumulation in a zone ahead of the crack tip and that the crack growth can be intermittent. This model was based on experimental observation on several
aluminium and titanium alloys. This has suggested that the model is related to the slip mechanism rather than to the material, it is thus generally applicable.

1.4.3 Submicroscopic cleavage and reverse shear model

The need to consider the intrinsic events ahead of the crack tip has been underlined also by Lal (1994), who proposed a new mechanistic approach to the fatigue crack growth process, supposed to be driven by two processes. When a critically stressed zone ahead of the crack tip \( V_c \), is very small, for example smaller than a grain, the advancing crack front discriminates the structural inhomogeneity in the form of defect free and defective regions in a grain. In this limit, even small defects such as dislocations are excluded and plasticity is extremely restricted. Crack growth will occur by brittle separation of atomic bonds, following a submicroscopic cleavage (SMC) mechanism controlled by the parameter \( K_{\text{max}} \). If the concentration and mobility of dislocation increase in \( V_c \), microplastic deformation will increase, local strength will decrease and fracture will occur more by slip than cleavage. This mechanism is called reversed shear (RS) and the growth controlling parameter is \( \Delta K \). An increase of \( \Delta K \) in Stage I crack growth increases the size of \( V_c \), and introduces more defects in it, leading in a reversed shear mechanism. This process continues in Stage II but, with a further increase of \( \Delta K \), \( V_c \) may include a large number of grains and its extent will approach a size characteristic of static fracture. The control parameter becomes again \( K_{\text{max}} \). Since \( K_{\text{max}} \) approaches closer to fracture toughness \( K_c \), intergranular separation and microvoids may initiate, indicating the entrance in the Stage III of crack growth. A schematic of the SMC model is shown in Figure 1.4.10: the atoms over a critical distance \( \rho_r^* \) ahead of the crack tip get elastically displaced by \( u \) in the direction of tensile loading. For \( u \) greater than a critical displacement \( u^* \), \( da/dN \) is proportional to \( u \) at a given distance from the crack tip. Also, \( u \) is a function of loading (specifically of \( K_{\text{max}} \)) and of material properties. The basic material factor influencing it will be the bonding forces between atoms. The macroscopic effect of these forces is best reflected and measured by the value of the Young's modulus \( E \).
In the RS mechanism model of Figure 1.4.11, the atoms are subjected at a reversed shear displacement $\Delta v$ along the slip bands emanating at $\pm 45^\circ$ from the crack tip. $\Delta v$ depends on the stress intensity range and material properties, and includes plastic displacements.
Similarly to the SMC mechanism, for shear displacement greater than a critical value $\Delta \nu^*$, the fatigue crack growth rate is proportional to $\Delta \nu$, which is in turn proportional to $\Delta K$. The plastic deformation occurring in $V_c$ is due at a local level to the dislocation movements, without involving much interaction within themselves or the material microstructure. The large elastic constraint on a relatively small plastic zone could justify the elastic character of deformation. The atomic displacement $\Delta \nu$, again will involve atomic bonding forces and the plastic deformation may be in proportion to the elastic deformation. In this case the shear modulus $G$ should be proportional to $E$, the latter becoming once more the most influential material property. In this model, then, the LEFM FCP can be represented by a local tip elastic strain and since the plastic zone size depends on the yield strength, this conclusion is supported by the experimentally observed lack of influence of $\sigma_y$ on the FCP. This theoretical model constitutes an interesting view of the fatigue crack propagation, but from the examples previously studied, it emerges that the influence of plasticity cannot be neglected so drastically.

1.5 Analysis of traditional models

The models described here are limited and both macroscopic and microscopic approaches alone cannot explain completely FCP and fracture mechanisms.

The microscopic models, although established that the mechanisms driving FCP happen in a small volume of material ahead of the crack tip and are influenced by microstructural properties such as slip bands, grain boundaries and defects, do not provide a quantitative description of the phenomena.

The phenomenological models are based on purely empirical consideration and cannot be used as prediction tools if the mechanical properties of the materials are changed or if parameter such as temperature, $R$ and frequency vary. The granular, inhomogeneous and anisotropic structure of the material is not considered and although parameters such as $C$ and $m$ in the Paris-Erdogan law (equation [1.3.1]) are experimentally recognised to be dependent on the material, the microstructural parameters involved cannot be identified.
Dislocation-based models are developed on physically sound ground. However, the crack growth mechanisms are highly simplified, since they assume for example a circular plastic zone shape and plane stress conditions, or displacement varying linearly from the distance from the crack tip. The result is an attempt to build a micro-macro model, but with an artificial link between the two scale lengths. The advantage of these models compared with the phenomenological description is that the microstructural parameters controlling $C$ and $m$ are identified to be dependent on material properties such as the dislocation motion.

In the LCF models, the concept of a process zone crucial for FCP is introduced. However, the process zone so defined has no interpretation of its significance and these models represent still a phenomenological rather than mechanistic approach. FCP laws such as in equation [1.3.22] are in fact expressed by means of the Coffin-Manson parameters, which are not related to the material microstructural features.

Finally, a real material is not linear elastic or perfectly plastic, as assumed in the models described above. A better understanding of FCP and fracture processes is then dependent on a complete and accurate description of the deformation fields at the crack tip, over a region covering a few grains of material.

### 1.6 Crack tip elastic-plastic strain field

One of the major breakthroughs in describing the elastic-plastic fields ahead of a crack tip was achieved independently by Hutchinson and by Rice and Rosengren. In their description of the deformation field developed during the tensile loading of a crack, the material is considered to be non-linear, obeying to Ramberg-Osgood relationship:

\[
\frac{\varepsilon}{\varepsilon_{ys}} = \frac{\sigma}{\sigma_{ys}} + a \left( \frac{\sigma}{\sigma_{ys}} \right)^n
\]

[1.6.1]
where $\varepsilon_{ys} = \sigma_{ys}/E$ with $E$ the Young's modulus and $n$ the hardening exponent. If the non-linear strain is small compared to the linear contribution, the simplified expression for equation [1.6.1] is:

$$\frac{\varepsilon}{\varepsilon_{ys}} = a \left( \frac{\sigma}{\sigma_{ys}} \right)^n$$  \[1.6.2\]

Using the $J$-integral definition, since the integral is path-independent, it is possible to take a contour $C$ at a radius $r$ from the crack tip. Equation [1.1.7] can then be written as:

$$J = \int_{-\infty}^{\infty} \left( W \cos \theta - T \frac{\partial u}{\partial x} \right) r d\theta$$  \[1.6.3\]

where $\theta$ is the polar angle, $W$, $T$, and $\partial u/\partial x$ are functions of $r$ and $\theta$. Recalling from equation [1.1.6] that $W$ has dimension $\sigma_y \varepsilon_y$, on dimensional grounds $T \partial u/\partial x$ also has dimension $\sigma_y \varepsilon_y$. The integrand in [1.6.3] cannot depend on $r$ since integral $J$ is path-independent, therefore it must be:

$$\sigma_{ij} \varepsilon_{ij} \propto \frac{1}{r}$$  \[1.6.4\]

By substituting equation [1.6.4] in equation [1.6.2], the stress and strain HRR fields at the crack tip are:

$$\sigma_{ij} = \sigma_{ys} \left( \frac{J}{a \sigma_{ys} \varepsilon_{ys} l_n r} \right)^{\frac{1}{n+1}} \tilde{\sigma}_{ij}(\theta, n)$$  \[1.6.5\]
In, $\sigma_{ij}(\theta, n)$ and $\tilde{\varepsilon}_{ij}(\theta, n)$ are tabulated functions\textsuperscript{44} and the amplitude of the dominant singular term depends upon the plane strain or plane stress conditions. The result of this analysis leads to the prediction of a plane strain crack tip tensile stress singularity that is higher than in plane stress conditions. The HRR therefore models, if only approximately, the deformation at the crack tip of material such as aluminium, in which the plane strain fracture toughness is lower than the plane stress value. It does not model for example mild steel mode of deformation\textsuperscript{42}. The applicability of this model is restricted to monotonically loaded cracks, and it represents accurately the fields solution only in a very small portion of the plastic zone, while a complete description requires the solution over the entire plastic zone and surrounding elastic region\textsuperscript{42}.

Observations based on torsion and indentation tests in which specimen size effects were observed, led to the acknowledgment that strain-gradient effects become significant in plasticity in regions of high strain concentrations\textsuperscript{45}. In the case of the tip of a crack where very high strain concentrations are expected, the strain gradient effects become significant when the size of the fracture process zone at the crack tip is comparable to an intrinsic material length $l_m$, typically of the order of microns\textsuperscript{46,47}. The introduction of a microstructural parameter constituted an important step forward in linking microstructural and macrostructural material behaviour. An analytical form for the elastic-plastic fields at the tip of a crack in plane strain condition was obtained:

\[ \varepsilon_{ij} = \alpha \sigma_{ys} \left( \frac{J}{\alpha \sigma_{ys} \varepsilon_{ys} l_m^r} \right)^{\gamma/n+1} \tilde{\varepsilon}_{ij}(\theta, n) \]
\[
\begin{align*}
\epsilon_{rr} &= -\epsilon_{\theta\theta} \\
\epsilon_{r\theta} &= \frac{n+2}{n+1} \left( \frac{n+1}{n} \right)^n \left( \frac{B_I^2}{\sigma_0^2} + \frac{B_{II}^2}{\sigma_0^2} \right) \times \\
\end{align*}
\]

[1.6.7]

where \(\sigma_0\) is the yield strength. \(B_I\) and \(B_{II}\) are similar to mode I and mode II stress intensity factors in classical elastic fracture mechanics. Equation [1.6.7] is derived in plane strain condition and therefore this model does not describe surface strain fields, which in principle are dealt as in plane stress condition. However, thickness effects influence the surface strain distribution, leading to an intermediate condition between plane stress and plane stress. For this reason and due to the lack of a more appropriate description, this model has been considered here to approximately describe surface strain field.

For quasi-static crack growth, the singular term governing the crack tip deformation has been described by the logarithmic equation:

\[
\epsilon_y = \left( \ln \frac{C_0}{r} \right)^{n-1} \tilde{\epsilon}_y (\theta, n)
\]

[1.6.8]

\(C_0\) is an unknown constant representing a scale length parameter, which should not be a function of \(\theta\).

Improvements in strain fields modelling require accurate experimental data, both to establish the regions of validity of the proposed models and to eventually build-up new mechanistic predictions, which would also identify the role of the microstructural parameters involved in fracture and fatigue mechanisms.
Chapter 2

Literature review II: Experimental techniques

The purpose of this chapter is to describe some of the experimental techniques used for deformation fields analysis at the tips of cracks. In particular, the working principle of the moiré interferometry technique will be discussed and then compared to other methods.

2.1 Introduction

As outlined in the previous sections, improvements in understanding fatigue and fracture mechanics depend on an accurate experimental description of the deformation fields over a small volume of material at the crack tip. Techniques that are traditionally applied for this purpose include photoelasticity and caustic methods. The first however is better for transparent or birefringent materials and the second, although can be applied in reflection, is not a whole field technique since it can only give information along a curve (known as the initial curve) around the crack tip. A wide range of more flexible techniques to measure displacement and strain fields has been developed. For example, applications of neutron scattering and atomic force microscopy can be found. The first, however, requires a neutron source, while in the second case, the problem is the poor mechanical stability of the sample. Another method able to reveal the size and shape of the plastic zone and to evaluate quantitatively the magnitude of plastic deformation inside the plastic zone is the microhardness technique. This method is based on the strain hardening or softening in metals which occurs with cyclic loading of the plastic zone ahead of a fatigue crack. The microhardness, which is a material property depending on the density and arrangement of dislocations, is directly related to the strain history of the specimen. The most popular methods, however, are Scanning Electron Microscopy (SEM) combined with stereo-imaging or fine-line/grid methods, techniques based on speckle phenomena, moiré interferometry and also Raman spectroscopy. These techniques allow deformation measurement for virtually any material, over a region ranging from...
a few hundred microns to a few millimetres ahead of a crack. They are therefore particularly suitable for advances in understanding fatigue and fracture. Details of these methods will be presented in the next sections.

2.2 SEM based techniques

The working principle of the Scanning Electron Microscope is rather complicated and a detailed description of the instrument is not appropriate here. Briefly, in a scanning electron microscope the area or microvolume of material to be examined is irradiated with a finely focussed electron beam. The signals produced correspond to specific emission volumes within the sample and include secondary electrons and backscattered electrons, Auger electrons, characteristic X-rays and photons. These signals can therefore be used to analyse many specimen characteristics, such as composition, surface topography and crystallography. The particles produced during the scattering that are of greatest interest in SEM are the secondary electrons, generated by the excitation of the electrons of the conduction bands, and the backscattered electrons, which are the result of elastic scattering. The signals detected from the emission of these particles vary as a result of differences in surface topography as the impinging electron beam is swept across the specimen. The secondary electron emission is confined to a volume near the beam impact area, allowing images of the surface features to be obtained at relatively high resolution. Due to the large depth of field of the SEM, as well as to the shadow relief effect of the secondary electron contrast, the images obtained have a three-dimensional appearance. On the other hand, the backscattered electrons contain information about the tilt angle and about the nature of the specimen over a range of depth. To form the SEM image, the electron beam is scanned across the specimen surface in a x-y pattern, while a cathode ray tube is scanned in the same x-y pattern. In this way a one to one correspondence between the set of beam locations on the specimen and the points on the cathode ray tube is established. The survey map is obtained by scanning an adequate number of specimen areas and using a combination of low-magnification and high-magnification imaging.

For deformation measurements, the SEM images can be analysed in combination with a stereo-imaging technique, which was specifically developed to
determine displacements at and near fatigue crack tips. SEM images of the crack tip region are obtained at various loads, using a loading stage within the scanning electron microscope. The magnification used is high enough to accurately locate the crack tip position, but low enough to include some of the deformation zone ahead of the crack tip as well as a portion of the crack flank. The stereo-imaging technique determines differences between two photographs by comparing them in a stereo-viewer. This visual system causes displacements in an object, seen with both eyes, perceived as depth. The surface displacements are then interpreted as being the third dimension and can be measured using simple tools such as parallax bars or more complex and accurate device such as machine vision systems. Recent applications of this technique include for example measurements of near-tip strain fields in Al-Fe-X alloys\textsuperscript{59}, and local constraint near fatigue cracks in alloys and particulate composites\textsuperscript{60}.

The SEM technique can be used also in conjunction with a micro-line/grid method, which is itself a method for deformation measurements usually used when larger strain are involved\textsuperscript{61-63}. The micro-lines or micro-grids are printed onto the specimen surface by photolithography techniques\textsuperscript{64}. The surface of the specimen is polished and then coated by a photo-resist, or photochemical reactive resin. The specimen is heated to cure the resist and the surface exposed to the light through a photo-mask, or a glass plate with micro-lines or grids. The exposed part of the resist is removed in the developer and vacuum evaporated metal is deposited onto the surface. The remaining resist is then removed in a solvent. This technique was applied, for example, to micromechanical characterisation of local deformation in laminates over a region of \textasciitilde 250 \textmu m. The spatial accuracy of these techniques is virtually limited only by the SEM accuracy. However, a major drawback is the limitation on the specimen geometry and size imposed by the specimen holder.

2.3 Speckle metrology

A class of greatly employed techniques for surface deformation measurements is based on the speckle phenomenon\textsuperscript{65}, which is due to the interference of scattered laser light illuminating a rough surface with height variations greater than the wavelength $\lambda$ of the illuminating light (Figure 2.3.1). The pattern formed consists of dark and bright spots or speckles, which are randomly distributed in space.
Bright speckle is obtained if the contribution of the scattered light at the point \( P \) (Figure 2.3.1) adds up constructively, otherwise the speckle is black. The effect can be obtained also with white light illumination by artificially creating a random speckle pattern\(^{66,67}\). Speckle photography, speckle interferometry and digital image correlation techniques are all methods that extract information about surface deformation from speckle patterns.

Speckle photography\(^ {68}\) involves recording single or double exposure photographs of the undeformed and deformed specimen using laser illumination. The displacement field is then mapped out by measuring point-by-point the separation of the two speckle patterns on the developed film. This technique has been applied, for example, to measure the crack tip displacement fields under static loading conditions\(^ {69,70}\), the transient displacement field around a crack due to stress wave loading\(^ {71}\) and stress intensity factors\(^ {72}\).

Speckle interferometry relies on the interference between a speckle pattern and reference wave, or between two speckle patterns. The displacement field is deduced from the fringe pattern obtained by measuring directly the phase change at each pixel using phase shifting techniques. Laser speckle interferometry has been recently applied to fatigue life estimation\(^ {73}\) and to study effects of tensile strain on fatigue failure\(^ {74}\). Other applications also included the measurements of crack face displacement\(^ {75}\).
The speckle patterns relative to deformed bodies can be also analysed directly through two-dimensional correlation, via comparison of digital video images of the undeformed and deformed configuration\(^7\). Examples of application of this 'computer vision' technique are the study of stable crack growth\(^7\), characterisation of crack tip deformation fields in alloy at high temperature\(^7\), and around a propagating crack tip\(^9\).

The advantage of using speckle-related techniques is that little or no specimen surface preparation is needed. However, an undeformed reference state is not available. Moreover, a speckle pattern is intrinsically affected by noise, which needs to be removed before extracting displacement data.

### 2.4 Raman spectroscopy

If a specimen is irradiated with intense laser light of frequency \(v_0\) in the UV-visible region\(^8\), the scattered light observed in a direction perpendicular to the incident beam consists of a strong Rayleigh scattering, which has the same frequency as the incident beam (\(v_0\)), (Figure 2.4.1). The other weaker signal is due to the Raman scattering that has frequencies \(v_0 \pm v_m\). \(v_m\) is the molecular vibrational frequency measured as a shift from the incident beam frequency.

![Figure 2.4.1: Schematic representation of Raman scattering.](image)
Raman spectra are normally observed for vibrational and rotational transitions. When this technique is applied to deformation measurement, the frequency shift observed between the undeformed and deformed specimen is proportional to the strain or to the stress. This technique allows in situ observation of microstructural strain variation, but it is not suitable for strain measurement in stainless steels for which the Raman lines are broad, due to the crystallographic orientation of individual grains. This introduces a big uncertainty in the frequency shift determination, and therefore in the stress or strain measurement.

2.5 Moiré interferometry

Moiré interferometry is now a well-established technique that provides whole field measurement of surface deformation. This method, characterised by high sensitivity and high resolution, is now successfully applied to a variety of problems, including crack tip analysis, surface displacement at nonambient temperatures, deformation of electronic packaging, microstructural strain analysis and also to study mechanical properties of human teeth.

2.5.1 Working principle

In this technique a diffraction grating is applied to the specimen and it deforms together with the loaded specimen. The specimen grating is viewed together with a superimposed reference grating, usually a virtual grating. The latter grating is formed by two beams of coherent light, reaching the specimen at angles $+\alpha$ and $-\alpha$ (Figure 2.5.1).
The two intersecting beams form alternating lines of constructive and destructive interference on the specimen surface, acting as a virtual reference grating, with pitch $g$ and frequency $f$ given by:
where \( \lambda \) is the wavelength of the coherent light employed here. The virtual reference grating interacts with deformed specimen grating and forms a moiré pattern, which is recorded by a camera. If the two grids are in phase the light transmission is at a maximum (bright fringe), and if they are out of phase the transmission is a minimum.

The initial frequency of the specimen grating \( F_s \) is half the frequency of the reference grating, i.e. \( f/2 \). When the lines of the two gratings are parallel, the light diffracted in the first order emerges essentially normal to the specimen. The mutual interference produced from these two coherent beams gives a uniform intensity throughout the field (null field), since the angle of intersection of the beams is zero. If the specimen is subjected to forces that stretch it uniformly in the \( x \) direction such that the uniform normal strain \( \varepsilon_{xx} \) is constant, the frequency of the specimen gratings decreases to \( F'_s \).

\[
F'_s = \frac{f / 2}{1 + \varepsilon_{xx}}
\]  \[2.5.2\]

Light from the first order of diffraction of beam \( A \) does not emerge perpendicular to the grating, but at an angle \( \beta \), given by the diffraction equation:

\[
\sin \beta_m = \sin(-\alpha) + m\lambda F'_s
\]  \[2.5.3\]

where \( m = \pm 1 \) are the diffraction orders. When \( \varepsilon_{xx} \) is small:
\[
\sin \beta_i = -\lambda f \frac{E_{xx}}{2} \quad \text{for beam } A \tag{2.5.4}
\]

\[
\sin \beta_{-i} = \lambda f \frac{E_{xx}}{2} \quad \text{for beam } B
\]

These two coherent beams propagate toward the camera with angular separation \(2|\beta_1|\), and the result is an interference pattern with uniformly spaced fringes parallel to the \(y\)-direction. This pattern represents a contour map of in-plane displacement given by:

\[
u_x = N_x f
\tag{2.5.5}
\]

where \(u_x\) is the \(x\)-component of displacement at any point in the field, \(N_x\) is the fringe order at the same point in the fringe pattern. The same procedure is applied when the specimen is stretched in the \(y\)-direction. If the reference grating lines are perpendicular to the \(y\)-axis (Figure 2.5.1), now the \(y\)-component of the displacement, \(u_y\), can be calculated from:

\[
u_y = N_y f
\tag{2.5.6}
\]

where \(N_y\) is the fringe order. For small strain, it is therefore obtained:

\[
\epsilon_{xx} = \frac{\partial u_x}{\partial x} = \frac{1}{f} \frac{\partial N_x}{\partial x}
\]

\[
\epsilon_{yy} = \frac{\partial u_y}{\partial y} = \frac{1}{f} \frac{\partial N_y}{\partial y}
\tag{2.5.7}
\]

\[
\gamma_{xy} = \left( \frac{\partial u_x}{\partial y} + \frac{\partial u_y}{\partial x} \right) = \frac{1}{f} \left( \frac{\partial N_x}{\partial y} + \frac{\partial N_y}{\partial x} \right)
\]
If the specimen grating with lines initially parallel to the x direction is rotated by an angle $\Omega$ about the z-axis, the fringe gradient is given by:

$$\frac{\partial N_z}{\partial y} = f\Omega_z,$$ \hspace{1cm} [2.5.8]

Several interferometer configurations have been developed, the simplest being the two beams set-up, which corresponds essentially to the diagram of Figure 2.5.1. The basic system requires two mutually coherent light beams A and B directed onto the specimen and a camera system, with collecting and imaging lenses. There are several ways to produce the two beams. For example, a mirror perpendicular to the specimen could be used to reflect half of a laser beam onto the specimen, reversing the angle of incidence in the process.

Figure 2.5.3: An arrangement to produce two incident beams onto the specimen.
Alternatively, the two beams can be produced by a beam splitter, or they can be delivered directly by optical fibres, which is the system used in this work. This last method offers the advantages of greater set-up flexibility.

Generally, both the \( u_x \) and \( u_y \) components of displacement are desired. Together they fully define the in-plane deformation of the surface and are sufficient to determine the \( x \), \( y \) and shear strain components (equations [2.5.7]), and the rotation \( \Omega \):

\[
\Omega = \frac{1}{2} \left( \frac{\partial u_x}{\partial y} - \frac{\partial u_y}{\partial x} \right) \tag{2.5.9}
\]

The two beams configuration allows the measurement of both components if the system can be rotated by 90° and a cross grating is applied to the specimen. However, it is possible to design optical systems that enable measurement of both \( u_x \) and \( u_y \) by selectively blocking the beams, or by using a four beam fibre-optic system. A portable moiré interferometer that allows measurement of the in-plane displacement components has also been designed.

2.5.2 Specimen preparation techniques

There are essentially two types of technique currently employed to obtain a specimen grating, (sometimes called "replication processes"): firstly the method employed here, and secondly the production of a grid directly on the specimen surface.

The first of these two methods utilises a mould, which is itself a grating obtained by exposing a photosensitive layer of material on a flat glass substrate to a grid pattern. The photosensitive material could be a high-resolution photographic emulsion or a photoresist (photosensitive polymer). The latter allows one to obtain higher diffraction efficiency, which means lower power lasers and shorter exposure times can be used. A holographic mould is a particular example of a phase grating, obtained by exposing the photosensitive plate to a virtual grating produced by illuminating with a dual beam interference pattern.
Before the replication process, the specimen surface must be ground and polished to obtain a flat surface. To replicate the grating onto the specimen surface, a liquid adhesive is allowed to flow between mould and specimen. When the liquid cures and the mould is removed, the specimen is coated with a plastic layer that has the furrowed surface of a grating. The grating must adhere to the specimen and deform together with it, without delamination or slippage relative to the specimen surface.

There are various thermosetting liquids employed such as silicone rubber, epoxy or acrylates. The drawbacks in using silicone rubber are that this polymer does not bond well to most substrates and it is necessary to pre-treat the specimen surface with a silicone primer. Bonding problems were noticed in this work also when an acrylate polymer (Acryfix) was used to produce gratings onto the polished and etched surface of metal specimens.

The grating surface should be highly reflective, so that the fringe pattern is relatively free of speckle noise due to the scattering of light from the specimen surface. This can be achieved for example by applying a reflective metal film by evaporation onto the grating surface of the mould. In this case, the weakest interface in the layers specimen-metal-mould occurs between the mould and the metal film. The reflective film acts as a parting agent, allowing one to remove the mould. Alternatively, the metal film can be evaporated directly onto the specimen once the grating has been replicated. This is the technique employed in this work, as will be discussed later. Aluminium and gold are the metals usually employed for this purpose, depending on the polymer that constitutes the grating. Aluminium, in fact, quickly tarnishes when evaporated directly onto silicone rubber and sometimes can result in pin-holes in the grating when epoxy is used, which is not desirable especially when a moiré microscope is employed.

Finally, an important requirement for a specimen grating is that it should be sufficiently thin to avoid the shear lag effect. This represents the shear stresses attenuation as they propagate through the grating thickness, which could represent a problem, since although the interest is in the deformation of the specimen surface, with moiré interferometry it is the deformation of the grating surface that is measured. The shear lag attenuates the largest strains near a discontinuity or strain gradient and its effect is significant in a region of about one grating thickness in width around severe strain concentration. Effects of the grating thickness on the measurement of strain
around embedded fibre optics sensors have been discussed and compared to finite element calculations in Reference 101.

The alternative method of forming the grating directly onto the specimen surface represents also a solution to shear lag problems. The "zero-thickness" gratings\textsuperscript{102,103}, for example, are amplitude gratings formed by an array of metallic small dots separated by bars of zero thickness. When the specimen is deformed, they move freely to follow the surface distortions. This method is also more suitable for high temperature applications than plastic gratings, which will degrade. To form "zero-thickness" gratings, a layer of metal (30-50 nm, usually gold) is deposited on the polished surface of the specimen. A layer of photoresist applied onto the metal is then exposed to a virtual grating, producing a furrowed surface. By dry ion etching, the photoresist is eroded uniformly and subsequently the metal film is also eroded in the valleys. The remaining photoresist is chemically stripped and a metal grid of dots remains onto the specimen surface.

In this work, as will be illustrated in details in the experimental set-up section, the method chosen to produce a specimen grating was the replication technique. By using a transparent thermosetting liquid, it is possible to see through the grating and image the specimen microstructure.

2.5.3 Advantages and disadvantages

As emphasised in previous sections, advances in understanding of the mechanisms governing fatigue and fracture depend on accurate deformation evaluation. From the equations [2.5.7] it can be deduced that the sensitivity factor in strain measurement is given by $1/(f\Delta)$, where $\Delta$ is the gauge length in the strain measurement (i.e. the length on the specimen across which the strains are averaged). One of the characteristics that therefore supports the choice of moiré interferometry as an ideal tool for strain analysis is that it is very sensitive and it can measure small displacement. The sensitivity $s$ is conventionally defined as the specimen displacement that produces a fringe phase change of $2\pi$, and is given by half the grating pitch:
The theoretical upper limit is therefore given by the upper limit for the grating frequency \( f_{\text{max}} = 2/\lambda \), which for visible light is about 4000 lines/mm. For a specimen grating of frequency \( F_s = 1200 \) lines/mm, which was used in this work, it is \( s = 0.417 \mu \text{m} \) per fringe order. Moiré interferometry is also characterised by high displacement resolution. In usual engineering applications, where the displacement varies smoothly between neighbouring fringes, visual interpolation to \( 1/5f \) or \( 1/10f \) of a fringe order is reliable. For \( F_s = 1200 \) lines/mm, the displacement resolution is better than \( 0.1 \mu \text{m} \).

Examples of high sensitivity moiré interferometer can be found in References 87, 88, 93, 104. To enhance further the interferometer sensitivity, an immersion moiré microscope has been designed based on the principle that by filling the gap between the interferometer and the specimen grating with an immersion fluid of index of refraction \( n \), the frequency of the virtual reference grating \( f \) becomes:

\[
f = \frac{2n \sin \alpha}{\lambda}
\]

and the sensitivity \( s \) increases by a factor \( n \).

Another method used for sensitivity enhancement is the optical/digital fringe multiplication, which being related to fringe analysis will be presented in the appropriate section (Chapter 3).

A further advantage of moiré interferometry, in particular over speckle interferometry, is the high signal-to-noise ratio. The fringe pattern is characterised by high contrast and visibility, and therefore noise reduction procedures are usually not necessary. This useful feature is due to the specimen grating quality, and this aspect will be discussed in detail in the next section.

A problem which can occur during surface analysis is due to accidental rigid-body rotations, which occur during application of loads about an axis parallel to the grating lines. If this happens, the resulting out-of-plane slope is seen as a foreshortening of the specimen grating, which produces an apparent compressive
strain. For example, a rigid body rotation about the $y$-axis, $\Omega_y$, introduces a displacement $u_x$. The pitch $g_s$ of the specimen grating appears to change by an amount:

$$\Delta g_s = g_s \Omega_y^2 / 2 \quad [2.5.12]$$

If the rotation $\Omega_y$ is not too large, the grating experiences an apparent compressive strain

$$\varepsilon_{xx} = -\frac{\Delta g_s}{g_s} = -\frac{\Omega_y^2}{2} \quad [2.5.13]$$

The result is then a second order effect on the fringe gradient, which can be neglected if this is small compared to the strain induced fringe gradient. Otherwise, correction can be made considering that the angle of accidental body rotation $\Omega_y$ is given by:

$$\Omega_y = \frac{\delta}{FL} \quad [2.5.14]$$

where $FL$ is the focal length of the collimating lens and $\delta$ is the distance that the light reflected in the zeroth order of the specimen grating travels in the focal plane when the specimen rotates by $\Omega_y$.

Although the advantages listed make moiré interferometry an attractive method, there is a series of drawbacks, which do not exactly allow the classification of moiré interferometry as a user-friendly technique. Environmental disturbances, as for any interferometric technique, constitute for example a source of concern. The air currents are disturbing because the index of refraction of air varies with pressure and temperature. These are both present in air currents, which may therefore change the optical path of one of the beams relative to the other, causing the shift of the virtual grating and consequently the shift of regions of constructive and destructive fringe pattern. However, J. McKelvie demonstrated that if the angle of incidence $\alpha$ were $\sim 89^\circ$, the emergent angle would be insensitive to changes of $\alpha$, and therefore to thermal...
effects. He therefore devised an optical arrangement for a moiré interferometer substantially immune from thermal disturbance.

In the set-up used in this work, vibrations of optical elements or in-plane motion of the specimen can cause the fringes to move at the vibration frequency. However, these are common problems when using interferometric techniques and eventually they can be solved. Major drawbacks are related to the constraints on the specimen geometry, due to the complexity of the optics\textsuperscript{105}. This problem is not however present in the interferometer employed here, as will be discussed later. The interferometer alignment is time consuming and a careful specimen preparation is required, which is not the case for techniques such as speckle interferometry.

To conclude, it must be mentioned that when microscopic moiré systems are used, the level of details that the technique may reveal legitimately should be estimated\textsuperscript{106}.

The problem can be visualised\textsuperscript{107} by introducing a modified diffraction equation ([2.5.3]) to describe a grating strained by a strain frequency $F_c$:

$$\sin \beta_{mn} = \sin \alpha + m\lambda F_c + nm\lambda F_c$$ \tag{2.5.15}

where $n=0, \pm 1, \pm 2...$ are the diffraction orders related to the strained grating. Typically, effects due to orders greater than the second are negligible.

The presence of the diffraction orders $n$ during a strain variation create a disturbance in the wavefront collected by the camera and a beam divergence as illustrated in Figure 2.5.4.
The effect is more severe for the second order diffracted beam, which is even more divergent and with greater disturbance. The lens therefore sets the limits on the observations that can be made. A lens that is bigger than the field of view is desirable, since the acceptance angle (i.e. the range of angles about the optical axis which is collected by the lens) varies across the lens. However, if the lens collects second diffracted order there will be a distortion in the moiré pattern. The maximum resolvable frequency $f_r$ is then given by the classical relation:

$$ f_r = \frac{2\sin(\theta/2)}{\lambda} \quad [2.5.16] $$

where $\theta$ is the acceptance angle. The distance within which there is no information will then be:

$$ \Delta_{\text{max}} = \frac{1}{f_r} \quad [2.5.17] $$
Chapter 3

Literature review III: Fringe analysis

The object of this chapter is to introduce methods of extracting deformation information from the fringe pattern obtained via interferometric measurement. After a brief introduction on more traditional procedures based on intensity analysis, phase measurement techniques will be discussed. Particular emphasis will be given to temporal phase-stepping methods and to unwrapping algorithms.

3.1 Introduction

When interferometric techniques are used to investigate deformations in a structure, the information is contained in a recorded fringe pattern and needs to be extracted by some fringe analysis method. This difficult task can be performed by using techniques based on intensity analysis. The fringe peak position in the recorded interferogram is detected and digitized, the pattern is subsequently analysed by fringe ordering and numbering.

To detect accurately the fringe peak position, the interferogram requires preprocessing to improve the signal-to-noise ratio and to enhance the fringe contrast. By setting an appropriate threshold, a black and white only binary image can be obtained for which the fringe edge determination is straightforward. An example of this technique was applied to the immersion interferometer mentioned in Chapter 2.

Once the peaks are detected, it is necessary to proceed with fringe numbering and ordering. This is usually performed manually, and several problems often arise at this stage. In the case of a moiré pattern, for example, the zero-order fringe can be assigned arbitrarily and adjacent fringes differ by one or more fringe order. This is not true, however, in regions in which the pattern represents local minima or maxima in the displacement components being measured, where adjacent fringes can have the same fringe order. Furthermore, a sign convention for the fringe gradient is required, the gradient being positive if the fringe order is increasing as the pattern is scanned in the
positive \( x \) direction. The knowledge of the fringe gradient sign is required for the fringe ordering, but cannot be determined unambiguously from the fringe pattern alone. The sign can often be deduced from the experimental conditions, e.g. loading, geometry. It can be also determined experimentally by observing the fringe movement when the specimen is slightly moved.

Intensity based techniques do not usually require special adaptation of the interferometer and fringe numbering is relatively straightforward when the characteristics of the interferogram are highly predictable. However, a limitation for this approach arises from the difficulty in reliable automatic fringe numbering based on intensity data alone in more unpredictable and complicated interferograms. An alternative method is to adopt phase measurement techniques, which measure directly wavefront phase in an interferometer by introducing a phase difference between the test and reference laser beams. The direct measurement of phase information offers many advantages compared to intensity methods, particularly when deformation measurements are involved. It is, in fact, a high precision method (potentially ten to a hundred times higher than the precision obtained by digitising fringe positions); it consists of a relatively simple process; and finally, the wavefront phase is directly proportional to the displacement to be determined:

\[
\begin{align*}
  u_i &= \frac{s}{2\pi} \Delta \varphi_i(x, y) \\
  &\quad i = x, y, z
\end{align*}
\]  [3.1.1]

where \( s \) is the displacement sensitivity (equation [2.5.10] for \( x \) and \( y \) and \( \lambda/2 \) for \( z \)).

Phase measurements can be performed by temporal or spatial methods\(^{109,110}\), which both involve a phase shift between the test and reference beams\(^*\). With the temporal methods, the phase information is contained in a time sequence of phase shifted fringe patterns. The term phase stepping is often used when the phase is held constant for the integration time of the recording camera, and then increased by a known phase increment ready for the next frame. Phase shifting is used when the phase modulator

\(^*\) Of course in moiré interferometry both interfering beams are reflected by the specimen. Nevertheless the analysis of the interference fringes is the same.
continues the linear ramp throughout the exposure time. The main practical consequence is that phase shifting results in lower modulation than phase stepping, because the cosinusoidal fringe signal is convolved with the time window duration, equal to the exposure time of the camera. However, the subsequent processing steps are identical.

An early application of phase shifting technique to moiré interferometry can be found in reference\(^\text{111}\). Phase stepping can be achieved for example by tilting an optical flat placed in one beam\(^\text{112,113}\). Phase modulation can be also obtained using polarised light, by rotating a half-wave plate\(^\text{114}\), or an analyser\(^\text{115}\), or a combination of both\(^\text{116}\). A straightforward phase stepping technique is to insert in one of the beam paths a mirror driven by a piezo-electric transducer (PZT)\(^\text{117}\).

The basic principle of the spatial methods is to record simultaneously all the phase shifts in the spatial position of an interferogram on a single photographic array. This can be achieved adapting the phase stepping technique by introducing the phase shift in parallel channels\(^\text{118,119}\). Alternatively, a spatial-carrier technique can be used, which does not employ phase stepping devices. The phase shift, in fact, is introduced by a small tilt between the reference and test beams. In moiré interferometry one would achieve this by tilting one or both the illuminating beams.

Fourier Transform and Phase-Stepping methods are the most extensively used techniques for phase measurement and they will be described in the following paragraphs.

### 3.2 Fourier Transform Method

This method is included in the general class of spatial-carrier phase measurement techniques. The basic idea is to introduce a carrier frequency \(f_0\) into the fringe pattern\(^\text{120}\). The fringe pattern intensity is given by:

\[
l(x, y) = I_0(x, y) + I_0(x, y) f(x, y) \cos[\varphi(x, y) + 2\pi f_0 x] \tag{3.2.1}
\]
where $f_0$ is the carrier frequency in the $x$ direction, $I_0$ is the background intensity, $\gamma$ is the fringe contrast and $\phi$ is the phase to be measured. Omitting the $x, y$ dependence, and defining $c$ as:

$$c = \frac{1}{2} I_0 \gamma e^{i\phi}$$  \[3.2.2\]

the equation [3.2.1] can be written as:

$$I = I_0 + ce^{2\pi i f_0 x} + ce^{-2\pi i f_0 x}$$  \[3.2.3\]

To obtain the phase $\phi$, equation [3.2.3] is first Fourier transformed with respect to $x$:

$$\tilde{I}(f_x) = A(f_x) + C(f_x - f_0) + C*(f_x + f_0)$$  \[3.2.4\]

where the upper case letters denote the Fourier spectra. Assuming that the amplitude and phase are slowly varying functions of position (the distance $x$ over which the changes are appreciable is so that $1/x << f_0$), the Fourier spectra of equation [3.2.4] are separated by the carrier frequency $f_0$ (Figure 3.2.1).
Using either of the two spectra, for example $C(f_x-f_0)$, and translating it by $f_0$ toward the origin, $C(f_x)$ is obtained. By carrying out an inverse transform of $C(f_x)$, the complex function [3.2.2] is obtained and the phase to be determined is simply:

$$\varphi = \tan^{-1} \frac{\text{Im}[c]}{\text{Re}[c]}$$  \[3.2.5\]

The advantage of using the FT method is that it requires only one image to map the phase. However, this technique is most suitable for regular fringe patterns, with little contrast or brightness variation across the image. It is therefore problematic to use when discontinuities such as cracks are present and when high spatial resolution is needed.

### 3.3 Phase-Stepping Method

To perform phase measurement with the phase-stepping method, the phase is stepped by a known amount between intensity measurements by introducing a phase stepper device in one of the two beams. Each data frame of measured intensity can be written as:

$$I = I_0 [1 + \gamma \cos(\varphi + \alpha)]$$  \[3.3.1\]
where $\alpha$ is the known modulation. In equation [3.3.1] the three unknowns are $I_0$, $\gamma$ and $\varphi$. To determine $\varphi$, therefore, a minimum of three intensity measurements is required whilst varying $\alpha$. This constitutes a disadvantage compared to the Fourier Transform technique. However, phase stepping techniques can measure the phase with a higher precision. On the other hand, the measurement accuracy can be affected by various factors, among which particular attention should be given to mis-calibration of the phase step, non-linearity effects and environmental sources of error. Effects due to vibrations and air turbulence can be minimised by good mechanical insulation, while the choice of the phase stepping algorithm used to solve the equation [3.3.1] can influence the system sensitivity to error related to the phase step. Some of the algorithms commonly used will be described in the following sections.

3.3.1 Three-frame technique

In this method, three fringe patterns are used to extract $\varphi$ from the equation [3.3.1]. If the phase shift between each image is for example $\pi/2$, and therefore $\alpha_1 = \pi/4$, $\alpha_2 = 3\pi/4$, $\alpha_3 = 5\pi/4$, the intensities relative to the three phase-stepped interferograms are:

$$I_1 = I_0 \left[ 1 + \gamma \cos \left( \varphi + \frac{\pi}{4} \right) \right] = I_0 \left[ 1 + \frac{\sqrt{2}}{2} \gamma (\cos \varphi - \sin \varphi) \right]$$

$$I_2 = I_0 \left[ 1 + \gamma \cos \left( \varphi + \frac{3\pi}{4} \right) \right] = I_0 \left[ 1 + \frac{\sqrt{2}}{2} \gamma (-\cos \varphi - \sin \varphi) \right]$$

$$I_3 = I_0 \left[ 1 + \gamma \cos \left( \varphi + \frac{5\pi}{4} \right) \right] = I_0 \left[ 1 + \frac{\sqrt{2}}{2} \gamma (-\cos \varphi + \sin \varphi) \right]$$

[3.3.2]

The phase at each point is then given by:
3.3.2 Carré technique

This technique is independent of the phase shift value²² and therefore is not sensitive to phase shift calibration errors. The method requires the acquisition of four interferograms phase shifted by \( \alpha \). Assuming that the phase shifting device is linear, the set of equations obtained can be written:

\[
I_1 = I_0 \left[ 1 + \gamma \cos \left( \varphi - \frac{3\alpha}{2} \right) \right]
\]

\[
I_2 = I_0 \left[ 1 + \gamma \cos \left( \varphi - \frac{\alpha}{2} \right) \right]
\]

\[
I_3 = I_0 \left[ 1 + \gamma \cos \left( \varphi + \frac{\alpha}{2} \right) \right]
\]

\[
I_4 = I_0 \left[ 1 + \gamma \cos \left( \varphi + \frac{3\alpha}{2} \right) \right]
\]

The phase shift \( \alpha \) and the phase \( \varphi \) are calculated from these equations ([3.3.4]):

\[
\alpha = 2 \tan^{-1} \left[ \frac{3(I_2 - I_3) - (I_1 - I_4)}{\sqrt{(I_2 - I_3) + (I_1 - I_4)}} \right]
\]

\[
\varphi = \tan^{-1} \left[ \frac{\alpha}{2} \left( \frac{(I_1 - I_4) + (I_2 - I_3)}{(I_2 + I_3) - (I_1 + I_4)} \right) \right]
\]
3.3.3 Four-frame technique

This algorithm involves the analysis of four sets of intensity measurements. When the phase difference between them is $\pi/2$ ($\alpha=0, \pi/2, \pi, 3\pi/2$), the intensity for each interferogram is given by:

$$I_1 = I_0 [1 + \gamma \cos \varphi]$$

$$I_2 = I_0 \left[1 + \gamma \cos \left(\varphi + \frac{\pi}{2}\right)\right] = I_0 [1 - \gamma \sin \varphi]$$  \[3.3.6\]

$$I_3 = I_0 [1 + \gamma \cos (\varphi + \pi)] = I_0 [1 - \gamma \cos \varphi]$$

$$I_4 = I_0 \left[1 + \gamma \cos \left(\varphi + \frac{3\pi}{2}\right)\right] = I_0 [1 + \gamma \sin \varphi]$$

By rearranging equations [3.3.6], it is easy to express the phase as:

$$\varphi = \tan^{-1} \left(\frac{I_4 - I_2}{I_1 - I_3}\right)$$  \[3.3.7\]

3.3.4 Five-frame technique

This method involves acquiring five images phase-stepped relative to each other. Choosing $\pi/2$ phase shift ($\alpha=-\pi, -\pi/2, 0, \pi/2, \pi$), the intensity equations have the simple form:
Ideally, the first and the last data frame should be the same, but the presence of phase shifter errors can cause differences between them. The resulting phase can then be written, by analogy with equation [3.3.7] as:

\[
\phi = \tan^{-1}\left(\frac{2(I_4 - I_2)}{-2I_3 + I_2 + I_1}\right)
\]

[3.3.9]

Compared to the four-frame algorithm, this method has significantly more tolerance of mis-calibration errors: phase errors scale now as the square of the mis-calibration factor rather than linear variation given by the four-frame method. The disadvantage is that it requires more computational and operational time. The five-frame technique was the method employed for all the experimental results employed here. Figure 3.3.1 shows a wrapped phase map obtained with this technique corresponding to vertical displacement component \( u_y \) at the tip of a crack in a fatigued stainless steel, subjected to a remote loading of 6 kN. Black and white represent phase values of \(-\pi\) and \(+\pi\), respectively. The phase map mainly contains only one region of bad data, due to grating deterioration in correspondence of a crack (boxed region in Figure 3.3.1).
Figure 3.3.1: Wrapped phase map obtained with five-frame technique. The boxed region indicates the crack.
3.3.5 "2+1" technique

This is a method designed to minimise the errors due to vibration and air turbulence. It requires three data frames:

\[ I_1 = I_0 [1 + \gamma \cos \varphi] \]

\[ I_2 = I_0 \left[1 + \gamma \cos \left(\varphi - \frac{\pi}{2}\right)\right] = I_0 [1 + \gamma \sin \varphi] \quad [3.3.10] \]

\[ I_3 = \frac{1}{2} I_0 [1 + \gamma \cos \varphi] + \frac{1}{2} I_0 [1 + \gamma \cos (\varphi + \pi)] = I_0 \]

The third frame is the average of the first two frames with a \( \pi \) shift between them, i.e., the dc intensity, which can be acquired at any moment. Vibration and air turbulence can then affect only the first two sets of measurement. If these two frames are taken quickly enough, the errors caused by environmental problems can be greatly reduced. This method leads to the following equation for the phase:

\[ \varphi = \tan^{-1} \left( \frac{I_2 - I_3}{I_1 - I_3} \right) \quad [3.3.11] \]

3.3.6 General algorithms

Defining \( M \) as the total number of frames acquired, the general equation for phase calculation can be written as\(^{125}\):

\[ \varphi = \tan^{-1} \left[ \frac{\sum_{i=0}^{M-1} I_i b_i}{\sum_{i=0}^{M-1} I_i a_i} \right] \quad [3.3.12] \]
where \( a_i \) and \( b_i \) are real sampling coefficients and \( I_i \) is the intensity of the \( i^{th} \) frame.

Techniques such as the four-frame can be generalised as \( N_f \) - frame technique. It consists of taking \( N_f \) frames, each one having a relative phase shift of \( \alpha = 2\pi i / N_f \). In this case is \( M = N_f \). The errors due to higher harmonics in the fringe pattern (caused for example by detector non-linearity) will decrease by increasing the number of frames \( N_f \). It was found that \( N_f > j + 1 \) frames are necessary for the elimination of the effects of higher harmonic components up to the \( j^{th} \) order\(^{126}\).

A general algorithm, designed to minimise phase shift calibration errors\(^{19}\), is the \( (N_f + 1) \)-frame technique \((M = N_f + 1)\) in which an extra frame having a \( 2\pi \) shift relative to the first one is acquired. The five-frame technique described above is one example of the \( (N_f + 1) \)-frame technique. With this respect, it is clear that the five-frame technique is less sensitive to phase shift errors compared to the four-frame, since linear phase errors are cancelled to the first order.

However, the \( (N_f + 1) \)-frame technique does not work well when both phase shift mis-calibration and higher harmonics are present simultaneously. The phase calculated through this algorithm is therefore affected by residual errors, due to higher harmonic components generated by constant phase shift error. In this case, it was found that the minimum number of frames required is \( 2N_f - 2^{126,127} \).

Table 3.1 summarises commonly used algorithms\(^{128}\). Algorithms 1 and 2 of the general class of \( N_f \) - frame technique require a minimum number of three and four frames respectively. Algorithm 2 corresponds to the four-frame technique described previously and equation [3.3.12] corresponds to equation [3.3.7]. Algorithms 3 (the five-frame technique described in equation [3.3.9]) and 5 are examples of \( (N_f + 1) \)-frame technique\(^{124,129}\), while algorithm 4 and 6 are examples of methods insensitive to both higher harmonics and linear phase shift error\(^{124,126}\).
Finally, it is worth mentioning that a technique related both to FT and phase-stepping interferometry is to introduce a known phase shift between successive pixels in an interferogram with a carrier frequency. Then, instead of recording at least three phase-shifted images, three or more sequential pixels are taken to calculate the phase using the corresponding phase-stepping equation. It can be shown that the spatial resolution obtained with this method is mathematically equivalent than what can be achieved by FT.

### 3.4 Phase unwrapping

The solution for the phase angle, obtained by phase measurement techniques described in the previous section, is an inverse tangent function of the image intensity values. There are, therefore, intrinsic discontinuities in the calculated phase each time the phase $\varphi$ changes by $2\pi$. In the fringe analysis literature this is known as the 'wrapped phase', since it is wrapped onto the range $(-\pi, \pi)$. The final step in fringe analysis therefore consists of unwrapping the phase along a path through the data. This path can either be along one or more of the spatial axes, or along the time axis. The
basic principle of this so-called integration process is to estimate the phase gradient between two adjacent pixels in an image or between two successive phase values calculated at the same pixel. If the gradient value exceeds a set threshold, then a phase discontinuity is assumed to lie between these two points and the phase jump is corrected by adding or subtracting $2\pi$, according to the sign of the phase gradient. The phase at any point is consequently given by $\phi + 2\pi N_e$, where $N_e$ is a 'fringe order counter'.

A condition necessary for a successful unwrapping process is the accurate identification of the discontinuities. If the experimental data have a low signal to noise ratio, the detection of discontinuities could be difficult. In addition, if the fringe pattern contains defects, it is not possible in general to unwrap the phase across them and maintain the correct phase relation. These defects can be structural, such as cracks and holes, or related to the specific interferometric technique. For example, in moiré interferometry, they can be related to the quality of the specimen grating. Unwrapping errors could also be caused by a high density fringe pattern, since a high number of phase jumps need to be accurately resolved in a small area. An increasing density in the fringe pattern often cannot be avoided when the structure analysed is deformed.

All these conditions determine the choice between the various proposed unwrapping algorithms, since what differentiates them are the strategies employed to recognize discontinuities and error sources. A key factor is the determination of the most appropriate integration path, because the phase difference between two points is independent of the integration path chosen, as long as it does not cross invalid data.

When the signal to noise ratio is high, algorithms such as "sequential linear scanning" are sufficient to obtain a reliable unwrapped phase map. Each row of the data array is unwrapped independently and their relative phase relationship is determined by unwrapping along a central column. Each pixel in the image array can be unwrapped sequentially along the horizontal and vertical directions by reference to the previous pixels. This "multiple scan direction technique" provides an error check, and therefore it helps to eliminate invalid points from the unwrapping process. Alternatively, the phase gradient could be measured with a "spiral scanning method" between a central pixel and 3x3 neighbour pixels already unwrapped. The condition for these methods to be successful is that at least one adjacent pixel is not masked, i.e.
considered to be bad points. If there is a cluster of masked defects, the approach is to unwrap along a path around the masked region.

An improved "low-cost" and noise immune technique is the "nearest neighbour" algorithm. This method is based on the requirement that, given the phase at the pixel \((x_0, y_0)\), the phase at any other point \((x, y)\) in the image should be defined uniquely, independently of the unwrapping path. If the path is closed, the integral of the phase jumps should be zero. A non-zero value indicates an area in which discontinuities occur. The integration path is then said to contain residuals and the sign of the integral is defined as polarity, or charge, of the residual. Since the residuals tend to be in pairs of opposite sign called dipoles, although isolated sources can occur near the boundary, cut lines are placed between dipoles in the phase map. These branch-cuts act as barriers to unwrapping to avoid path dependent results. Each residual is allowed to be at one end of a cut, with the other end attached to a source of opposite sign, or to the boundary of the phase map. The minimisation of the cut-length is the criterion used to decide how to pair the residuals. In this way, several different cut placement routes will in general give the same minimum cut length. The choice of the route, however, will affect only the unwrapped phase in the region containing the corrupted phase information. The path chosen for the calculation is therefore not important. To guarantee the correct cut distribution, this method requires that the separation between dipoles is always larger than the spacing between the sources making up the dipoles. The algorithm could fail in the situation in which the boundary is very close to a dipole.

The "modified nearest neighbour" algorithm is a development of the previous method that provides better results at high source densities. The main feature is that all the sources that might end up as dipole pairs are joined; simple rules are then used to select the cuts to be broken that will leave dipoles. This algorithm, like the previous, is based on a local search method and the result could end with long and physically unacceptable cuts.

With the aim of developing a global, whole field search method, a "minimum cut-length" algorithm has been recently implemented. The criterion for placing the cut length to match the discontinuity sources consists of a global minimisation of the sum of the cut-length over the entire phase map.

Figures 3.4.1 and 3.4.2 show the distribution of the branch cuts placed between discontinuity sources obtained unwrapping the phase map of Figure 3.3.1 with the
modified nearest neighbour and the minimum cut-length methods, respectively. For clarity, since the majority of discontinuities are placed in the region of bad data boxed in Figure 3.3.1, the branch cut map corresponding to that region is shown magnified. Using the modified nearest neighbour algorithm, the cuts placed between dipoles are longer compared to the cuts obtained with the minimum cut-length method. The consequence on the unwrapped phase map can be seen by comparing the unwrapped phase resulting from the application of the modified nearest neighbour method (Figure 3.4.3) and the unwrapped phase obtained employing the minimum cut-length (Figure 3.4.4).
Figure 3.4.1: Branch cut map obtained using the modified nearest neighbour algorithm to unwrap the phase map of Figure 3.3.1. The region of bad data is shown magnified.
Figure 3.4.2: Branch cut map obtained using the minimum cut-length algorithm to unwrap the phase map of Figure 3.3.1. The region of bad data is shown magnified.
Figure 3.4.3: Unwrapped phase map obtained with modified nearest neighbour algorithm.

Figure 3.4.4: Unwrapped phase map obtained with minimum cut-length algorithm.
For the experimental results discussed in this work, the phase is calculated relatively to a reference state. For computational purposes, it is more convenient to unwrap directly the wrapped phase difference introduced by the deformation rather than subtract the individually unwrapped phase maps. This procedure will be described in detail in the experimental set-up section (Chapter 4, Section 4.6). Figure 3.4.5 shows the wrapped phase difference for the same experimental conditions as in Figure 3.3.1, in which the reference was not subtracted. This example of phase map has been chosen because it contains a disturbance of twice the frequency of the reference fringe pattern (Figure 3.4.6).
Figure 3.4.5: Striations in the wrapped phase difference map.

Figure 3.4.6: Reference $y$ fringe pattern.
This effect has been observed previously\textsuperscript{124, 137} and it was found to be related to movement of the phase-stepping device. Although movement of the phase-stepping device is a likely cause for the observed disturbance also in this work, vibration affecting one or more of the five frames acquired cannot be totally discarded. In this work, fortunately, it was observed that this problem affected occasionally only the reference map and this provided an easy solution. The unwrapped reference map, in fact, can be represented approximately with the equation of a plane:

$$\varphi = ax + by + c$$ \hspace{1cm} [3.4.1]

A 'clean' reference map can be obtained by performing a best fit of this plane. The phase difference introduced by the deformation can be obtained by unwrapping the deformed phase map and then subtracting the fit of the unwrapped reference.

Figures 3.4.7 and 3.4.8 show respectively the unwrapped phase difference, where the disturbance is present, and the unwrapped phase map obtained by subtracting the fit of the reference map. These latest maps constitute also an example of the possibility of placing a cut along the crack as barrier to the unwrapping.
Figure 3.4.7: Unwrapped phase difference. The line is the cut placed to stop unwrapping across the crack.

Figure 3.4.8: Result of subtracting the fit of the reference phase map to the unwrapped phase of Figure 3.3.1.
Another approach to the phase unwrapping problem is to formulate it as a least squares problem\textsuperscript{138}, in which a solution surface is 'fitted' to the wrapped phase data. Substantially, this method consists of applying algorithms that minimise the difference between the gradients of the wrapped phase and the solution, where the gradients of the wrapped phase are simply defined as the wrapped phase differences. The difference between least squares methods and 'path-following' methods presented previously is that the former do not deal directly with the residual problem, because they obtain solutions by integrating through the residuals to minimise the gradient differences. The least squares method therefore amounts to solving discretised partial differential equations and it can be either unweighted or weighted.

In the unweighted approach, the sum of the squares of the gradients between the wrapped phase and solution is minimised. The unwrapped phase map can be found, for example, directly by means of the fast Fourier transform algorithm. However, with this method corrupt phase values generally corrupt the unwrapped result on a global scale. The weighted least squares algorithms, on the other hand, use predetermined weights to accommodate the residuals and thus are able to zero-weight the corrupt phase values and produce accurate solutions. They however require quality maps or masks to define the weights.

A different approach has been recently proposed\textsuperscript{139} in which the phase at each pixel is measured as a function of time. Unwrapping is then carried out along the time axis for each pixel, independently from the others and therefore, errors do not propagate out from regions of high noise corrupting the rest of the image. In the case of phase-stepping, this method is intrinsically simple: the phase discontinuity in time,\[
\Delta \varphi (x, y, t) = \varphi (x, y, t) - \varphi (x, y, t-1),
\]
can be calculated directly from the intensity frames at different times\textsuperscript{140}. Using the four-frame technique, the phase gradient can be expressed by:

\[
\Delta \varphi (t) = \tan^{-1} \left[ \frac{\Delta I_{42}(t)\Delta I_{13}(t-1) - \Delta I_{13}(t)\Delta I_{42}(t-1)}{\Delta I_{13}(t)\Delta I_{13}(t-1) - \Delta I_{42}(t)\Delta I_{42}(t-1)} \right]
\]

[3.4.2]
where $I_{ij}(t) = I_i(t) - I_j(t)$, $i,j = 1, \ldots, 4$. Provided the deformation is sufficiently small (i.e. the phase change between times $t-1$ and $t$ lies in the range $(-\pi, \pi)$, the unwrapped phase is given by:

$$\varphi(x, y, s) = \sum_{i=1}^{s} \Delta \varphi(x, y, t)$$

[3.4.3]

The main advantage of the temporal unwrapping is that, being a one-dimensional unwrapping along the time axis, phase errors are constrained within the high-noise regions of the phase map and they are not able to propagate to the low-noise regions. This method, however, requires a significant computational effort and rapid sampling of the phase as function of time. For the purposes of the static experiments carried out in this work, the minimum cut-length algorithm was chosen as the best method for phase unwrapping.
Chapter 4
Experimental set-up

This chapter contains the description of the experimental set-up employed to measure deformation fields at the crack tip of fatigued stainless steels. It includes details on a phase-shifting high-magnification moiré interferometer, specimen preparation, alignment procedure and validation tests.

The interferometer described was initially developed at Cambridge University, Cavendish Laboratory, for microstructural strain analysis in composite materials. A second interferometer has been constructed at Loughborough, with a five-fold increase in pixel count for the image sensor. The technique has also been improved primarily through optimisation of the grating replication process.

4.1 Interferometer set-up

The interferometer consists of a microscope head performing both the function of illumination and observation of the fringe patterns resulting from the superposition of the reference grating and the specimen grating, as described in Chapter II. Figure 4.1.1 is a schematic drawing of the set-up. The light source is a 10 mW He-Ne laser (Melles Griot, \(\lambda=632.8\) nm). Spindler & Hoyer micro-bench components, mounted at the laser output, hold a polarising beam splitter (PBS) that splits the light into two beams. A polariser (P) and a half-wave plate (HWP) control the beam intensity and the split ratio.
Figure 4.1.1: Schematic diagram of the high magnification moiré interferometer.

The system includes also phase-stepping optics. A 45° glass wedge (W1) is placed across the path of one of the beams. W1 is translated by a Physik Instrument P.820.10 piezo-electric translator (PZT), a low voltage device with a specific displacement of 15 μm for 10 V of applied voltage. The PZT is driven by an A/D converter interfaced to a Sun Sparcstation10 via a GPIB communication card. A second identical wedge W2 brings the beam parallel to its original direction. In this way, no beam tilt is produced if the PZT tilts during translation and there is no lateral beam translation, provided that the wedge is angled correctly in the beam (at an angle of 63.7°)\(^\text{137}\).

The beams are delivered to the microscope head by two single-mode optical fibres OF1 and OF2 (Fiber Delivery System). These fibres are polarisation preserving.
with built-in launching lenses, and they are held in four-axis mounts (Figure 4.1.2) placed at the laser output.

Figure 4.1.2: Photograph of the four-axis mounts holding the fibres OF₁ and OF₂.

The entire microscope head is built with Spindler & Hoyer micro-bench components and it allows some degrees of freedom. The output fibre end, also indicated as OF₁ and OF₂ in Figure 4.1.3, can be rotated to orient correctly the light polarisation. The two beams can be also adjusted perpendicularly along x and y (Figure 4.1.3) and therefore, correctly positioned relative to the lenses OL₁ and OL₂. The position of these lenses, in turn, can be varied along z to obtain the optimal expansion of the beams relative to the field of view. The uses of not collimated beams causes slight variation of the illumination direction across the field, whose effects on the strain evaluation will be discussed in section 5.2. The micrometers controlling OL₁ and OL₂ allow the beams' dimensions to be varied from fractions of a millimetre to several millimetres. The mirrors M₁ and M₂ direct the beams onto the specimen (S) and from their interference, a virtual grating of twice the specimen grating frequency (1200 lines/mm) is formed. The mirrors are allowed tilt and rotation movements for alignment purposes. For the same reason, the distance between mirrors and fibers ends can be varied.
The first order beams diffracted from the specimen grating are collected by a microscope objective MO (Ealing Electro Optics) and imaged by a high resolution Kodak Megaplus 1.4 camera. The field of view obtained when a 6x microscope objective is employed is 486x370 \( \mu m^2 \). The extension tube shown in Figure 4.1.3, when removed allows one to use a 3x microscope objective without changes in the interferometer alignment. This solid state camera consists of a camera head and a control unit. The camera head features a Charge-Coupled Device (CCD) array of 1340 vertical shift registers; the first 23 are optically opaque to provide a black video reference. Each register contains 1037 elements, of which the first and the last are also optically opaque. In total, there are 1360(H)\( \times \)1036(V) light sensitive elements, or pixels. The centre-to-centre pixel spacing of 6.8 \( \mu m \) is equal to the pixel dimension, allowing a 100 % fill ratio. The light sensitive area is then maximised and the sampling errors reduced. The camera has a built-in electromechanical shutter that opens exposing the sensor, it then closes while the images are digitised and transferred to host memory using a Direct Memory Access card (EDT Ltd SDV board).

The total size of the microscope is of 320x160x180 mm\(^3\) including camera. This compact design allows the microscope head to be easily mounted on an X-Y-Z translation stage in front of a tensile test machine (Instron 4466), where the specimen can be loaded. The Instron is a desktop twin screw design with a maximum load capacity of 10 kN. The motor mounted optical encoder has an effective crosshead position control resolution of 0.057 \( \mu m \), corresponding to 7% of the specimen grating pitch. In practice, few machine-induced vibration problems were encountered. The loading machine and microscope are mounted on an optical table with air-damped legs to reduce ground-borne vibration.

The X-Y-Z translation stage means that the position of the microscope head relative to the specimen can be accurately controlled, allowing the specimen surface to be mapped without any change in the optics. Both the \( x \) and \( y \) component of the displacement field can be obtained; however, since changing from one to the other involves rotating the microscope mounting plate through 90°, they cannot be measured simultaneously. Figure 4.1.3 is a photograph of the experimental set-up aligned for measurements of \( y \)-component of the deformation fields.
Another major feature of this moiré microscope is that, by switching from laser to white light illumination, it is possible to obtain the displacement field in exact registration with the underlying specimen microstructure. From the analysis of the microstructure contrast described in section 4.6, it emerged that the best illumination for observation of the microstructure is given by a dark field system. With this microscope, a type of illumination close to the dark field can be obtained by an oblique illumination system, consisting of a white light source with two multi-mode optical fibre bundles (Fiberoptic-Heim). The specimen is illuminated from the two sides (Figure 4.1.4), in the direction parallel to the grating lines in order to prevent light diffracted by the specimen grating entering the microscope.
The sensitivity of the technique, defined as the specimen displacement that produces a fringe phase change of $2\pi$, is 0.418 $\mu$m/fringe and is given by half the specimen grating pitch.

Figure 4.1.4: Photograph of the experimental set-up mounted in front of the Instron. The black box is the Kodak camera. Note the diffraction colours from the specimen.

As a concluding remark, the design of this interferometer answers one of the criticisms sometimes made against moiré techniques for strain analysis, i.e. that the complexity of the optical set-up exerts constraints over the specimen geometry.
4.2 Grating replication technique

Some of the techniques employed to apply a grating onto the specimen surface have been reviewed in Chapter II. The method used here consisted of replicating a master grating by moulding it with a thermosetting liquid polymer. The master is a 50×50 mm² holographic phase grating of frequency 1200 lines/mm, from Ealing Electro Optics. Such a grating is expensive and needs to be used more than once. It has been previously observed that silicone rubber is a non-adhesive material. If a silicone sub-master is produced as an intermediate step, it is therefore possible to obtain from the sub-master a number of specimen gratings, without damaging the holographic mould.

4.2.1 Sub-master preparation

The substrate used for the sub-master preparation was a 75×50×6 mm³ glass slide. To allow the bonding between silicon and substrate, the surface was first coated with a thin layer of silicone primer (Dow Corning 92-023). The polymer was applied with a pipette on the clean glass, continuously from one edge of the glass to the other. The slide was kept nearly vertical to obtain a uniform distribution of the liquid. The primer is then allowed to cure for approximately 15 minutes at 50°C.

The silicone rubber (Sylgard 184, Dow Corning) was mixed with its curing agent and an accelerator (QFC 3-6559, Dow Corning) in the ratio of 10% and 5% by mass, respectively. Theoretically, it should be possible to cure the polymer without accelerator, but the polymerization was not always complete. The mixture is first degassed, the bubble free compound is then simply poured onto the substrate and squeezed between glass and holographic master, applying a pressure by mean of dead weight (~2 kg). The silicone is allowed to cure, usually overnight, at a temperature between 50°C and 60°C. Master and glass slide are then separated by inserting a blade between them, so as to introduce a crack. By gently tapping on top of the blade, the crack propagates and the two surfaces separate smoothly.

4.2.2 Specimen grating

The specimen surface first undergoes a hand grinding process and is then finished with diamond polishing. To replicate the sub-master grating onto the specimen
surface, the epoxy liquid plastic PL-3 (Photoelastic Division, Measurements Group Inc.) was employed. This is a transparent material of low modulus of elasticity (0.0014 GPa), with excellent bonding properties on surfaces.

Resin and hardener, in the ratio of 1:1.5 by weight, are first degassed in two separate test tubes and heated to a temperature between 52° C and 57° C. The materials become less viscous, allowing easier mixing of the components. The mixing process should be carefully carried out. A clear, transparent grating is in fact obtained only if the resulting mixture is uniform. A good indication is usually given by observing the mixture colour, since residual unmixed resin has a lighter colour. This process needs to be performed relatively quickly, usually a few minutes, because the polymerisation is an exothermic reaction, activated by the heating produced when the two components mixed together reach the temperature of 60°. The time required for mixing cannot therefore be longer than the time required for the process to be activated. It is important to mix thoroughly but without introducing air bubbles. This can be achieved by mixing the components in a test tube with a spatula whose blade length is equal to the mixture level and avoiding whipping movements.

Both specimen and grating are also first heated to a temperature of 60° C. The liquid polymer is transferred onto the specimen by extracting the spatula from the test tube when the mixture is ready, bringing it into contact with the specimen and letting the mixture flow onto the surface. The sub-master is then placed over it (Figure 4.2.1).
Figure 4.2.1: Preparation of specimen grating: a) the epoxy is transferred onto the specimen surface and b) squeezed between specimen and sub-master. After curing c) the sub-master is removed and d) a gold layer is evaporated onto the specimen grating.

This method of application also minimises the introduction of air bubbles in the liquid and it is possible because of the resin's low viscosity, which allows the use of a small amount of mixture to obtain a thin grating.

The specimen grating should be a thin layer because moiré techniques measure deformation on the grating surface. If the grating is too thick, the measurement will not represent accurately the specimen surface deformation, but it will be influenced by shear lag effects, whereby shear stresses in the grating attenuate as they propagate through the grating thickness. This effect affects the results, particularly near discontinuities such as a crack, as a rule of thumb over a region of width approximately equal to one grating thickness. By controlling the pressure applied during the curing process (1 MPa), it is possible to obtain a repeatable grating of thickness approximately 4μm. Reliable results can hence be expected outside a region of about width approximately 4μm around severe strain concentration areas.
The applied pressure needs to be uniform over the entire specimen surface for two reasons. A non-uniform pressure would cause a non-uniformity of the grating thickness over the specimen surface. The second problem arises when the grating is applied to the stainless steel compact tension specimen, for which the grating is applied over a quite large surface (approximately 50×50 mm²). If the pressure were not uniform, the epoxy curing process would be slower in the regions subjected to higher pressure. In the instruction for mixing type PL-3 liquid plastic, in fact, it is stated that the polymerization time is shortened by higher temperatures and/or by casting thicker sheets. This was also observed during this work and the risk is therefore to find patches of epoxy not completely cured which would result in a grating not completely replicated, sometimes in the less desirable areas, such as at the crack tip. This problem was avoided by applying the pressure by means of dead weights of surface larger than the specimen surface.

When the curing process is completed, specimen and sub-master are separated using the same technique employed to separate master from sub-master.

The grating thickness is measured with a Nikon interferometric microscope. The micrometer eyepiece allows the measurements of the average spacing \( d \) between monochromatic Fizeau fringes observed with an interference filter (\( \lambda = 514 \text{ nm} \)). Removing the filter, two sets of white light fringes can be seen, corresponding to the top and bottom grating surfaces.

The number of fringes contained in the distance \( D \) between the two central dark fringes is:

\[
N = \frac{D}{d}
\]

[4.2.1]

The grating thickness is then given by:

\[
t = \frac{N\lambda}{2n}
\]

[4.2.2]

where \( n = 1.5 \) is the epoxy refractive index, assuming that the phase change of the light at the air-epoxy interface is the same as that at the epoxy-metal interface.
A metallic coating on the specimen grating enhances the intensity of the moiré fringes relative to the speckle pattern due to the specimen surface. This was obtained by evaporating a small amount of gold onto the specimen grating in an Edwards' vacuum chamber. Particular attention to the gold thickness was given when analysing the stainless steel specimens and where the displacement fields were to be correlated to the underlying microstructure. In section 4.6 the steps followed for optimisation of the gold thickness will be described.

4.3 Interferometer alignment

The interferometer design allows many degrees of freedom, as previously described. This characteristic helps to configure the correct set-up and thereby to produce the high quality fringe pattern required for an accurate fringe analysis. On the other hand, too many degrees of freedom make the alignment more difficult to achieve. However, the task becomes less laborious if it is carried out in consecutive steps. Firstly, the fibres launch needs to be aligned, and its efficiency optimised. This is done by regulating the screws in the four-axis mount (Figure 4.1.2). To ensure that the polarisation of the fibres output is the same, a polaroid is placed in front of the laser beam and the input fibre ends rotated until the best light extinction is achieved. The next step is to align the beams so that they travel parallel to the optical axis. This is achieved by altering the adjustment screws of the fibre ends holder in the microscope head (Figure 4.1.3) until each beam centres a blanking disc with a central recess, which can be mounted on each side of the microscope head.

The successive alignment procedure is performed using a silicone sub-master coated with a partially transmitting gold layer and mounted on a rotation stage and aligned perpendicular to the microscope optical axis. This is achieved by using another laser pointed through the open microscope tube (without camera) with the beam entering and exiting the tube on the optic axis. The grating is then aligned so that the beam from this laser is reflected back in itself. The microscope, with a 3x objective (MO) is first focussed on a small cut made on the grating and the cut centred in the field of view. The microscope objective is removed and the mirror position is then regulated so that the incident beams are diffracted from the region of the grating containing the cut into the centre of the camera. If the alignment is correct, a null field pattern should be visualised. The vertical fringe pattern of Figure 4.3.1, corresponding
to the alignment for $x$-component measurement, was obtained by introducing a small mis-alignment in one of the beams. The fringes are of very high contrast and the defect in the field of view is the cut in the grating used for the alignment.

![Fringe pattern obtained from the test grating after the alignment procedure.](image)

The overall interferometer alignment remained stable, even when the interferometer configuration was changed to $y$-component measurements. As mentioned before, an extension aluminium tube (Figure 4.1.3), allows one to change the microscope objective from $3 \times$ to $6 \times$ without need to change the laser beams alignment. However, a source of alignment problem derived from the loosening of the screws in one of the four-axis mounts. The launching and beam efficiency was, therefore, periodically lost.

The fringe pattern obtained with the test grating was used to evaluate the image quality and to identify sources of noise. The optical table minimises ground-borne vibration. The system is sensitive to air turbulence, but its influence on the fringe
pattern was minimised by screening the microscope head and by acquiring quickly the images. The frame rate is given by \(1/(145 \text{ msec} + \text{exposure time})\) and with a typical exposure time of 50 msec it was possible to acquire five frames per second, an acceptable rate for the experiments described here.

Another disturbance to the fringe pattern was caused by the opening and closing of the camera shutter, which originates an acoustic wave propagating through the adjacent optical fibres. Shielding the camera with an aluminium case solved this problem.

### 4.4 PZT calibration

The use of a PZT as phase stepping device requires a calibration procedure to find the relationship between the driving applied voltage and the phase step. The calibration algorithm employed here has been implemented by Dr N. Alcala\textsuperscript{144}. It consists of evaluating the average \(S\) of the squared difference between a reference intensity \(I_0\) and an intensity \(I\), obtained after an unknown phase step, which is produced by the known applied voltage \(V\):

\[
S = \langle (I_0 - I)^2 \rangle \tag{4.4.1}
\]

The quantity \(S\) is a sinusoidal waveform, function of the voltage \(V\). After few cycles, \(S\) can be fitted by the function:

\[
S(V) = A + B \cos(\varphi(V)) \tag{4.4.2}
\]

\(A\) and \(B\) are constant and \(\varphi(V)\) is a slowly varying function of \(V\). The function \(\varphi(V)\) can be approximated by:

\[
\varphi(V) = a + b(V - V_o) + c(V^2 - V_o^2) \tag{4.4.3}
\]
where $V_0$ is the voltage for the first frame. By fitting $S(V)$ over two cycles (Figure 4.4.1), the constants $a = -7.98$ rad, $b = 2.21$ rad/Volts and $c = -2.23 \times 10^2$ rad/Volts$^2$ were obtained. The second order term $c$ is very small, therefore the phase can be assumed to be a linear function of the applied voltage (Figure 4.4.2).

Figure 4.4.1: Second order polynomial fit of $S(V)$. The dots represent the experimental data and the line is the fit.
4.5 The rotation test

With the test grating ready on the rotation stage after the interferometer alignment, it was convenient to perform a first experiment to test automated phase stepping and fringe analysis algorithms. A rotation of the grating will introduce a linear phase change in the pattern of Figure 4.3.1. The phase difference was measured using automated temporal five-frame phase stepping. The five-frame algorithm, as described previously, is less sensitive to PZT calibration errors than the four-frame technique. It also eliminates first order miscalibration errors, sufficient for the purpose of the experiments described here. The phase stepping and image acquisition was performed automatically using software written in Fortran, which was already in use but was modified here to allow the recording of white light images. The software includes a parameter file containing the results of the PZT calibration constants and the camera exposure time, allowing easier access for modifications. Using \( \pi/2 \) phase
shifts between frames, the phase difference is given by the equation \[3.3.9\]. The resulting intensity was stored in arrays proportional to the sine and cosine of the phase value.

The images were analysed again using Fortran software, which allows choosing the phase unwrapping method between the Nearest neighbour, Modified nearest neighbour and Minimum cut-length. To remove the \(2\pi\) discontinuities in the wrapped phase map of Figure 4.5.1, the Minimum cut-length algorithm was employed. The result is shown in Figure 4.5.2. This method was chosen because it offers a global minimisation process, it allows cuts to be placed, stopping the unwrapping process from occurring across defects (such as cracks). Regions of invalid data can also be masked.
Figure 4.5.1: Wrapped phase map obtained after introducing a rotation to the fringe pattern in Figure 4.3.1. The white colour corresponds to \( \pi \) phase values, the black colour to \(-\pi\).

Figure 4.5.2: Result of the unwrapping process applied to the phase map in Figure 4.5.1
The displacement, related to the phase through the equation [3.1.1], can be described with the equation of a plane \( u_x = ax - \Omega y + c \). When pure rotation is involved and considering the variation of the observed displacement along the \( y \) axis, the slope \( \Omega \) of the straight line obtained will correspond to the rotation angle. Figure 4.5.3 shows \( u_x \) for \( x = 180 \mu m \), obtained by rotating the grating of 0.4° (diamonds) and 0.9° (circles) from the original position of Figure 4.3.1.

![Graph showing displacement vs. angle](image)

**Figure 4.5.3:** \( u_x \) component of the displacement in function of \( y \) obtained by rotating the test grating by 0.4° (diamonds) and 0.9° (circles). The continuous lines are fits.

The continuous lines are linear fits performed on the data, with \( \Omega = 0.01580 \pm 1 \times 10^{-5} \) rad for the circles and \( \Omega = 0.00700 \pm 1 \times 10^{-5} \) rad for the diamonds, where the error estimate was obtained from the linear fit. These values are consistent with the approximately known rotation angles.
4.6 The Brazilian test

To check the interferometer ability to detect deformations and therefore to measure strains, a second test was carried out.

The diametral compression test, or Brazilian test, is based on the idea that tensile stresses develop when a circular disc is compressed between two diametrically opposing forces. The maximum tensile stresses (and hence tensile strains) grow perpendicularly to the loading direction and are proportional to the applied load.

![Diagram of the disc in compression test](image)

Figure 4.6.1: The disc in compression test.
The stresses are\(^{145}\):

\[
\sigma_{xx} = \frac{2P}{\pi Dt} - \frac{2P}{\pi t} \left\{ \frac{x^2(R-y)}{[x^2 + (R-y)^2]^2} + \frac{x^2(R+y)}{[x^2 + (R+y)^2]^2} \right\}
\]

\[
\sigma_{yy} = -\frac{2P}{\pi Dt} - \frac{2P}{\pi t} \left\{ \frac{(R-y)^3}{[x^2 + (R-y)^2]^2} + \frac{(R+y)^3}{[x^2 + (R+y)^2]^2} \right\}
\]

\[
\tau_{xy} = \frac{2P}{\pi t} \left\{ \frac{x(R-y)^2}{[x^2 + (R-y)^2]^2} + \frac{x(R+y)^2}{[x^2 + (R+y)^2]^2} \right\}
\]

\[\text{[4.6.1]}\]

\(P\) is the applied load, \(D\) and \(t\) the specimen diameter and thickness respectively, \(\sigma_{xx}\) and \(\sigma_{yy}\) are the normal stresses in the direction perpendicular and parallel to the loaded diameter, and \(\tau_{xy}\) is the shear stress.

Assuming a point load on a thin disc, the normal stress \(\sigma_{xx}\) at the centre of the disc is tensile and constant, while \(\sigma_{yy}\) is compressive:

\[
\sigma_{xx} = \frac{2P}{\pi Dt} \quad \text{[4.6.2]}
\]

\[
\sigma_{yy} = -\frac{6P}{\pi Dt}
\]

In the elastic range, assuming plane stress conditions (ie. \(\sigma_z=0\)) the \(x\) component of the strain is given by:

\[
\varepsilon_{xx} = \frac{1}{E} (\sigma_{xx} - v\sigma_{yy}) \quad \text{[4.6.3]}
\]
where $\nu$ is Poisson’s ratio and $E$ is Young’s modulus. By substituting equations [4.6.2] into equation [4.6.3], the following result is obtained:

$$\varepsilon_{xx} = \frac{1}{E} \sigma_{xx} \left(1 + 3\nu\right)$$  \[4.6.4\]

The Brazilian test was performed on a Perspex disc, of diameter $D = 18.9$ mm and thickness $t_{disc} = 4.8$ mm.

Using the micrometers of the X-Y-Z translational stage, the interferometer was moved to visualise the centre of the disc in the middle of the field of view. Phase-stepped fringe patterns were acquired for each compressive load applied to the specimen that was placed in the Instron 4466, stopping the motion of the crosshead. The phase difference introduced during the disc compression was measured relative to an unloaded state and related to the $u_x$ displacement through the equation:

$$\Delta \varphi_j = \tan^{-1} \left[ \frac{N_2 D_1 - N_1 D_2}{D_1 D_2 + N_1 N_2} \right]$$  \[4.6.5\]

where

$$N_i = 2 \left( I_i^{(4)} - I_i^{(2)} \right)$$

$$D_i = -2 I_i^{(3)} + I_i^{(5)} + I_i^{(1)}$$

with $i=1,2$ corresponding to the reference and loaded state respectively and $I_i^{(1)}, \ldots, I_i^{(5)}$ referring to the light intensity measured at each frame. Subscript $j$ refers to the displacement component ($x$ or $y$) being measured.

For small strains $\varepsilon_{xx}$ and $\varepsilon_{yy}$ are given by:
The experimental strain $\varepsilon_{xx}$ was estimated by calculating the average gradient of $u_x$ in a small region of $50 \times 50 \, \mu m^2$ around the centre of the field of view, to take into account the uncertainty of the position of the centre of the disc.

The Young's modulus was then calculated by performing a linear fit of the data reported in Figure 4.6.2, giving the value $E = 4.43$ GPa, which is higher than the expected $E$ for Perspex (around 3-3.5 GPa, depending on loading rate). The experimental result for $E$, however, has been obtained assuming point contacts. With extended contacts, it has been shown\(^{146}\) that the stress is reduced by a factor:

$$\left(1 - \left(\frac{b}{R}\right)^2\right)$$

[4.6.7]

For a contact length of $b = 2$ mm, the stress is reduced by a factor of approximately 0.8, which would give a value for the Young's modulus of $E = 3.54$ GPa.
4.7 Stainless steel specimen preparation

As described previously, the moiré interferometry technique requires the application of a diffraction grating onto the surface of the specimen. It was also established that, when replication techniques are used, it is convenient to use a silicone sub-master instead of a holographic master. However, precise requirements are imposed on the stainless steel specimen preparation, since a major feature of the interferometer is the imaging of the specimen microstructure in exact registration with the deformation fields.
The geometry chosen for the specimen is a Compact Tension (CT) of 6 mm thickness designed according to British Standards\textsuperscript{147}. Figure 4.7.1 illustrates the specimen and Figure 4.7.2 the pin loading grips, made in medium carbon steel.

Figure 4.7.1: CT specimen geometry.
The grip thread allows the extensions to be fitted to both the Instron 4466 and the Instron 8032, a servo-hydraulic tensile machine more suitable for fatigue tests. Furthermore, it is possible to obtain a better perpendicular alignment of the specimen in front of the microscope head.

To reveal the specimen microstructure, the surface is first hand ground and polished to a 1μm finish. Due to the surface dimension and the nature of stainless steel, this is a long and literally painful process. Two or three days are needed to grind the surface, and the polishing can take four to five hours. The microstructure can then be made visible by chemical etching, using a solution of 17% Glycerol, 50% Hydrochloric acid and 33% Nitric acid. The etching process was also found to be
crucial in maintaining good adhesion between the grating and the substrate, particularly with the high strains encountered at the crack tip.

To enhance the fringe visibility, a gold coating is evaporated onto the specimen grating once it has cured. The gold thickness needs to be controlled to obtain the best compromise between fringe visibility and microstructure images of good contrast. If the gold thickness is too low, the grating reflects only a small fraction of the incident light, and scattering from the underlying microstructure introduces significant speckle noise into the fringe pattern. If on the other hand the gold layer is too thick, the white light will not penetrate sufficiently to produce microstructural images of acceptable contrast.

A systematic study has therefore been carried out to quantify the thickness required. Several layers of gold were evaporated in succession onto the grating. The layer thickness was controlled by a film thickness monitor (Edwards FTM2) consisting of a 6 MHz oscillating crystal placed in the vacuum chamber. As the gold is deposited onto the crystal, the oscillating frequency shift is proportion to the amount of material evaporated. For each gold thickness \( t_g \), both a set of phase-stepped fringe patterns and a microstructure image were recorded.

A good estimate of the fringe noise level is obtained by estimating the noise in a phase map, obtained from the phase-stepped fringe patterns using the procedure described in the next section. Fits were carried out over continuous regions of the wrapped phase map, rather than over the unwrapped phase map, to prevent possible phase unwrapping errors from influencing the results. The noise level will then be given by the standard deviation, \( \sigma_p \), from the best fit of the approximately linear variation of phase values over a cross-section. A large \( \sigma_p \) will propagate affecting the unwrapping process and hence influence strongly the displacement and strain data. A measure of the microstructure contrast is given by the standard deviation of the acquired white light image intensity, \( \sigma_i \), normalized by the averaged light intensity, \( <I> \). The contrast improves with increasing values of \( \sigma_i/<I> \).

The microstructure images were acquired using a conventional optical microscope (Olympus BX60, Universal Plan Semi Apochromat objective), for several types of illumination systems: bright field (BF), dark field (DF) and crossed polariser and analyser (Pol+An). The results are shown in Figure 4.7.3. The dark field illumination gives the best microstructure contrast. Between 2.5 nm and 4 nm, the
standard deviation $\sigma_\phi$ reaches a plateau, while $\sigma_\phi$ does not decrease dramatically with increasing $t_\phi$, due to the surface finish produced by the polishing and etching processes.

It has been observed that scattering from an unpolished surface introduces higher speckle noise into the fringe pattern; a polished, but not etched surface introduces an extraneous fringe pattern due to the grating thickness, that disturbs the moiré pattern.

To determine the most appropriate value for the gold thickness, it must be taken into account that the white light illumination system used in the moiré set up is not a perfect dark field system, therefore the microstructure contrast will be lower than the values estimated using the Olympus microscope. For this reason it has been decided to choose a value for gold thickness $t_\phi$ around 3 nm, which represents a reasonable compromise between maximising the microstructure contrast whilst minimising the phase noise. If the thickness of the gold thickness was increased, the
microstructure contrast observed with the moiré set up was poor. Furthermore, after a few days, furrows formed on the gold surface.
Chapter 5

Results I: Monotonically loaded crack tip fields

In this section, the displacement and strain fields obtained for monotonically loaded cracks in compact tension specimens\(^{142, 148}\) will be presented and discussed. In particular, the experimental strain fields will be compared with analytical models presented in the fatigue and fracture review sections.

5.1 Introduction

The rate of crack propagation during fatigue in a metal is determined by the stress field ahead of the crack tip and by the response of the crystal matrix and other microscopic features to this stress field. This mutual interaction over a region covering a few grains ahead of the crack tip gives rise to the elastic-plastic strain field, the knowledge of which would then play a crucial role in improving the understanding of fatigue mechanisms. Theoretical models describing the strain field at the crack tip are only available for monotonically loaded cracks. They have been described in section 1.6, however their validity is still debated. Improvements of these models are dependent on a detailed experimental description of deformation at the crack tip\(^{29, 149}\).

Various experimental techniques have been employed in the attempt to validate the description of strain fields in terms of the HRR (equation [1.6.6]), and of the logarithmic type of singularity (equation [1.6.8]). The results obtained lead to somewhat discordant conclusions. For example, combined SEM and stereomaging studies\(^{59}\) on Al-Fe-X alloys found a HRR field singularity for stationary cracks and a logarithmic type field for growing cracks. This dependence, however, ceases to exist in the fracture process zone, identified as the region about 5 to 20 \(\mu\)m from the crack tip. Similar results were obtained for 304 stainless steel by a speckle pattern technique over a field of view of approximately 1 mm\(^{77, 150}\), with the vertical component of displacement and strain showing HRR behaviour for stationary cracks in regions where the plastic strains strongly dominate the
elastic strain component (i.e. $\varepsilon_{yy}^{\text{el}}/\varepsilon_{yy}^{\text{pl}} < 0.03-0.05$). The field, however, was not found to describe the horizontal deformation components. For loaded cracks in 316 stainless steel, HRR behaviour was encountered using a Displacement And Strain Image Evaluation System (DAiSIES)\textsuperscript{105}. As a final example, the behaviour of an aluminium single crystal\textsuperscript{98} was found to differ substantially from the polycrystal specimen\textsuperscript{149}; in both cases, the material behaviour was examined by finite element analysis and moiré interferometry. The lack of agreement between these experimental results is a clear indication that more investigation is needed, for both a wider range of materials and experimental conditions. To the best of our knowledge, there has been no attempt to verify equation [1.6.8], which describes near-tip strain fields by taking into account strain gradient effects.

In this section the strain fields at the tip of monotonically loaded cracks will be examined for three types of stainless steels, namely two different grades of austenitic stainless steels, 316 and 316 L, and a 2205 duplex stainless steel (22% Cr, 5% Ni). The 316 L stainless steel was purchased from East Midlands Alloys and Metals, while the 316 and the 2205 were provided by Avesta Sheffield Ltd.

The in-plane displacements at the crack tips have been measured by using the high magnification moiré interferometer described in the experimental set-up section, applied to metals for the first time to the best of our knowledge.

Furthermore, whole field strain maps have been compared for the first time to the recent theoretical model that consider strain gradient effects.

5.1.1 Materials

The major properties of the austenitic and duplex stainless steels studied here are summarised in Table 2.
<table>
<thead>
<tr>
<th>Stainless steel</th>
<th>Typical composition (wt. %)</th>
<th>$\sigma_y$ (MPa)</th>
<th>$E$ (GPa)</th>
<th>$n$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Duplex 2205</td>
<td>0.02 0.17 22</td>
<td>510</td>
<td>200</td>
<td>0.2→0.4$^a$</td>
</tr>
<tr>
<td>Austenitic 316</td>
<td>0.08 0.10 16→18</td>
<td>276</td>
<td>200</td>
<td>5$^b$</td>
</tr>
<tr>
<td>Austenitic 316 L</td>
<td>0.03 0.10 16→18</td>
<td>241</td>
<td>200</td>
<td>-</td>
</tr>
</tbody>
</table>

Table 2: Table of major properties of the stainless steels studied (Avesta Sheffield Ltd., Technical services). The reference values for the hardening exponent $n$ are from Reference 105 ($^a$) and Avesta Sheffield Ltd., Technical services ($^b$). The latter reference value is the conventional hardening exponent, defined as $1/n$.

The two austenitic steels, the 316 and 316 L, have a similar face centered cubic (f.c.c.) crystalline structure (Figure 5.1.1). Their difference is mainly in the carbon content, which is lower in the 316 L grade. The 316 grade offers better corrosion resistance compared to the 316 L grade.

A heat treatment to enlarge the grain size was carried out for the 316 and 316 L specimens studied here, resulting in grain dimension of the order of approximately 25 μm and 40 μm respectively. The duplex steel is not an extensively studied material. The interest in understanding duplex stainless steel behaviour arises because of its properties of high mechanical and fatigue strength, and especially high corrosion resistance. The duplex steel composition consists of a mixture of a softer body centered cubic (b.c.c.) $\alpha$-phase (ferrite) and a harder f.c.c. $\gamma$-phase (austenite). The material was received as rolled plate. The rolling process induces a characteristic band microstructure, consisting of elongated grains (Figure 5.1.2).
Figure 5.1.1: Microstructure of the austenitic 316 stainless steels (field of view 250×150 μm²).

Figure 5.1.2: Duplex microstructure. Note the band structure, in which the ferrite (black grains) and the austenite (white grains) phase can be identified (field of view 250×150 μm²).
This material has anisotropic mechanical properties, depending on the orientation of the rolling direction compared to the load direction. The specimen grain size of the 2205 steel was of the order of 5 \( \mu \text{m} \) and less. A heat treatment for grain size enlargement was not carried out in this case, in order to avoid significant changes in the material microstructure. For this specimen, therefore, white light observation with the moiré microscope was performed with the sole objective of precisely locating the crack tip and repositioning the interferometer in front of the specimen when the latter is under load.

5.1.2 Experimental procedure

Before the surface preparation described in sections 4.2 and 4.7, each specimen was fatigue precracked to 3 mm crack length in the testing machine Instron 8032. For the 2205 steel, the precrack was grown parallel to the rolling direction. The load applied was a sine waveform of mean level 5 kN, amplitude 4 kN and frequency 5 Hz. These values correspond to a stress intensity range \( \Delta K = 47.7 \text{ MPa}\sqrt{\text{m}} \) \((R = K_{\text{min}}/K_{\text{max}} = 0.11, \text{ and where } K_{\text{max}} \text{ and } K_{\text{min}} \text{ are respectively the maximum and minimum stress intensity factors})\). The stress intensity range applied here is a factor three higher than the value usually employed to ensure that plastic deformation ahead of the crack tip is negligible \(^{05, 77}\). The reason for these high values employed were that computer controlled precracking facility was not available, and the specimen was to be constantly monitored with a travelling microscope during fatigue. At low stress amplitude, after a week of monitoring still there was no sign of a crack. The amplitude was therefore raised. The austenitic specimens were then subjected to a stress relieving heat treatment at 600 °C for three hours while the duplex was annealed at 570°C for five hours. Temperature values and annealing times were provided by the Technical Services of Avesta Sheffield.

The procedure fatigue precrack-annealing-grating application was preferred to the grating application-fatigue precrack for mainly two reasons. Firstly, it was found that, due to the high stress intensity values used to fatigue precrack, with the grating application-fatigue precrack procedure the grating quickly deteriorated. Secondly, ahead of a fatigue precracked specimen that has not been annealed, the HRR model and the
model with strain gradient effects cannot be strictly applied. They are, in fact, based on the J-integral concepts, which is not valid for unloaded cracks\(^1\) (Section 1.6).

After the etching and grating application, the specimen was placed in the Instron 4466 testing machine. A set of laser light and white light reference images were acquired. The specimen was thus subjected to Mode I loading of \(K = 35.8\) MPa\(\sqrt{m}\) (maximum load of 6 kN). The load to be applied for the monotonic test was decided after preliminary experiments, designed to determine roughly at which load deformations in the fringe pattern were observed at the crack tip without damage in the grating. Images were also acquired for intermediate loads. In this way, when the maximum load was reached, it was possible to reposition the interferometer so that the field of view would cover the same region of material as in the reference image.

The \(x\) and \(y\) displacement induced by the applied load and relative to the unloaded reference condition are calculated from the phase difference \(\Delta \varphi\) with the procedure described in section 4.6. For each set of five phase-stepped images, a white light microstructure image was also acquired by blocking the laser beams and switching to white light illumination. A problem arising when planning the experiment was to decide whether to perform the interferometric measurement under displacement controlled or load controlled conditions. Load controlled conditions are in general preferable. However, in this case it would have been necessary to set the Instron 4466 to cycle mode. Then, load controlled conditions are achieved by choosing two close values as cycle limit and a suitable crosshead speed. However, if the crosshead moves between the acquisition of one of the five frames acquired during the phase-stepping, significant errors would be introduced in the calculated phase. Furthermore, in this work is the displacement that is measured and therefore displacement controlled conditions are more appropriate.

On the other hand, in displacement controlled conditions the load drops, due to the material relaxation. It was observed, however, that the load starts dropping after about a few minutes after stopping the crosshead. The acquisition time for the five frames was typically of approximately 1 second, for an exposure time of 50 ms. It was thus decided to acquire the images immediately after stopping the crossheads, since the load variation during the acquisition time was not significant.
5.2 316 stainless steel

For displacement measurement in the 316 steel, two specimen were used, one for \( u_x \) and one for \( u_y \) measurement. Figure 5.2.1 shows the wrapped phase map relative to the \( u_x \) displacement component of the specimen loaded to \( K = 35.8 \text{ MPa}\sqrt{m} \).

The phase map relative to \( u_y \) displacement is not shown here since the relative reference map was found to be affected by the 'striations' discussed in the fringe analysis section. Figure 5.2.2 shows instead the relative fringe pattern. The field is asymmetrical but it also contains a horizontal component due to contribution of rotation.

The \( u_x \) and \( u_y \) displacement contour maps, together with the underlying microstructure, are represented in Figure 5.2.3 and Figure 5.2.4 respectively. The contour interval is 1 \( \mu \text{m} \). These plots have been obtained post-processing the data with Matlab.

The regions marked A and B show a very distinctive behaviour: the contour line changes direction at the grain boundary. This observation constitutes further evidence of the importance in considering the discontinuous structure of materials when the deformation at the crack tip is analysed\(^{32}\). In Figure 5.2.4, just below the crack tip, there is a region of noisy data. This is presumably due to a local grating debonding, which was subsequently confirmed by a blue dye penetration technique.
Figure 5.2.1: Wrapped phase corresponding to horizontal displacement component $u_x$ in the 316 steel. Black and white represent phase values of $-\pi$ and $+\pi$, respectively.

Figure 5.2.2: Fringe pattern corresponding to vertical displacement component $u_y$ in the 316 steel.
Figure 5.2.3: $u_\alpha$ contour map for 316 stainless steel loaded at $K=35.8$ MPa$\sqrt{m}$. $u_\alpha$ values are obtained from the wrapped phase map shown in Figure 5.2.1. The contour interval is 1 $\mu$m. Regions marked A and B show a change in direction of the contour lines at the grain boundaries. The circle marks the crack tip.
Figure 5.2.4: u, contour map for 316 stainless steel loaded at $K=35.8$ MPa$\sqrt{m}$. The contour interval is 1 $\mu$m. Regions marked A and B show a change in direction of contour lines at the grain boundaries. The circle marks the crack tip.
It seems appropriate here to determine the level of details that can be legitimately resolved in the displacement maps shown here by estimating the minimum resolvable distance (equation [2.5.17]). The angle $\theta$ in equation [2.5.16] is related to the numerical aperture $NA=0.15$ of the 6× microscope objective in the moiré set-up by the relationship $\sin\theta=NA$. The minimum resolvable distance is then $\Delta_{\text{min}} = \frac{\lambda}{2NA} = 2 \, \mu m$. The variation in the displacements are observed over the grain size of approximately 25 $\mu m$, it is therefore concluded that distortions in the moiré pattern due to diffracted second order do not affect the experimental results shown here.

For small strains, the strain components $\varepsilon_{xx}$ and $\varepsilon_{yy}$ are related to the in-plane displacement components by equations [2.5.7]. The y-component of the strain, $\varepsilon_{yy}$, was therefore obtained from the displacement values by fitting a plane of the form:

$$u_y = Bx + Cy + D \quad [5.2.1]$$

to sub-images of the displacement map. The gradient of the plane along the relevant axis then provides an estimate of the strain component at the central pixel of the sub-image. This procedure provides an average strain value over the sub-image, the dimensions of which define the effective gauge length. In the current example, the 71×71 pixel sub-image size corresponded to a gauge length of approximately 25 $\mu m$, which is of the same order of magnitude as the grain size for this material. In effect the moiré interferometer provided a two-dimensional array of 14×19 independent miniature strain sensors. During the strain calculation, the rigid body rotation contribution to the displacement field contained in the coefficient $D$ is eliminated.

$\varepsilon_{yy}$ is shown in Figure 5.2.5, represented as isostrain contour map, of contour interval 0.01 strain, superimposed on the underlying specimen microstructure. The crack tip is at the origin of the x-y system of co-ordinates (Figure 5.2.6). The strain field has a butterfly shape and is asymmetrical, although the specimen was subjected to remote pure Mode I loading. Deviations from the shape of contour lines expected
for a homogeneous material cannot be attributed with certainty to the granular structure of the material, since the pixel window used here for the strain calculation is comparable to the specimen grain size. An estimate of the error on the strain can be given as follow. If the displacements are independent random variables with standard deviation $\sigma_u$, the standard deviation of the strain values estimated with the above procedure is given by\textsuperscript{128}.

$$\sigma_e = \frac{\sqrt{3}}{2} \frac{\sigma_u}{P^2 L} \quad [5.2.2]$$

where $L$ is the distance between camera pixels, as measured on the specimen surface, and $2P+1$ is the number of side pixels of the sub-image. In the present case is $P=35$ and $L=0.36 \, \mu m$. With a typical $\sigma_u=10 \, nm$ it is obtained $\sigma_e=0.02 \, m\text{strain}$. However, when trying to achieve such high accuracies, systematic errors associated with the optical set-up can become significant. These error sources have been mentioned previously (such as phase-stepping miscalibration, vibration and accidental rigid body rotation).
Figure 5.2.5: $\varepsilon_{yy}$ contour map superimposed on the microstructure for 316 stainless steel. The contour interval is 0.01.

Figure 5.2.6: System of co-ordinates with origin at the crack tip.
Another source of error could be due to the use of non collimated beams (section 4.1). The effect is represented schematically in Figure 5.2.7, in which $\alpha$ is the same as in Figure 2.5.1 and the same $x$, $y$ and $z$ direction are intended.

Figure 5.2.7: Schematic diagram to evaluate the influence of non collimated beams. The dotted lines represent the collimated beams.

If the beam is not collimated, the angle of incidence onto the grating will be given by $\alpha-\delta$, where $\delta$ is the angle subtending the field of view, as indicated in Figure 2.5.7. The effect on the edge of the field of view will then be similar to an out-of-plane rotation. As discussed in section 2.5.3, and considering the same axis, the grating will experience an apparent compressive strain $\varepsilon_{xy}=-\delta/2$. Let $a$ be the width of the
collimated beam, it is $a=480 \mu m \cdot \sin \alpha$. If $d$ of approximately 10 cm is the distance of the objective lens OL from the centre of the field of view, since $d$ is much greater than the field of view, it is approximately $\sin \delta \sim \delta = a/(2d)=18 \times 10^{-4}$. Thus the maximum apparent compressive strain at the edges of the field of view will be of the order of $1 \mu \text{strain}$.

A realistic estimate of the uncertainty on strain is of the order of 0.4 mstrain, obtained by dividing $\sigma_o$ by the gauge length of 25$\mu$m used here.
5.3 316 L stainless steel

Figure 5.3.1 shows the wrapped phase map relative to the $u_y$ displacement component of the specimen loaded to $K = 35.8$ MPa/\( \sqrt{m} \).

Figure 5.3.1: Wrapped phase corresponding to vertical displacement component $u_y$ in the 316L steel. Black and white represent phase values of $-\pi$ and $+\pi$, respectively (field of view 486x370 \( \mu m^2 \)).
In this case the displacement field superimposed on the underlying microstructure is not shown, as it is more interesting to examine the strain fields, in which the rotation contribution is cancelled.

The strain fields at the loaded crack tip is represented, as for the 316 specimen, as an isostrain contour map, of contour interval 0.01 strain, superimposed on the underlying specimen microstructure (Figure 5.3.2). The crack tip is at the origin of the x-y system of co-ordinates (Figure 5.2.6). The distribution is once again characterised by a butterfly shape strain field but it shows a different behaviour from the 316. Two areas of intense localised strain ($\varepsilon_{yy}$ between 4% and 10%) can be seen behind the crack tip, over approximately 50 $\mu$m radius. In the whole field there is the evidence of a butterfly shape field, with two symmetrical fans of approximately 1% strain. A qualitatively similar strain distribution was found at the crack tip in thin sheet of an aluminium alloy$^{152}$. 

125
Figure 5.3.2: $\sigma_y$ contour map superimposed on the microstructure for 316 L stainless steel. The contour interval is 0.01. Regions marked A, B and C indicates where contour lines follow grain boundaries.
The contour line corresponding to 0.01 strain clearly follows the grain boundaries (regions marked A, B and C in Figure 5.3.2). The effect seems genuinely due to the granular structure of the material, since the grain size is larger than the size of the window employed for strain calculation. It is possible that microstructure effects are more pronounced for this steel compared to the 316 because of the larger grain size of the 316 L specimen. The strain was calculated in the small strains approximation (equation [2.5.7], however the strain values obtained here are not small. A more accurate measure of strain would be given by the Lagrangian tensor\textsuperscript{105}:

\[
\eta_{xx} = \frac{\partial u_x}{\partial x} + \frac{1}{2}\left[\left(\frac{\partial u_x}{\partial x}\right)^2 + \left(\frac{\partial u_x}{\partial x}\right)^2\right]
\]

\[
\eta_{yy} = \frac{\partial u_y}{\partial y} + \frac{1}{2}\left[\left(\frac{\partial u_y}{\partial y}\right)^2 + \left(\frac{\partial u_y}{\partial y}\right)^2\right]
\]

\[
\eta_{xy} = \frac{1}{2}\left(\frac{\partial u_x}{\partial y} + \frac{\partial u_y}{\partial x}\right) + \frac{1}{2}\left(\frac{\partial u_x}{\partial x} + \frac{\partial u_y}{\partial y}\right)
\]

[5.3.1]

Unfortunately, the components of the Lagrangian tensor cannot be calculated here, since it is not possible to measure both components of displacement for the same specimen. This would require a second set of fibres and other optical components shown in Figures 4.1.2 and 4.1.3, plus the use of a crossed line grating onto the sample.
5.4 2205 duplex stainless steel

The wrapped phase map for the loaded 2205 steel is shown in Figure 5.4.1.

Figure 5.4.1: Wrapped phase map relative to the 2205 steel loaded at $K=35.8$ MPa m. White colour corresponds to $\pi$ phase value, black to $-\pi$. 
Figure 5.4.2: $\sigma_{yy}$ at the tip of a crack in the 2205 duplex steel, which is characterised by a symmetrical distribution. As for the previous contour maps, the contour interval is 0.01.
Figure 5.4.2 shows the $\varepsilon_{yy}$ field superposed to the underlying microstructure, with the same contour interval used for the austenitic steels. The crack tip is again at the origin of the $x$-$y$ system of co-ordinates. Although the grain size is too small to distinguish the single grains, the band structure parallel to the crack line is still visible. In this case, as for the 316 L, the highest strain concentration is behind the crack tip ($\varepsilon_{yy}$ between 2% and 5%), over 25 $\mu$m radius. In the whole field, as for the 316 L, there is the evidence of a butterfly shape field, with two symmetrical fans of approximately 1% strain. The strain distribution, however, covers a region that is smaller than in the austenitic case. This is expected since the yield strength (510 MPa) for duplex stainless steel is higher than the value for austenitic steels (less than 300 MPa). (These values were provided by Avesta Sheffield Ltd., Technical services).

A common feature of the strain distributions relative to the three steels is the higher strain concentration behind the crack tip. The higher strain concentration should be at the crack tip, where there is a traction free surface. This behaviour could not be justified by the presence of branching behind the crack tip, as proved by the crack line reconstruction in the duplex steel shown in Figure 5.4.3. Grating effects are also discarded, since it is a thin grating. Furthermore, as mentioned earlier, a similar distribution was observed by Dawicke & Sutton (1994)\textsuperscript{152}. In that case there was no grating applied onto the specimen, since the experiments were carried out by digital-image correlation. Notch effects were also rejected because the precrack length is 3 mm, and therefore the crack tip is sufficiently far from the notch. It is important to note that in the Dawicke & Sutton work, the effect was more evident when high stress levels were applied to produce the crack. It is thus possible that this effect is due to the high stress intensity range applied here. The theoretical models describing an asymptotic behaviour of the crack tip fields, such as HRR, are derived for a crack in an infinite plane. Traditional fracture mechanics approaches recognise that the stress and strain fields remote from the crack tip may depend on the specimen geometry. However it is assumed that, under small scale yielding conditions, the near-tip fields have a similar form in all configurations governed by a single parameter (e.g. $K$, $J$ or $CTOD$) that can be used as geometry independent fracture criterion. This is not valid any longer for non-hardening materials under fully plastic conditions. In the present
case, thus, the compact tension geometry of the specimens could be responsible for the maxima of the strain field behind the crack tip, since with this geometry and the high load applied, shear and bending is introduced. The bending could cause a compressive action ahead of the crack tip, which would result in lower strain values compared to those reached behind the crack tip.

Considering only the data ahead of the crack tip, the evaluated strains have been analysed in terms of the theoretical models described in section 1.6.
Figure 5.4.3: Reconstruction of the precrack in the duplex.
5.5 HRR field

The fact that the HRR field is inherently symmetrical about the crack line means that it cannot be expected to provide good agreement with the 316 data. All three data sets are analysed here, however, for the sake of completeness. The $\varepsilon_{yy}$ values are plotted as a function of the distance from the crack tip, $r$, in a bi-logarithmic scale with the origin located at the crack tip (Figure 5.2.6). Figures 5.5.1, 5.5.2 and 5.5.3 are the results for the 316, 316 L and 2205 specimen respectively. If a HRR field exists, a linear relationship would be obtained between $\log \varepsilon_{yy}$ and $\log r$ with a slope $-n/(n+1)$ independent of $\theta$ (equation [1.6.6]). The plots in Figures 5.5.1, 5.5.2 and 5.5.3 show that the experimental values for $\theta=30^\circ$, 45$^\circ$, 60$^\circ$ do not satisfy this linear relationship over all the regions investigated. Some plots also show an upturn in strain at large $r$ values, instead of decreasing monotonically. This is due to the high strain level above and below the crack discussed previously. The fit line has been obtained by using the polynomial fit routine of Matlab and using the first eight data in the case of the 316 steel. The uncertainty on $\ln(\varepsilon_{yy})$ was estimate from the polynomial routine to be $\pm 0.05$. 

133
Figure 5.5.1: Bi-logarithmic plot of $e_{22}$ against the radial distance $r$ from the crack tip in the 316 steel, obtained for $\theta = 30^\circ$ (circles), $45^\circ$ (crosses), $60^\circ$ (diamonds). The dotted and continuous line represents the HRR fields, corresponding respectively to $\theta = 45^\circ$ and $\theta = 60^\circ$.

Figure 5.5.2: Bi-logarithmic plot of $e_{22}$ against the radial distance $r$ from the crack tip in the 316 L steel, obtained for $\theta = 30^\circ$ (diamonds), $45^\circ$ (circles), $60^\circ$ (squares). The continuous and dotted line represents the HRR fields, corresponding respectively to $\theta = 30^\circ$ and $\theta = 60^\circ$. 

134
Figure 5.5.3: Bi-logarithmic plot of $\kappa_{ij}$ against the radial distance $r$ from the crack tip in the 2205 steel, obtained for $\theta = 30^\circ$ (circles), $45^\circ$ (squares), $60^\circ$ (diamonds). The continuous and dotted line represents the HRR fields, corresponding respectively to $\theta = 45^\circ$ and $\theta = 60^\circ$.

A linear least-squares fit was performed for small $r$, in the regions where the plots indicate a linear behaviour. The results obtained for the parameter $n$ are reported in Table 3.
Table 3: Values of \( n \) calculated from fits of the theoretical HRR crack tip model to the measured strain fields for three stainless steel materials along two angular directions. The reference values are from Reference 105 (†) and Avesta Sheffield Ltd., Technical services (‡). The latter reference value is the conventional hardening exponent, defined as \( 1/n \).

<table>
<thead>
<tr>
<th></th>
<th>( \theta=30^\circ )</th>
<th>( \theta=45^\circ )</th>
<th>( \theta=60^\circ )</th>
<th>Reference value</th>
</tr>
</thead>
<tbody>
<tr>
<td>316</td>
<td>0.46</td>
<td>2.36</td>
<td></td>
<td>5†</td>
</tr>
<tr>
<td>316 L</td>
<td>0.04</td>
<td>0.08</td>
<td></td>
<td>-</td>
</tr>
<tr>
<td>2205</td>
<td>0.25</td>
<td>0.25</td>
<td></td>
<td>0.2→0.4‡</td>
</tr>
</tbody>
</table>

The value for the hardening exponent \( n \) is not available for the 316 L specimen. A comparison with the expected value for \( n \) can however be made for the 316 and the 2205 specimens. For all the materials, \( n \) is not independent on \( \theta \). There is no agreement between the \( n \) obtained for the 316 specimen for \( \theta=60^\circ \) and the value \( n=5 \) found in the literature\(^{105}\), while the values \( n<1 \) obtained for the remaining cases are unacceptable. The values for \( n \) vary in fact between 1 and \( \infty \), which corresponds respectively to the limit of elastic and plastic behaviour. These results suggest that the crack tip does not conform to HRR behaviour over the entire region ahead of the crack tips. Deviations of the angular dependence were found also in Reference 8, for example. This behaviour is attributed to three-dimensional effects associated with the stress state of a free surface \([8,9]\). However, this deviation could be also due to the inhomogeneity of the material, which is not taken into account in the HRR models. Furthermore, for the HRR model to be valid, its range of dominance should be large in order to control the fracture process zone, in which microstructural response, large deformation and non proportional loading conditions invalidate the theoretical formulation\(^{153}\). Finally, the non conformity to the HRR model of the experimental
data was also discussed by Hutchinson\textsuperscript{42}, that stated: "plane strain analysis is not applicable to the mild steel model of deformation; but it does model, if only approximately, the deformation at the tip of a crack in a material such as aluminium".

5.6 Logarithmic type of singularity

The experimental data were tested for conformity to the logarithmic model (equation [1.6.8]). For this $\dot{\varepsilon}_{yy}$ is plotted as a function of $r$ at the same fixed angles as before ($\theta=30^\circ$, $45^\circ$, $60^\circ$). Within this representation, equation [1.6.8] is a power law with exponent $n/(n-1)$. The results are presented in a semi-logarithmic plot in Figures 5.6.1, 5.6.2 and 5.6.3, corresponding to the 316, 316 L and 2205 respectively. A non-linear least squares fit was performed for $\theta=45^\circ$ on the data relating to the 2205. As shown in Figure 5.6.3, there is a partial agreement in the central region of data for the 2205 specimen (residual $R=0.6$), but the value for $n$ resulting from the fit for $\theta=45^\circ$ was 0.7, which is unacceptable. The fit was not performed on the data relating to 316 L and 316 since there is no evidence of regions with logarithmic behaviour for this specimen (Figure 5.6.1 and Figure 5.6.2).

It is therefore concluded that a description of the crack tip singularity in terms of a logarithmic law is also not appropriate over the whole field.
Figure 5.6.1: Semi-logarithmic plot of $\varepsilon_{yy}$ against $r$ for $\theta=30^\circ$ (circles), $45^\circ$ (crosses) and $60^\circ$ (diamonds) for 316 steel.

Figure 5.6.2: Semi-logarithmic plot of $\varepsilon_{yy}$ against $r$ for $\theta=30^\circ$ (circles), $45^\circ$ (squares) and $60^\circ$ (diamonds) for 316 L steel.
Figure 5.6.3: Semi-logarithmic plot of $\varepsilon_{yy}$ against $r$ for $\theta=30^\circ$ (circles), $45^\circ$ (squares) and $60^\circ$ (diamonds) for 2205 steel. The continuous line is the theoretical power law for $\theta=45^\circ$. 
5.7 Strain gradient effects

Strain gradient effects can be expected to dominate the strain field behaviour in regions with high strain concentration, such as at the crack tip. They are in general important when the length scale over which plastic deformation occurs is small, typically on the order of microns. In the present case, the material at the crack tip is plastically deformed in a very small volume (Figure 5.2.5, 5.3.2, 5.4.2), and it is therefore reasonable to assume that the effects of strain gradient are large.

For this reason, the strain values obtained have been compared with the analytical strain field (derived for plane strain condition) described by Huang et al. The equation for $\varepsilon_{yy}$ was obtained by transforming the equation [1.6.7] to Cartesian co-ordinates:

$$\varepsilon_{yy} = \left(\frac{\sqrt{3}}{2}\right)^{n+1} \left(\frac{n+1}{n}\right)^{n} \left(\frac{B_{I}^{2}}{\sigma_{0}^{2}} + \frac{B_{II}^{2}}{\sigma_{0}^{2}}\right)^{n+1}.$$

[5.7.1]

In the case of the 316 stainless steel, the superposition of Mode I and Mode II fields has been considered to obtain an asymmetrical field, which is unattainable with pure Mode I loading.

A two variables non-linear least squares fit of the experimental data was attempted, with parameters $B_{I}/\sigma_{0}$, $B_{II}/\sigma_{0}$ and $n$. The fit was done with Matlab and using the built in least squares function.

The initial value for parameter $n$ was once again set to 5. To estimate initial values for $B_{I}/\sigma_{0}$ and $B_{II}/\sigma_{0}$, a linear least squares fit was performed, considering a single amplitude component at a time in the equation [5.7.1].

140
The contour map obtained is shown in Figure 5.7.1, with $n=3$, $B_1/\sigma_0=0.704 \mu m^{1/4}$, $B_2/\sigma_0=0.497 \mu m^{1/4}$ and obtained from the fit. The value obtained for $n$ disagrees with the value found in literature, although it should be remembered that $n$ is generally calculated from stress-strain curves obtained from macroscopic tensile tests and may not necessarily be valid on the scale length considered here. There is still a discrepancy between fitted and measured strain values (Figure 5.2.5) and the shape of the contours does not reproduce the experimental contours. When only the positive strain values at the crack tip were considered, the fit did not improve. Hence, the model for near-tip strain fields in material with strain gradient effects does not describe the field calculated in this experiment. It must be said that there is an ambiguity in the estimation of the crack tip position. The position marked with a circle in Figure 5.2.4 is suggested from the displacement map, and also from the strain map in Figure 5.2.5. This ambiguity arises firstly because the intersection of the crack with the surface may not coincide with the position of the crack line well within the material, and secondly regions of the grating undergoing localized debonding appear dark in the microstructure images and can therefore be mistaken for the crack as discussed previously (Figure 5.2.4). In the present case, this second possibility can be discounted because the grating was subsequently removed and no significant changes to the microstructure appearance were found and the crack tip position observed on the surface coincided with the position given in Figure 5.2.4. However, a least squares fit using also the crack tip coordinates as fit parameters was performed, but the results were unacceptable, since it gave a negative value for $B_1/\sigma_0$.

To reduce the number of parameters involved in the fit procedure, a further test was carried out by fixing $B_1/\sigma_0$ and $B_2/\sigma_0$. A least squares fit of the data was then performed by shifting the $x$ origin with respect to the previous position $x=0$ (and therefore varying the crack tip $x$ position) and using $n$ as parameter. The results are shown in Table 4, where $R$ indicates the residual obtained for each fit. The smaller residual corresponds to the original crack tip position.
Table 4: Values for $n$ and corresponding residual $R$ obtained by fitting the data at different crack tip $x$ origin.

<table>
<thead>
<tr>
<th>Crack tip $x$ origin</th>
<th>$n$</th>
<th>$R$</th>
</tr>
</thead>
<tbody>
<tr>
<td>+70µm</td>
<td>6.5</td>
<td>11.5</td>
</tr>
<tr>
<td>+40µm</td>
<td>6.0</td>
<td>11.5</td>
</tr>
<tr>
<td>+20µm</td>
<td>3.0</td>
<td>7.3</td>
</tr>
<tr>
<td>0</td>
<td>3.0</td>
<td>6.2</td>
</tr>
<tr>
<td>-20µm</td>
<td>3.3</td>
<td>6.4</td>
</tr>
<tr>
<td>-50µm</td>
<td>3.2</td>
<td>6.7</td>
</tr>
</tbody>
</table>

For 2205 steel it seemed appropriate to model the strain field by considering Mode I loading only and using the same procedure as for the 316. In equation [5.7.1] $BIFO=0$ can be set to zero and $n$ and $B_I/\sigma_0$ are the parameters to be determined from the experimental data. An initial estimate for the parameter $B_I/\sigma_0$ was determined by a linear least squares fit to be 0.3. Using this value and an initial exponent $n$ of 3 (within the quoted range), a non-linear least squares fit was performed on the experimental data, which produced values of $n=0.9$ and $B_I/\sigma_0=0.06 \mu m^{1/0.9}$. The resulting strain field is shown in the contour plot of Figure 5.7.2 with the same contour interval of Figure 5.4.2:

Clearly, the material behaviour during this experiment is not well represented by the strain gradient model of equation [5.7.1]. The disagreement is seen more in the shape of the strain field, than in the actual strain values. Furthermore, the value of the work-hardening exponent $n$ results less than one, which is unacceptable.
Figure 5.7.3 shows the result of the fit performed on the data relative to 316 L. In this case, as for the 2205 steel, $B_{ll} = 0$. The fit produced values of $n=0.7$ and $B_1/\sigma_0=0.02 \mu m^{1/1.7}$. Disagreement with the predicted behaviour is found for this material also.
Figure 5.7.1: Fit of $\varepsilon_{yy}$ for the strain-gradient-effect model to the contour of Figure 5.2.5 (316 steel).

Figure 5.7.2: Fit of $\varepsilon_{yy}$ for the strain-gradient-effect model to the contour of Figure 5.4.2 (2205 steel).
Figure 5.7.3: Fit of $\varepsilon_p$ for the strain-gradient-effect model to the contour of Figure 5.3.2 (316 L steel).
5.8 Conclusions

The experiments described in this section show that the set-up constructed in this work, complete with automated phase stepping and fringe analysis, is particularly suitable for quantitative analysis of surface map at the crack tip.

The strain field at the loaded crack tips of three austenitic and duplex stainless steels was measured over a field of view of 370 × 480 µm², using a high magnification moiré microscope. Quantitative values for $\varepsilon_{yy}$ can be therefore provided in a 'process zone' ahead of a monotonically loaded crack tip, in which the mechanisms driving fatigue and fracture are supposed to take place. The analysis over this region constitutes thus a bridge between the microscopic and traditional macroscopic investigation. The experimental data obtained are particularly suitable for the validation of theoretical models and the predictions by three current crack tip models were tested against the results. HRR field, do not describe the strain field at the microscopic level. The analytical model proposed for near-tip field in materials with strain gradient effects is also unable to describe the experimental data. Deviations may be attributable to thickness effects, material inhomogeneities and specimen geometry, which are particularly influent on this scale. In particular, the specimen geometry could be also responsible for the maxima in the strain concentration located behind the crack tip. The logarithmic type of singularity also could not describe the experimental data.

The observed strain distributions are the result of the material response to the applied stress field. This response must be a function of material structure and its crystallographic properties and it is therefore not surprising that the results presented here disagree with models that assume homogeneous and isotropic materials. The influence of the crystallographic structure is observed in austenitic steels and is more evident in the 316L steel, which has larger grains. The granular structure causes deviations in the strain contours and this cannot be taken into account by existing strain fields models.

These results could be used to formulate new deformation equations that take into account slip plane orientations and obstacles to slip, e.g. grain boundaries,
precipitates, dislocation pile-ups, etc. The use of this technique seems thus promising in terms of providing the suitable experimental data for further development of near tip deformation field models and for a better insight into the microstructural mechanisms driving fatigue crack growth.
Chapter 6

Results II: Evolution of strain fields during fatigue

The purpose of this chapter is to illustrate and discuss the evolution of near-tip strain field in austenitic and duplex steels. The results obtained by high magnification moiré interferometry will be supplemented by transmission electron microscopy observation and fracture macrographs.

6.1. Introduction

As already mentioned, the rate of crack propagation in a metal is determined by the stress field ahead of the crack tip and by the response of the crystal matrix and other microscopic features to the applied stress field. These mutual effects give rise to a strain field ahead of the crack tip, part of which will be the consequence of plastic deformation, which in turn involves the movements of dislocations within the crystals ahead of the crack tip. As the incremental crack extension with each stress cycle is controlled by the amount of plastic deformation, it is clear that the extent and the intensity of the strain field ahead of the crack tip will determine the crack growth rate.

Stainless steels behaviour during fatigue has been extensively studied within the traditional approaches described in the review section. However, there are not many experiments that characterise quantitatively the strain field at the crack tip of this material during fatigue over a distance between a few microns and a few millimetres, which is the region in which the mechanisms driving fatigue take place. A few examples can be found in Reference 154 in which the plastic zone over a region of approximately 2 mm was measured ahead of fatigued cracks in steels. The plastic zone was estimated by 'conventional' moiré interferometry (where by conventional here is intended the high sensitivity moiré without microscope head). In Reference 77 the stable crack growth in 304 stainless steel was studied by the computer vision technique.
Since a reference state is available here, the high magnification moiré microscope employed allows one to evaluate quantitatively the elastic-plastic strain ahead of the fatigue crack tips in austenitic and duplex steels.

6.1.1 Experimental procedure

The specimens employed were the same as those described in the previous section. After the measurement performed ahead of monotonically loaded cracks (number of cycles $N=0$), the specimen can then be removed and fatigued in the more suitable Instron 8032 at a mean level of 4 kN, amplitude of 2 kN and frequency 3 Hz ($\Delta K=23.9$ MPa$\sqrt{m}$, $R=0.3$). At various stages of fatigue, the specimen can be placed back in front of the interferometer and realigned in the same position of the reference state by using the crack line position and particular crystallographic features as points of reference. The in-plane surface displacement of unloaded ($K=0$) and loaded ($K=35.8$ MPa$\sqrt{m}$) can then be measured relatively to the undeformed state. In the experiments discussed here, the displacements were therefore measured for $N=2\times10^3$ for the 316 specimen and for $N=0$, $N=2\times10^3$, $N=5\times10^3$, $N=8\times10^3$, and $N=11\times10^3$ for the 2205 specimen.

The strain $\varepsilon_{yy}$ is then calculated from the gradient of the displacement with the same procedure described before.

To supplement the information obtained by moiré microscopy, transmission electron microscopy (TEM) observation of samples taken at the tip of the fatigue crack in the 316 specimen have been performed.

Using TEM technique it is possible to observe the dislocation in a crystal. The dislocations are in fact lattice defects and their presence perturbs the path of the diffracted electron beam relative to its path in a perfect crystal.

A slice of approximately 1 mm thickness was first spark eroded from the surface analysed by moiré technique. Discs of 3 mm diameter were then obtained by spark erosion above, below and ahead of the crack tip. The discs were then polished to 0.17 mm thickness and electro-polished. The samples need to be thin film because the electrons have little penetrating power\textsuperscript{14}. The dislocation distribution was then observed at several location in each TEM sample since TEM images do not allow one
to identify the crystal observed and the dislocation distribution can change dramatically from one crystal to another. What it is observed in reality is a projected length of the dislocation line that lies along a particular plane.

Finally both the austenitic and duplex specimens were broken to perform a visual examination of the fracture surface.

6.2 316 stainless steel

The sequence during fatigue of the elastic-plastic strain field $\varepsilon_{yy}$ at the crack tip of the austenitic 316 stainless steel is illustrated in Figure 6.2.1 shown in a colour-coded map, with strain values increasing from blue colour to red colour. To compare the results, the strain field of Figure 5.2.5 is also shown as a colour map (Figure 6.2.1a). The strain was calculated for number of cycles $N=0$ and applied stress intensity factor $K=35.8$ MPa$\sqrt{m}$ (Figure 6.2.1a), $N=2\times10^3$ for $K=0$ (Figure 6.2.1b), and $N=2\times10^3$ for $K=35.8$ MPa$\sqrt{m}$ (Figure 6.2.1c). Each strain distribution is shown with the colour bar which represents the strain values corresponding to the colour code. At the first load ($N=0$ and $K=35.8$ MPa$\sqrt{m}$), the strain field corresponding to a monotonic load (Figure 6.2.1a) is characterised by a high intensity (approximately 6%) asymmetric distribution at the crack tip, with one larger lobe above the crack tip. Two fans of lower intensity (about 1%) extend in front of the crack tip and over the whole field.
Figure 6.2.1: Elastic-plastic $\varepsilon_{yy}$ strain fields at the crack tip of austenitic 316 stainless steel for a) $N=0$, $K=35.8$ MPa$\sqrt{m}$, b) $N=2 \times 10^3$, $K=0$, c) $N=2 \times 10^3$, $K=35.8$ MPa$\sqrt{m}$. The strain values corresponding to the colour code are indicated in each colour bar.
Figure 6.2.2: 316 microstructure underlying the strain fields of Figure 6.2.1c.
For $N=2 \times 10^3$, the strain (Figure 6.2.1b) appears to be non-uniformly distributed around the crack tip, reaching a peak intensity of about 3%. The specimen is not loaded, therefore this strain distribution can be associated to the accumulation of plasticity during fatigue. The crack has not advanced during fatigue, as shown in Figure 6.2.2, which represents the specimen microstructure for $N=2 \times 10^3$ and $K=35.8$ MPa$\sqrt{m}$. The corresponding strain distribution (Figure 6.2.1c) is characterised by three distinct regions.

Behind the crack tip the blue area of lower strain is maybe associated to the region of grating debonding discussed previously.

The upper region can be identified as a plastic wake, i.e. a region of cyclic plastic zone contained in the monotonic plastic zone\textsuperscript{154}. The larger lobe of Figure 6.2.1a is now completely dominating the field distribution. The two fans of Figure 6.2.1a develop above and below the crack line, becoming of higher intensity (2% to 3%) and occupying a larger region.

After fatiguing to $N=19 \times 10^3$, the crack grew through the whole field of view. Microstructure images were acquired (Figure 6.2.3) using the white light microscope (built-in in the moiré interferometer) and following the crack by moving the translation stage. These images thus represent the specimen microstructure viewed through the moiré grating. The last stage depicts intergranular type of crack growth (see for example the regions marked A and B respectively in Figure 6.2.3). The regions limited by circles C and D indicate the area where TEM specimens were taken.
Figure 6.2.3: Reconstruction of the crack growth. The different stages of fatigue are marked by arrows. The portions of circles indicate the regions C and D where TEM samples were taken. Regions marked A and B show intragranular crack growth.
The dislocations observed in region C are shown in Figure 6.2.4a and 6.2.4b. Two types of dislocation arrangements were observed. Figure 6.2.4a clearly shows an almost square grid of intersecting dislocations, while Figure 6.2.4b shows cell arrays consisting of almost closed loops, sometimes termed 'loop patches' (Liu (1994))\textsuperscript{155}. The arrangement of Figure 6.2.4a is consistent with the unzipping mechanism of crack propagation described in Section 1.4.2 (Figure 1.4.8)\textsuperscript{38,39}. The cracks "unzip" along intersecting alternate slip bands that are formed due to shear localisation forming a network ahead of the crack tip (Figure 1.4.9). The slip bands in this case would be orientated at approximately \( \pm 45^\circ \) relative to the vertical axis of Figure 6.2.4a, parallel to the dark lines which represent the dislocations. A similar network was found also at the tip of a fatigue crack in titanium alloys\textsuperscript{156}. A distribution forming loop patches was found in polycrystalline copper fatigued at low strain amplitude and the arrangement was believed to be responsible for irreversible plastic strain\textsuperscript{155}.

As shown in Figure 6.2.5, the dislocation lines change angle by about \( 45^\circ \) across the grain boundary. This is possibly due to dependence of the shear modulus on the grain orientation. Similar grain boundaries effects have been theoretically suggested by Richter et al\textsuperscript{32} to describe intergranular fatigue crack growth in f.c.c. metals.

On the other hand, the region D below the crack tip is characterised by a generally less dense and disordered dislocation distribution (Figure 6.2.6a). The only ordered area observed is shown in Figure 6.2.6b, in which the dislocations form once again a grid pattern, corresponding to approximately horizontal and vertical slip planes.

Another TEM specimen was taken in front of the crack tip. In this case were found only a few isolated dislocations at various points in the specimen. This confirms also the unzipping model and Neumann model\textsuperscript{37}, which is characterised by an almost slip free region ahead of the crack tip between the slip lines.
Figure 6.2.4: Dislocation distribution in region C. Instrument magnification 20k (field of view 3.5×2.5 μm²). a) grid array, b) loop patches.
Figure 6.2.5: Dislocation lines changing direction at grain boundary. Instrument magnification 20k.
Figure 6.2.6: Dislocation distribution in region D a) disordered distribution, instrument magnification 20k (field of view 3.5x2.5 \(\mu m^2\)), b) only region of order observed in the form of network array. Magnification 33k (field of view 2.1x1.5 \(\mu m^2\)).
The difference between the upper and lower dislocation distribution could be attributed here to a local material hardening. During fatigue, first slip bands in only one side of the crack line are activated. The dislocations arrange themselves in an ordered arrangement, either in loops or grids. When these structures are formed, the material is work-hardened and further deformations are not allowed. Slip bands are thus activated on the other side of the crack line and the dislocations start forming an ordered array. When this region of material also becomes work-hardened, the process continues again on other alternate slip bands.

To complete the characterisation of the material behaviour during this experiment, the specimen was broken and the fracture surface examined is shown in the macrograph of Figure 6.2.7. There is evidence of a slant mode of fracture during the precrack, with a pronounced shear lip on the side of the surface studied by moiré interferometry (indicated by an arrow in Figure 6.2.7). This mode of fracture could explain the asymmetry of the strain field shown in Figure 5.2.5. The high stress intensity range applied to precrack the specimen are in fact expected to favour slant mode of fracture. When lower stress range are applied, Mode I crack growth is favoured and the crack front is flat as shown in the region marked F in Figure 6.2.7 (corresponding to fatigue between \( N=2\times10^3 \) and \( N=19\times10^3 \) cycles). Visual examination of this region shows a rough surface, indicating a cleavage type of crack growth.
Figure 6.2.7: 316 fracture surface macrograph. The precrack region is marked P and fatigue region is marked F, the two regions are separated by the darker line. The shear lip is indicated by the letter S and is on the side of the specimen where displacement was measured. The arrow indicates the direction of crack growth.
6.3 316 L stainless steel

Crack branching during the first $N=2 \times 10^3$ fatigue cycle was encountered in this specimen, as shown in the microstructure image of Figure 6.3.1, obtained by using the Olympus microscope and the Megaplus camera.

In this case it is more interesting to show the fringe pattern relative to the fatigued specimen loaded at 6 kN. This pattern (Figure 6.3.2) shows 'sectors' around the upper crack tip. In each sector the fringes are almost parallel, very similarly to the results obtained ahead of a crack in a single crystal of aluminium. This behaviour was attributed to the orientation of preferred slip planes at the crack tip. In the present case, unfortunately, the presence of branches in the crack and the fact that the specimen is a polycrystal do not allow similar conclusions. However, these sectors were not observed at the crack tip of the 316 steel, which has similar crystalline structure to the 316 L. It could be possible that the sectors in this example are rather due to the interactions between the tips of the two branches.
Figure 6.3.1: Microstructure of 316 L showing branches B1 and B2 in the crack. (Field of view: 100×76 μm²)

Figure 6.3.2: Fringe pattern showing sectors of nearly parallel fringes (Field of view: 486×370 μm²).
6.4 2205 stainless steel

The behaviour of this material during fatigue has not been extensively studied. An example can be found in Reference\textsuperscript{157}, in which the duplex steel is analysed macroscopically. In Reference\textsuperscript{158} the microscopic approach was employed to study the material response to a load applied in the direction parallel to the rolling orientation.

Figure 6.4.1: Elastic-plastic $\varepsilon_{yy}$ evolution at the loaded crack tip of the 2205 duplex steel. $K$=35.8 MPa\textmu m. a) $N=0$ b) $N=2\times10^3$, c) $N=5\times10^3$, d) $N=8\times10^3$, e) $N=11\times10^3$. The colour code shown in a) is the same for each map.
Figure 6.4.1 illustrates the sequence during fatigue of the elastic-plastic strain field at the loaded crack tip ($K=35.8 \text{ MPa}\sqrt{\text{m}}$) for $N=0$ (first load), $N=2\times10^3$, $N=5\times10^3$, $N=8\times10^3$ and $N=11\times10^3$ (Figure 6.4.1a, 6.4.1b, 6.4.1c, 6.4.1d and 6.4.1e respectively). These fields are shown once again in colour-coded maps, with red colour corresponding to higher strain values and blue colour to lower strain values. The sequence of the underlying image microstructure is not shown, since the crack does not advance during fatigue. As $N$ increases, $\varepsilon_{yy}$ also increases and is characterised by a non-uniform behaviour. It is possible, in fact, to identify two distinctive regions in the strain distribution. The strain level of approximately 1% originally localised in the two symmetrical fans (Figure 6.4.1a) with respect to the crack tip expands to cover the whole field with increasing number of cycles. Two areas of intense localised strain ($\varepsilon_{yy}$ between 2% and 5%) can be seen above and below the crack tip over 25 $\mu$m radius. These regions undergo increasing deformation as $N$ increases, and assume a characteristic 'flame' shape that develops along a direction parallel to the crack line. For $N=11\times10^3$ cycles (Figure 6.4.1e), within approximately 50 $\mu$m from the crack tip, there is a prominent region of higher strain that becomes more elongated in a direction of approximately 45° to the crack line. The field symmetry is maintained, and under these cycling conditions, a non-propagating crack is observed.

Further information on the strain field evolution during fatigue can be drawn by examining the deformation of the unloaded specimen at the same stages of fatigue life considered for the loaded material. This is illustrated in the sequence in Figure 6.4.2, which represents the residual strain accumulated at the crack tip during fatigue.
Figure 6.4.2: Elastic-plastic ε_{yy} evolution at the unloaded crack tip of the 2205 stainless steel a) N=2\times 10^3, b) N=5\times 10^3, c) N=8\times 10^3, d) N=11\times 10^3. The colour code shown in a) is the same for each map.
The higher strain region closer to the crack tip for \( N=2 \times 10^3 \) (Figure 6.4.2a) is nearly uniformly distributed over a region of approximately 50 \( \mu \text{m} \) radius around the crack tip. In the remaining field of view there is an inhomogeneous residual strain distribution. As \( N \) increases (Figure 6.4.2b, 6.4.2c and 6.4.2d, corresponding to \( N=5 \times 10^3 \), \( N=8 \times 10^3 \), \( N=11 \times 10^3 \) respectively), there is a net asymmetrical separation of the strain distribution in two 'flames' extending along the crack line, while the strain in the remaining field maintains its inhomogeneity. It is possible that this behaviour may be associated with localised crystalline behaviour. The strain values below and above the crack tip increase, possibly indicating cycling work hardening behaviour of the material. The strain values below the crack tip are in essence unchanged between \( N=8 \times 10^3 \) and \( N=11 \times 10^3 \). However, above the crack tip they increase significantly and the separation in two distinct 'flames' start disappearing. The resulting asymmetric field is not observed when the specimen is under load.

Figure 6.4.3 shows the cross-sectional fatigue behaviour of \( \varepsilon_{yy} \) at the unloaded crack tip for \( x=50 \ \mu\text{m} \). The most prominent feature in the residual strain field evolution discussed previously is clearly observed. For \( N>5 \times 10^3 \) there is the formation of two peaks and strain directly ahead of the crack tip decreases for \( N=5 \times 10^3 \) and \( N=8 \times 10^3 \). Figure 6.4.4 is a plot of \( \varepsilon_{yy} \) as function of \( N \) for three different points in the field of view for both the unloaded and loaded crack tip. Clearly, the strain at a point near the loaded crack tip (diamonds) is higher than the corresponding value for the unloaded crack (circles). For a point far from the loaded crack tip (triangles), after a first increase, the strain reaches a constant value.
Figure 6.4.3: $\varepsilon_{yy}$ at the unloaded crack tip of the duplex steel in function of $y$ for $x=50$ μm.

Figure 6.4.4: $\varepsilon_{yy}$ for duplex in function of $N$ at three different points in the field of view.
A possible explanation for this behaviour is schematically represented in Figure 6.4.5. At the first load (Figure 6.4.5a), the circles represent region of high strain, plastically deformed, corresponding to the red regions in Figure 6.4.1a. The lines represent the slip bands at \( \pm 45^\circ \). The residual plastic zone after cycling is first uniformly distributed around the crack (Figure 6.4.5b, corresponding to the green region of Figure 6.4.2a). When the material is subjected to fatigue, slip planes are activated above and below the crack tip. The fatigue process, however, causes work hardening along these slip planes. The deformation, hence, moves along a softer direction (6.4.5c) causing a decrease in the angle between the zones of high strain. The 'dip' in the measured strain directly ahead of the crack that appears as fatigue takes place is therefore just a consequence of the decreasing angle between the high strain zones and the fact that no slip takes place along the crack line. This decreasing angle can be explained as follow. The shear strain reaches a maximum at \( \pm 45^\circ \) to the crack plane and, for an homogeneous material, the shear strain will also be a maximum in this direction. A material consisting of grains, however, has a discrete range of possible slip directions for each grain. For a given grain, with random orientation, slip will not generally occur at \( 45^\circ \) with respect to the crack plane, but on one of the nearest available slip planes. The slip directions for an ensemble of such grains will therefore be expected to occur over a range of angles centred on \( \pm 45^\circ \). As the crack advances, however, slip on the planes orientated at greater than \( 45^\circ \), or less than \(- 45^\circ \), will be hindered by the work-hardened material lying above and below the \( \pm 45^\circ \), thereby allowing slip to occur preferentially on angles lower than \( + 45^\circ \) or greater than \(- 45^\circ \). This mechanism is favoured by the particular anisotropy of this duplex steel, which has bands of alternate layers of softer and harder material parallel to the crack line, as described in section 5.1.1. It is reasonable to assume that the hard bands constitute a strong obstacle to slip at angles greater than \( + 45^\circ \) or lower than \(- 45^\circ \).

The explanation is therefore a combination of work hardening effects coupled with the material anisotropy and granular structure. This possible process implies a rich 'dynamic' behaviour, in which material properties and strain fields adjust themselves to accommodate fatigue.
In this case also, the study was completed by the observation of the fracture surface (Figure 6.4.6). Contrarily to the austenitic steel case, the crack front during the precrack is flat, even if the stress intensity amplitude applied was the same for both specimens. This is nevertheless acceptable, since the 2205 stainless steel yield stress (510 MPa) is higher than the 316 yield stress (less than 300 MPa). These yield stress values were provided by Avesta Sheffield, Technical Services.

Unfortunately, TEM observation could not be performed at the crack tip of this material, since the gear box driving the spark erosion unit broke down, and due to the age of this unit, it has not been possible to find a replacement in a reasonable time.
Figure 6.4.5: Schematic representation of the strain field evolution during fatigue at the crack tip of the duplex steel.
Figure 6.4.6: Macrograph of the 2205 fracture surface. The crack propagation direction is indicated by the arrow.
6.5 Comparison of the strain fields in the two steels

The 316 and 2205 steels behaviour shown in the previous sections can be compared only up to $N=2\times10^3$ cycles. The strain field is more intense ahead of the austenitic steel and it extends over a region that is larger than ahead of the crack tip in the duplex. This result is a consequence of the higher yield stress of the 2205 steel.

It is also interesting to compare the different strain field shapes to fields of similar shape found in the literature. The $\varepsilon_{yy}$ observed at the tip of growing cracks in aluminium alloy was determined ahead of an initial flat crack, obtained by fatigue precracking the specimen at low stress amplitude ($\Delta\sigma=34.5$ MPa). The experiment was also performed ahead of an initial slant crack, obtained by fatiguing at high stress range ($\Delta\sigma=172.5$ MPa). Ahead of the slant crack a single lobe of large strain above the crack tip was present and it was attributed to progressive yielding along a single 45° plane through the thickness. Ahead of the flat crack, the initial strain field formed a twin lobe shape. As the crack grew, one lobe began to dominate, preceding the path that the crack took during the transition to 45° slant fracture, after which the strain had qualitatively the same shape as the field ahead of the initial slant crack.

As observed previously by examining the fracture surface (Figures 6.2.7 and 6.4.6), it is clear that during the precrack the austenitic steel is characterised by a slant mode of fracture, while the crack front in the duplex is flat. The strain field at the crack tip of the austenitic field shows a larger dominant lobe, while the strain fields for the duplex show twin lobes. It is thus possible that the shape of the strain fields is related to the mode of fracture that predominates during fatigue, and therefore to the loading amplitude applied.

6.6 Conclusions

The behaviour during fatigue of an austenitic and a duplex stainless steel has been studied by various techniques. The $\varepsilon_{yy}$ strain fields were calculated by the displacement measured using a high magnification moiré interferometer for different numbers of cycles. The dislocation distributions ahead of the crack tip in the austenitic steel were observed by transmission electron microscopy. Conventional
optical microscopy was employed to observe the material microstructure and the fracture surfaces.

The cyclic elastic-plastic strain increases as the number of cycles increases. The inhomogeneous and non-uniform field is characterized also by a distinctive change of shape and a loss of symmetry with an increasing number of cycles both for austenitic and duplex steels.

The $\varepsilon_{yy}$ at the crack tip of the austenitic steel was characterised at first by a two lobes asymmetrical distribution. One lobe became dominant as the number of cycles increased. This behaviour is explained by a shear mode growth of the precrack. The dislocation distributions at the crack tip after fatigue show a possible unzipping mode of crack propagation on alternate slip bands controlled by local material work-hardening.

The evolution of the strain field ahead of the crack tip in the duplex specimen suggests also in this case that a local work hardening in the material influences the orientation of preferred slip bands, which in turn control the strain values. The strain fields clearly indicate that the crucial region involved in fatigue mechanisms is concentrated over a small region up to approximately 100 $\mu$m ahead of the crack tip. The field inhomogeneity could be associated with different local crystal properties, due to the two phase composition of the steel. Changes in the microstructure, resolvable by optical microscopy, were not observed during fatigue.

In conclusion, quantitative values for $\varepsilon_{yy}$ can be provided over a meso-scale ahead of the crack tip for fatigued stainless steels, by using high magnification moiré interferometry. The experimental data obtained are eventually suitable for the implementation of theoretical models. The observed strain distributions are the result of the complicated material response to the applied stress field over a large number of cycles, which can be detected by this experimental technique.
Chapter 7

Conclusions and future work

7.1 Conclusions

To address the need for accurate experimental data at and near the tip of cracks in stainless steel, a technique able to measure in-plane displacement fields over a few tens of crystal grains ahead of the crack tip has been developed in this work.

The development of the technique included the construction of a compact high magnification moiré interferometer, comprising a built-in microscope head and phase-stepping optics. By switching from laser illumination to white light illumination, displacement fields can be recorded in exact registration with the underlying specimen microstructure. The microscope design was based on an existing design, which has been improved here through a five-fold increase of the recording camera pixel count. This was achieved by using a high resolution Kodak Megaplus 1.4 camera that allowed the registration of microstructure picture of good quality and better resolution in the determination of fringe edges for phase calculation. The drawback in using this camera was the vibration originated by the electro-mechanical shutter that was, however, successfully eliminated.

The technique was optimised principally for application to stainless steels. In particular, to obtain clear stainless steel microstructure images, the choice of the white light illumination system and the preparation of the specimen surface constituted a substantial part of the work. Development of the technique also involved the design of specimen and grips that could be used with the configuration of the experimental set-up, resulting in virtually no constraint on the specimen geometry, which is often the case when stereoimaging techniques or other moiré configurations are employed. Further advantages of this technique are that a reference state is available and no post-processing steps are necessary to reduce the noise, in contrast to techniques based on speckle effects.
With a reference state, it has been possible to study loaded and unloaded specimens during fatigue.

The in-plane component of the displacement at the tips of cracks was measured by using temporal phase stepping and spatial unwrapping. Whole field contour maps were thus obtained over a sub-millimetre field of view, with a sensitivity of 0.417 μm/fringe and an accuracy better than 40 nm. The strains at crack tip, which represent the elastic-plastic deformation controlling fatigue, were thus calculated from the gradient of the displacements. Strain calculation also allowed the elimination of rotation component in the displacement and therefore clearer interpretation of results.

To summarise the advantages of this technique, it can be said that this is a very sensitive method to measure small displacements in a sub-millimetre field of view. It can image the underlying specimen microstructure in exact registration with the displacement fields and there are no constraints on the specimen geometry. The crack tip can be precisely located and a reference state is available, so that displacement can be measured during fatigue.

Certainly this technique has a few disadvantages. For example the alignment procedure is time consuming (and an operator that does not happen to be tall enough has to stand on a ladder). The specimen preparation is also time-consuming, sometimes requiring a few weeks. The size of the data files goes from over 1 Mbyte for each recorded fringe intensity to over 5 Mbyte for the calculated phase values, which gives storage problems when the data are recorded, for example, at several different loads. Computational time for the strain evaluation can be long, up to four hours with the gauge length employed in this work. These inconveniences are nevertheless minor compared to the advantages of the technique.

This technique has been applied here for the first time to the best of our knowledge to study deformation fields near the crack tip in austenitic and duplex stainless steels. The possibility of superimposing displacement and strain field to the underlying microstructure revealed possible effects of grain boundaries, grain size and twin crystals on the fields contours.

The highest strain concentration was observed behind the crack tip. This could be due to effects of the specimen geometry.
The experimental results obtained ahead of the tip of monotonically loaded cracks have been compared to the fields predicted by current theoretical models. The HRR model, for which the region of validity is a debated topic of research, was not found to represent the experimental data. The values obtained for the strain hardening exponent with fitting procedures were either in inhomogeneity, which are not taken into account by HRR fields. It must also be said that the HRR is valid for ductile materials and at condition that there is not appreciable crack tip plasticity.

A logarithmic type of singularity, which describes the crack tip fields on empirical basis, was also compared to the results. Disagreement was also found in this case. The results have been analysed disagreement with the expected values or less than one, which is physically unacceptable.

The origins of the disagreement could be identified in thickness and geometry effects, but also in the material anisotropy and here for the first time in terms of a recent model that takes into account strain gradient effects. This model also could not reproduce the experimental data, in particular with regards to the field shape and the values obtained for the hardening exponent.

The possibility of locating exactly the crack tip and the availability of a reference state allowed the quantitative evaluation of strain fields during fatigue. However, the lack of analytical models describing elastic-plastic near-tip strain fields during fatigue did not allow a quantitative interpretation of the experimental data. A qualitative interpretation of the evolution of these fields has nevertheless been attempted, by complementing the interferometric measurement with transmission electron microscopy observation and visual examination of the fracture surfaces. The variation in intensity and shape of the strain has then been related to local material work hardening and to the intensity of the applied loads. The alternate unzipping model of crack propagation, controlled by local work hardening has been proposed to explain the observed dislocation distribution ahead of the crack tip in one of the austenitic specimen. Unfortunately, a complete set of data for different stage of fatigue life could not be obtained for the austenitic steels. The problem faced was repeated crack branching when trying to precrack other specimen.
7.2 Future work

The technique described in this dissertation could be a powerful tool for improvement in understanding fatigue mechanisms. It is, in fact, able to provide the accurate experimental near-tip fields that are required to implement and develop theoretical models. Complete information on the near-tip fields requires the measurement of both in-plane displacement components. This can be achieved by replicating a crossed grating onto the sample and by constructing a second interferometer axis, with another set of optical fibres and optical components. The present set-up has been constructed with this addition in mind. The technique could then be completely automatised by automatically controlling the shutters to switch between the two axis of the interferometer and between white light and laser light illumination.

By automatising also the specimen polishing procedure, the technique would be less time (and fingers) consuming. Enquiries have been made towards the purchase of an automatic polishing equipment suitable for the specimen size involved in fatigue tests.

The particular features of this technique could also help to identify the microstructural parameters involved in fatigue and fracture processes. For example, experiments could be carried out at different precrack length, in particular using a precrack length of the order of a few microns. It is in fact expected that microstructural effects on crack propagation mechanisms are dominant for crack length of this order of magnitude. This kind of experiments, however, would be more feasible and realistic if an automatic precracking facility were available. Also, to obtain precrack length of the order of a few microns, it would be appropriate to use the spark erosion technique. In this way, the crack length could be carefully controlled and the residual stress around the crack tip would be virtually none, avoiding so the necessity of annealing procedure.

Furthermore, the technique characteristics such as the possibility of measuring near-tip deformation fields, to observe the underlying microstructure and practically no constraints on the specimen geometry, could allow one to investigate the material behaviour if parameters such as specimen geometry or applied loads were varied. By testing specimens of different thickness, thickness effects on surface deformation could also be investigated in detail.
Another idea for future exploitation of this technique could be to use the experimental data obtained as input values for finite element analysis\textsuperscript{99}. A pilot test could be performed by studying aluminum specimens, since aluminium is a better studied material. Subjected to successful results, the technique could then be confidently applied to the less studied stainless steels.
References


82. Renishaw plc, Technical Services.


Appendix I

List of publication

Conference papers:


Journal papers
