Calculation of Energy Release Rates For Delaminations in Composite Materials Using Phase Shifting Moiré Interferometry

A thesis submitted in partial fullfilment
of the requirements for the degree of
Doctor of Philosophy

Kenneth Eugene Perry, Jr.

Department of Mechanical Engineering
Division of Mechanics of Materials
University of Strathclyde
Glasgow, Scotland
1993

Declaration of Author's Rights

Declaration of originality

I declare that this document embodies the results of my own work and that it has been composed by myself alone. Following standard academic conventions, I have made due acknowledgement of the work of others.

Signed: _____

Date: 22 Sept 93

Contents

List of Figures

List of Tables

Abstract

An experimental technique was developed to measure the energy release rate for delaminations in composite materials. The method can be used in cases where conventional methods may not be applicable, such as problems involving multidirectional laminates or arbitrary specimen geometry.

The work involved the experimental and computational enhancement of the optical technique of moiré interferometry, including alternative grating technology, image processing of fringe patterns, the implementation of phase shifting and the automation of the data reduction process. The improvements provided a practical experimental means for determining full-field displacement fields surrounding delaminations, together with a convenient method for using these displacement fields to calculate energy release rates via the J-integral.

Three types of experiments were performed on carbon-fibre/epoxy unidirectional laminates to validate the new approach. The results compared favourably to those obtained by conventional methods. Several experiments were also performed on multidirectional laminates to demonstrate the extended capabilities of the new method.

It was concluded that the new method is not particularly well suited to the measurement of critical energy release rates, but it should be a valuable tool in the validation of computational methods.

Chapter 1

General Introduction

Composite materials have a long history of usage. Plywood, corrugated cardboard and concrete are examples of typical composite material systems. More recently, fibre-reinforced resin composites that have high strength-to-weight and stiffness-to-weight ratios have become important in applications such as aerospace design.

Laminated polymeric composites offer an unique advantage over conventional engineering materials, in that structural properties can be tailored to suit specific applications. This is accomplished by combining several materials in a prescribed manner, thus taking advantage of the individual properties of each constituent material.

However, the inherent anisotropy and the discrete layer-by-layer fabrication method of composite materials leads to mechanical behaviour and failure characteristics that are quite different from those of conventional homogeneous materials. Consequently, failure modes such as delamination are a characteristic problem limiting the capacities of composite structures.

In this chapter, an attempt is made to clarify the problem of delamination fracture, and to distinguish it from the general background of failure in composites. A brief discussion of composite materials and their fracture behaviour is first discussed, leading up to a clarification of the objectives of this work. The fracture mechanics concept of the energy release rate is then introduced, and finally, the experimental technique of moiré interferometry is briefly described.

1.1 Composite Materials and Delamination

The word 'composite' in composite materials signifies that two or more materials are combined on a macroscopic scale to form a useful material. For example, concrete is composed of sand and rock, bound together by a mixture of cement and water. Although the cement bond is weak, the compressive strength of concrete is derived from the rock. The cement merely acts as a 'matrix' material, binding the sand and rock particles in place.

There are three commonly accepted types of composite materials [1]:

- *fibrous composites,* which consist of fibres in a matrix
- *laminated composites,* which consist of layers of various materials
- *particulate composites,* which are composed of particles in a matrix

In this work, emphasis will be placed on the commonly utilized combination of the first two—particularly the laminated carbon-fibre/epoxy material system.

Lamination is the process by which unidirectional laminae are pressed into a laminate. Typically the individual laminae are orientated differently, see Figure 1.1, to provide specific strength or stiffness properties in certain prescribed directions. Other fabrication methods, such as winding and moulding can also be used when manufacturing composite components. From a macroscopic point of view, and regardless of the fabrication method, the mechanical behaviour of the resulting structure can be optimised for any combination of in-service requirements.

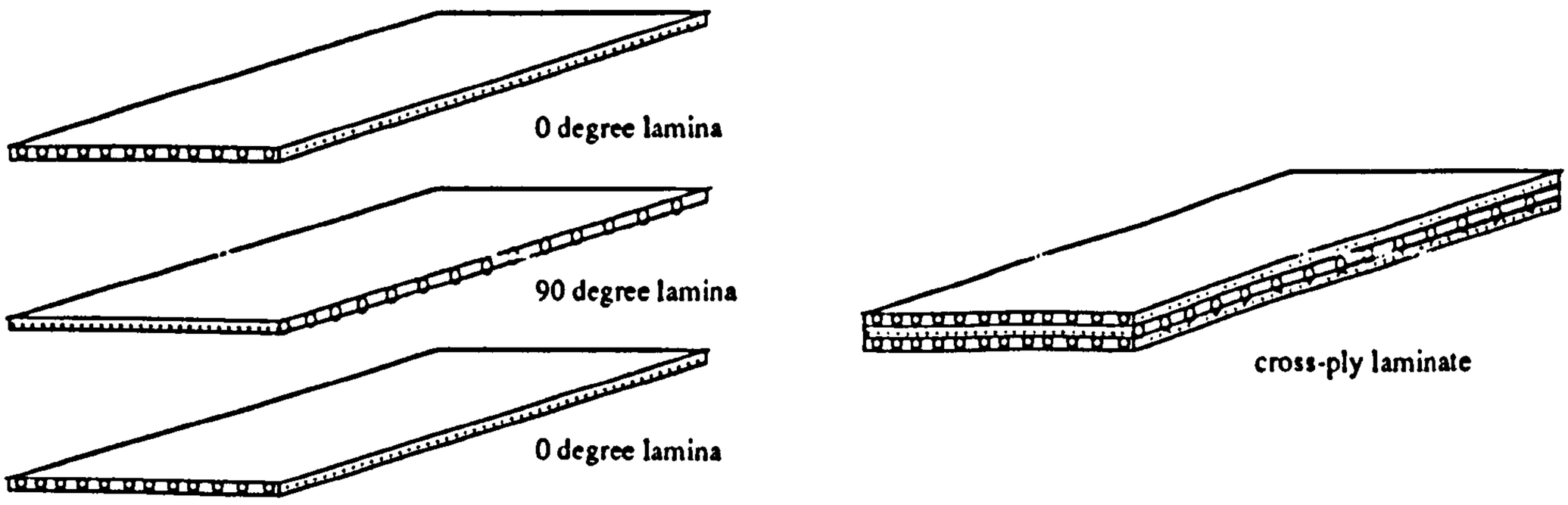


Figure 1.1: A typical cross-ply laminate design.

The strength and stiffness of carbon-fibre/epoxy materials are derived from the properties of carbon fibres. Long fibres are inherently much stronger than the same material in bulk form. For example, glass in bulk will fracture at stresses as low as only a few MPa, while thin glass fibres can have strengths of up to several GPa. This paradox is due to the more perfect structure of a fibre—the crystals are aligned in the fibre along the fibre axis. Moreover, the statistical likelyhood of encountering a flaw in the smaller cross-section of a fibre is reduced.

The primary function of the epoxy in carbon-fibre/epoxy materials is to bind the fibres and individual laminae together. Epoxy does a reasonable job and is readily available, but other matrix materials are commonly used. For instance, poly(ether-ether ketone) (PEEK) has a much higher fracture toughness than epoxy and composites with a PEEK matrix exhibit much higher delamination and impact resistance than those made with epoxy.

Although the matrix material is not directly responsible for bearing load, it does play a major role in the transfer of shear stresses between the fibres. The matrix material is also required to protect the composite from impact damage by absorbing energy during dynamic loading. The nature and extent of damage progression will be directly determined by the properties of the matrix material.

One of the fundamental difficulties with laminated structures is their tendency to *delaminate*. Delamination begins from internal defects, such as voids due to imperfect bonding between adjacent laminae or small cracks in the matrix material. Such defects can arise during the fabrication process or they can originate in service as a result of impact damage or excessive loading levels.

Another very common source of delamination in multidirectional laminates results from interlaminar stresses which can develop in a boundary layer region along the free edges of bonded dissimilar materials. This problem is commonly referred to as the 'free-edge' problem, and has received considerable attention, dating back to the early 1970's [2].

The first complete three-dimensional analysis of interlaminar stresses in laminated composites was presented by Pipes and Pagano [3]. Figure 1.2 illustrates their problem. They used the finite difference method to solve the elasticity problem for an orthotropic, finite width laminate subject to axial loading. Their results verified the existence of all three components of interlaminar stress, σ_z, τ_{zy}

and τ_{zx}, in a boundary region near the free edges of the specimen. Furthermore, these stresses were found to increase in magnitude toward the free-edge—possibly even becoming singular (or zero in the case of τ_{xz}) at the surface.

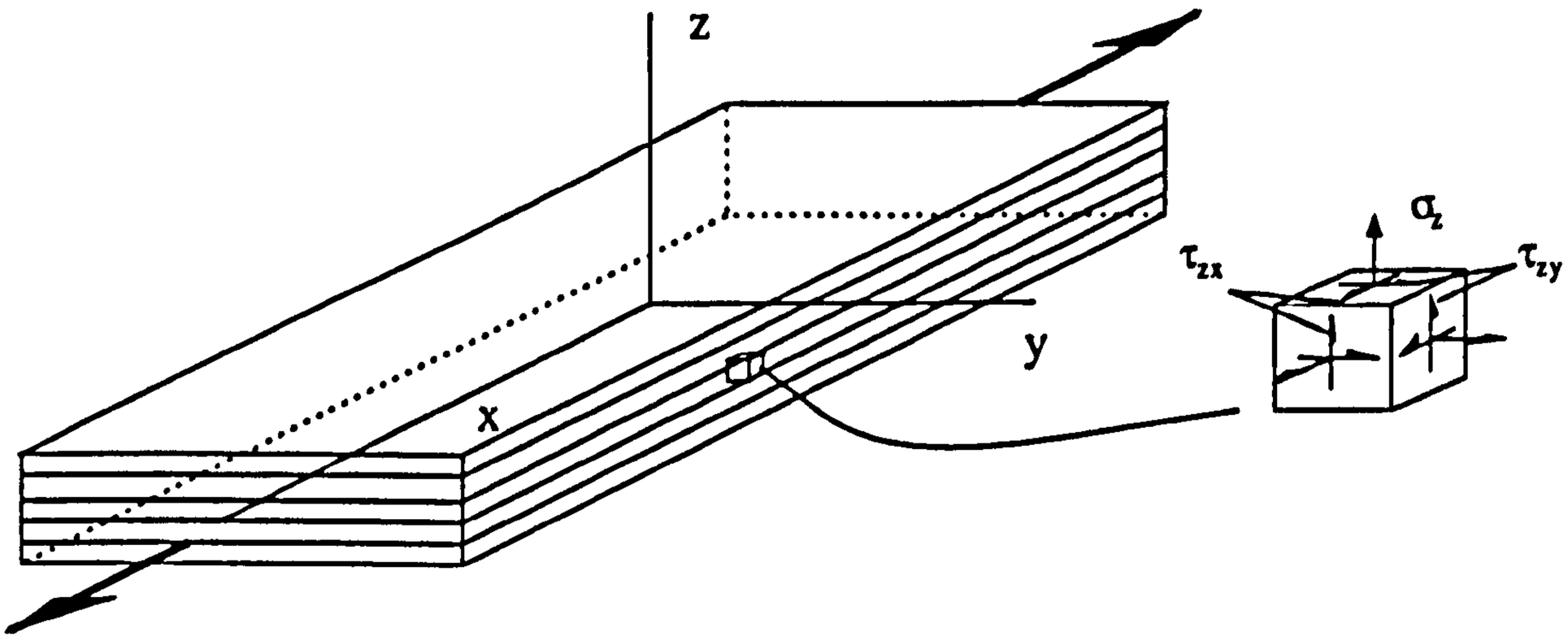


Figure 1.2: Notation for the free-edge stress problem.

While the exact nature of the stress field near the free-edge of a composite has remained a difficult theoretical undertaking, some general findings have been established. Most importantly, it has been determined that laminate *stacking sequence* plays a significant role in the nature and magnitude of the free-edge stresses. Over the years, various results for standard composite stacking sequences have appeared, indicating which laminate configurations should be avoided in design applications.

Delamination in composite structures leads to an immediate reduction in stiffness and load bearing capacity of the component, particularly in compression. Once initiated, delaminations tend to propagate under either static or fatigue loading conditions, further degrading the mechanical properties of the structure. Unlike fracture in homogeneous and isotropic materials, damage initiation and progression in composites is governed by a host of mechanisms, making both the onset *and* propagation of delaminations difficult to predict in practice. Indeed, the application of conventional fracture mechanics methodology to composite materials presents a great challenge.

The success of applied mechanics with regard to homogeneous materials depends upon the ability to describe material as a statistically uniform *continuum*. In such cases, an arbitrary infinitesimal volume will adequately represent mate-

rial properties throughout the entire body. Composite materials are inherently heterogeneous—and a true continuum analysis would require modelling at the level of the fibre/matrix interface. Such an analysis would be intractable, and have little significance to reality, since the distribution and interaction of flaws in a real composite must ultimately be treated as a random phenomena.

Researchers have adopted two different viewpoints from which to investigate the behaviour of composite materials. These viewpoints are summarized in the following definitions [1]:

- *Micromechanics*—The study of material behaviour wherein the *interaction* of the constituent materials is examined in detail as part of the definition of the behaviour of the *heterogeneous* composite material.

- *Macromechanics*—The study of composite material behaviour wherein the material is presumed *homogeneous* and the effects of the constituent materials are detected only as averaged *apparent* properties of the composite.

The micromechanics approach has been successful in predicting stiffness properties of composite materials, based on, for instance, the 'rule-of-mixtures' approach. On the other hand, macromechanics provides an adequate and useful description of composite material behaviour at the level of the structure. Thus, it is possible to *predict* lamina properties by the procedures of micromechanics, and use these properties in a macromechanical analysis of a structure.

Unfortunately, the problem of predicting general laminate strength properties has had much less success. Strength properties of composites are difficult to characterise because the nature of damage initiation and growth in laminates will depend both upon the random distribution of microflaws in a composite and upon previous damage patterns. Such a dependency renders a deterministic approach difficult.

In the specific case of delamination fracture, however, at least two simplifying assumptions are strongly justified—

- Delaminations generally propagate in a self-similar manner. Since the interfacial strengths of composite laminates are normally weaker than the surrounding material, the distributed type of damage that plagues laminate strength predictions does not significantly influence delamination growth.

This allows the engineer to study the growth of a delamination as an isolated event, with a known trajectory.

- For many types of matrix materials, delamination occurs as a brittle fracture of the interface. Thus, linear elastic fracture mechanics concepts (particularly the energy release rate concept) can be used to characterize interfacial fracture toughness.

Ideally then, it should be possible to perform simple experiments on well designed coupon samples to determine the energy release rate for any particular material system. These properties can then be incorporated into a computer model of a general structure with a simulated delamination to determine under what conditions the delamination would propagate. This is commonly done for conventional homogeneous engineering materials such as steel and aluminium. However, before this approach can be confidently adopted for composite structures (with regard to delamination fracture), the computer models used to determine the stress fields near delaminations must be validated.

The objective of this work, therefore, is *to develop an experimental method that is capable of measuring the energy release rate for an arbitrary delamination under general loading.* Then it will be possible to validate computer predictions of energy release rates for complex delamination problems.

1.2 Fracture and the Energy Release Rate

Fracture mechanics is an engineering discipline that primarily draws upon the fields of applied mechanics and materials science. In doing so, it attempts to make a connection between the morphology or structure of a material and its fracture properties. In its most basic form, fracture mechanics relates the maximum permissible applied loads acting upon a structural component to the size and location of a crack—either real or hypothetical—in the component. In this section, the basic ideas of fracture mechanics are presented, and discussed relative to their applicability to delamination fracture in composites.

Underlying fracture mechanics methodology is the idea that there exists a driving 'force' behind every crack and an inherent resistance of a material to impede crack growth [4]. The driving 'force' is the elastic strain energy supplied

to the crack by the action of external loads. When the elastic strain energy available to grow the crack exceeds the materials ability to resist fracture, the crack is assumed to propagate, and energy is released through the creation of new fracture surfaces. The rate at which energy is released during crack growth is an important fracture parameter, and is known as the energy release rate.

The formal definition of the energy release rate is arrived at by considering the energetics of a cracked body under load. The potential energy, Π, of such a body is written

$$\Pi = U_e - U_w,$$

where U_e and U_w are, respectively, the internal elastic strain energy and the external work done on the body.

Now consider an infinitesimal *virtual* increment of crack growth, da. Differentiating the above equation with respect to da yields

$$\frac{d\Pi}{da} = \frac{d}{da}(U_e - U_w) = \frac{dU_e}{da} - \frac{dU_w}{da}.$$

The term $d\Pi/da$ represents the energy released from a body during the virtual extension of a crack by da. This quantity of energy is available for driving the crack. For the case of brittle fracture when no energy is dissipated in non-recoverable deformation processes, such as plasticity or microcrack formation, the energy release rate, denoted G, is defined as

$$G \equiv -\frac{1}{b}\frac{d\Pi}{da} = \frac{1}{b}\left(\frac{dU_w}{da} - \frac{dU_e}{da}\right), \tag{1.1}$$

where b is the width of the specimen. If the energy available to drive the crack is less than that required to propagate the crack, then crack growth will *not* occur, and the corresponding value of G is referred to as sub-critical. Thus, G is seen to be a measure of the energy that is *available* to create new fracture surfaces—that is, it is the amount of energy that *would* be released from the body *should* the crack extend by da.

A materials resistance to fracture can be characterised by the value of G at fracture—the critical value, denoted G_c, and this value is assumed to be a material property. The utility of the energy release rate in characterizing fracture depends upon the ability to convert loads, crack lengths, etc., for any given situation, into

an equivalent expression for G. If G_c is known for a given material system, then it is possible to *predict* when crack growth will occur in a structure for any given loading situation, provided the appropriate expression for G is available.

As an example of how G can be used in practice, consider the symmetric double cantilever beam (DCB) specimen in Figure 1.3. From simple beam theory, the deflection of each arm can be written

$$\delta = \frac{8Pa^3}{Ebh^3}.$$

If during an increment of crack growth, the load point displacement δ remains constant, $dU_w/da = P\Delta\delta = 0$, and

$$\frac{d\Pi}{da} = \frac{dU_e}{da} = \frac{d}{da}\frac{P\delta}{2} = \frac{-12P^2a^2}{Ebh^3}.$$

where the minus sign denotes a net internal strain energy *decrease* due to the diminished load during fixed grip crack extension. The value of the energy release rate is found to be

$$G = -\frac{1}{b}\frac{d\Pi}{da} = \frac{12P^2a^2}{Eb^2h^3}.$$

Two alternative expressions for the energy release rate involving the compliance, C, of the specimen were presented by Irwin [5]

$$G = \frac{P^2}{2b}\frac{dC}{da} = \frac{\delta^2}{2bC^2}\frac{dC}{da}.$$

For the DCB specimen of Figure 1.3, the compliance can be written

$$C = \frac{8a^3}{Ebh^3},$$

and so, as before,

$$G = \frac{12P^2a^2}{Eb^2h^3}.$$

The compliance approach is particularly useful in practice, since C need only be determined as a function of a, either experimentally or analytically, for G to be found. This approach will be discussed in more detail in Chapter 2.

Thus far nothing has been said about the *nature* of the fracture—i.e., whether a crack will proceed in a stable or an unstable manner. This will be determined

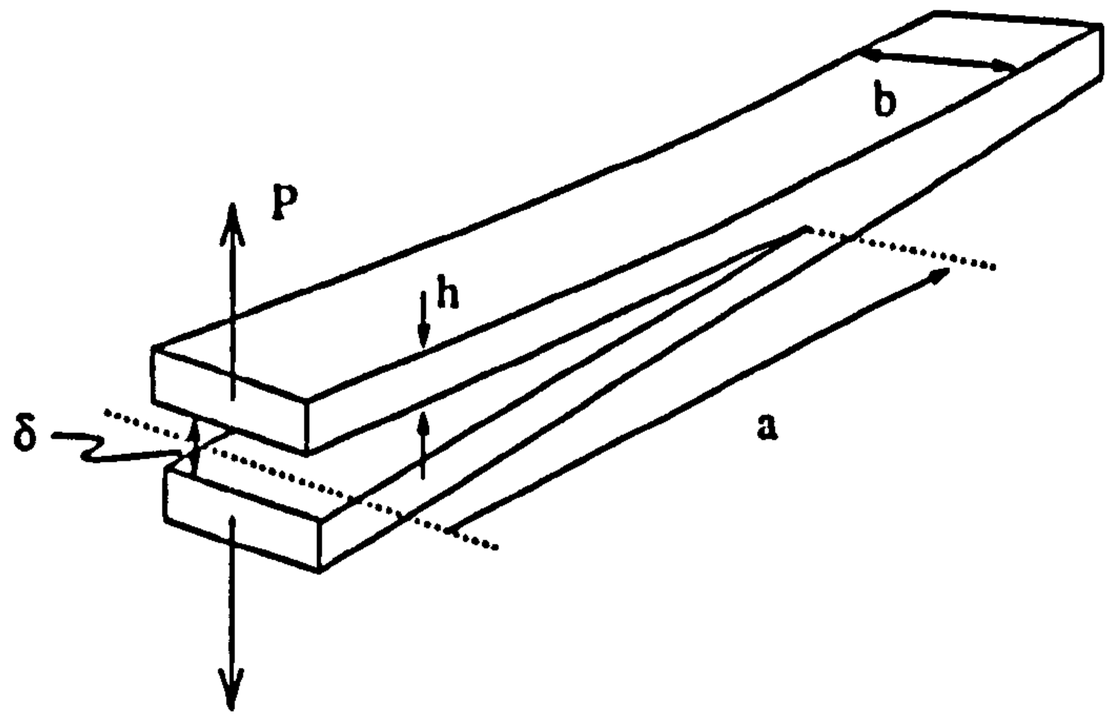


Figure 1.3: Symmetric double cantilever beam specimen.

by the rate of change of the available energy with respect to crack growth,

$$\frac{d}{da}\left(\frac{dU_w}{da} - \frac{dU_e}{da}\right).$$

When the above expression is positive, the crack driving force *increases* as the crack extends. If G_c is a constant, this implies that 'unstable' fracture will occur—i.e., there will continue to be more energy available to drive the crack, as it grows, than the crack growth process can absorb. If, however, the above expression is negative, or if G_c is not a constant, and

$$\frac{d}{da}\left(\frac{dU_w}{da} - \frac{dU_e}{da}\right) < \frac{dG_c}{da},$$

then the structure will manage to absorb the energy, and the crack will proceed in a stable manner, as more energy is fed in through the load-points.

One particular advantage of characterizing crack growth in terms of energy is that this approach effectively circumvents the difficulties associated with crack tip stress singularities. This is important since the local stress field associated with delamination fracture is difficult to characterize mathematically. When the total energy balance of a laminate is considered, the details of how the energy is distributed is not required, and no assumptions need to be made regarding the local crack tip stress and strain fields.

Another useful form of the energy release rate for a two dimensional body can be derived directly from Equation 1.1, by explicitly considering the internal elastic potential energy within a core region surrounding a delamination. By invoking

fundamental laws of solid mechanics, the path independent integral known as the J-integral can be shown to be exactly equivalent to the previous definition of G. By definition, J is written

$$J = \int_{\Gamma} (W n_x - T_x \frac{\partial u}{\partial x} - T_y \frac{\partial v}{\partial x})\, ds,$$

using the notation illustrated in Figure 1.4.

The physical significance of the J-integral is that it represents an expression for the conservation of elastic energy in the form of a contour integral enclosing a crack tip. The energy release rate, G, and J are identical—provided the material behaves in a linear elastic manner, and fracture is assumed to be brittle.

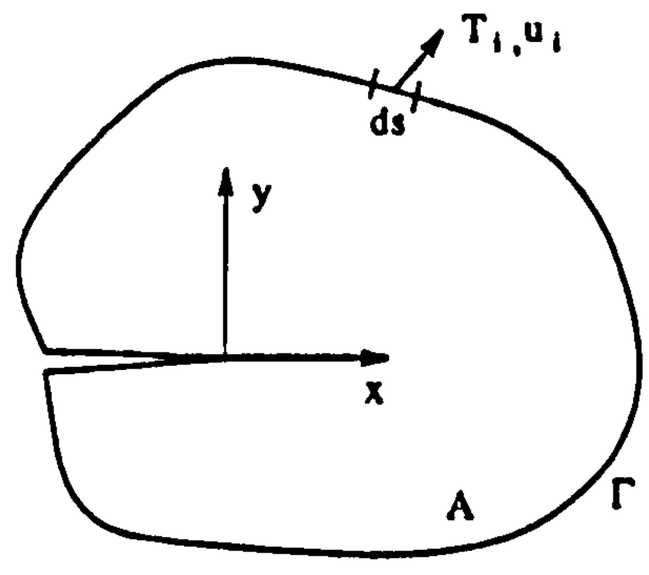


Figure 1.4: Area and contour notation for the J-integral.

Smelser and Gurtin [6] demonstrated that the J-integral remains valid for the case of bimaterial fracture, provided the dissimiliar materials are joined along a plane that lies parallel to the crack. It is also required that certain boundary conditions are met along the bond line—specifically, the forces must balance and the displacements should be continuous. Assuming that a body is composed of two materials and joined along the x axis, these conditions are:

$$\begin{aligned} \sigma_{xy}(x, 0^+) - \sigma_{xy}(x, 0^-) &= 0 \\ \sigma_{yy}(x, 0^+) - \sigma_{yy}(x, 0^-) &= 0 \end{aligned}$$

and

$$\begin{aligned} u(x, 0^+) - u(x, 0^-) &= 0 \\ v(x, 0^+) - v(x, 0^-) &= 0 \end{aligned}$$

Note that the continuity of displacements implies that the horizontal derivatives of $u(x,0)$ and $v(x,0)$ must also be continuous across $y = 0$:

$$\frac{\partial u(x,0^+)}{\partial x} - \frac{\partial u(x,0^-)}{\partial x} = 0$$
$$\frac{\partial v(x,0^+)}{\partial x} - \frac{\partial v(x,0^-)}{\partial x} = 0$$

The motivation for using the J integral as a measure of G in this work is discussed below.

1.3 Photomechanics and Local Displacement Fields

The direct definition of the J-integral provides a convenient means for computing the energy release rate, provided the displacements u_i and tractions T_i are available. One of the few photomechanics techniques that can supply such information is moiré interferometry.

Moiré interferometry is a full-field optical method that provides in-plane displacement measurements of deformed specimens. The displacement information is in the form of a pair of fringe patterns, representing the components of x- and y-field deformation. The method is based on optically interrogating a diffraction grating bonded directly to a specimen. The displacements on the surface of the specimen directly deform the diffraction grating, and a two-dimensional displacement map with a sensitivity on the order of 0.4–1 μm is readily obtained [7].

The moiré technique has been widely used in the field of experimental stress analysis for a number of years. In fact, one of the earliest experimental verification of the interlaminar stresses in a composite laminate, performed by Pipes, Byron and Daniel [8], utilized the geometric moiré method. Here the technique was used to examine the surface displacements of symmetric cross-ply laminate under axial tension. At various loads, moiré fringe patterns were recorded, as typified in Figure 1.5. The fact that the fringes were not perpendicular to the edge of the specimen indicated the existence of non-zero shear stresses at the free-edge of the laminate. Comparison of the experimental displacement results to theoretical elasticity predictions confirmed the existence of the free edge stresses.

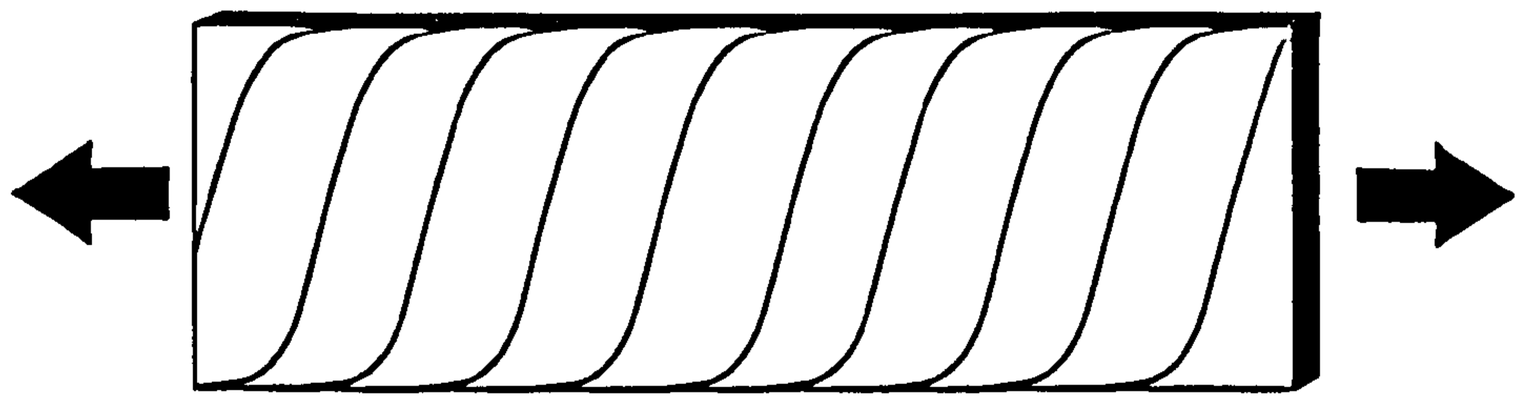


Figure 1.5: Schematic of typical moiré fringe (longitudinal) displacement contours, experimentally confirming the nonzero free-edge shear stress distribution in cross-ply laminates.

The idea of utilizing the detailed strain fields obtained from moiré interferometry to evaluate the J-integral was first established by Gray *et al.* [9]. Kang *et al.* [10] used white light moiré to evaluate the J-integral for stable crack growth in aluminium tensile specimens. As a consequence of the work by Kang *et al.* and also by Dadkah *et al.* [11], serious questions were raised in the computational mechanics community regarding the use of the J-integral as a single parameter to characterize ductile fracture.

With respect to the deformation of composite materials, a substantial amount of work has been done by Post (see for example [12]). Han [13] has developed a moiré interferometer ideally suited to the study of deformation at composite interfaces. His interferometer uses a unique form of phase shifting and an immersion imaging system to achieve high-sensitivity, noise rejected deformation measurements. The application of such a technique to the fracture problems with which we are concerned here, would, however, be very difficult.

Liechti [14] presented a novel photomechanics technique to measure the crack opening displacement in bimaterial blister specimens. The interferometrically determined displacements were used to compute stress intensity factors for three-dimensional blisters loaded by internal pressure.

The technique of moiré interferometry was also used in a limited capacity by Takahashi *et al.* [15] to experimentally determine crack opening displacement (COD) measurements of composite double cantilever beam specimens. These authors compared experimental measurements to theoretical predictions based on conventional and corrected beam theory.

An experimental technique such as moiré interferometry has the advantage of

providing a direct representation of the surface strains of an arbitrary specimen, under general loading. This is particularly useful when the use of theoretical predictions are not available or where they have not been verified. For this reason, an experimental photomechanics approach is required to provide much needed additional information.

However, a refinement of the moiré interferometry technique is necessary before reliable and meaningful results can be obtained in composite materials. One particular problem stems from the conventional method of introducing diffraction gratings onto the specimens. The standard method involves replicating the diffracting profile of a master gratings in a layer of epoxy or silicone on the specimen surface. If care is not taken, this method introduces a boundary layer, effectively the thickness of the epoxy or silicone, that will to some extent, smear-out the variations in the true surface strains of the deformed specimen. When an investigation is concerned with quantifying inter-layer and intra-layer stresses, a grating boundary layer with a thickness on the order of the plies can significantly obscure the results.

Another requirement for the deployment of moiré interferometry in the analysis of composite laminates is a refinement of sensitivity. Because of their brittle nature, composite materials will not deform large amounts before they fracture. The gauge length necessary to distinguish between inter-ply strains puts a further demand on the system. Thus, greater measurement sensitivity is required, which translates into the need to use high frequency diffraction gratings, and a shorter wavelength of light in the interrogation.

A final consideration with the practical implementation of any photomechanics technique is how to process the large amount of data that is produced. For this work, it will be necessary to compute J-integral values from pairs of u- and v-field fringe patterns. Assuming the use of a standard video acquisition system, each fringe pattern represents approximately 1/4 MBytes of information. Therefore, an important aspect of this work is to develop both the hardware and software needed to facilitate the data reduction process, so that the mechanics analysis can proceed uninhibited.

The use of any experimental optical method in the study of delamination fracture is faced with a serious limitation that can not be circumvented. Since moiré

interferometry can measure only surface strains, there is an immediate restriction as to where displacement information can be recorded. If the specimens of interest were thin, and the stress state one of plane stress, then the surface strains would be representative of those throughout the entire specimen. Unfortunately, the dimensions of a delamination specimen prohibit such a straightforward relationship between the surface strains and the strains on the interior of the specimen.

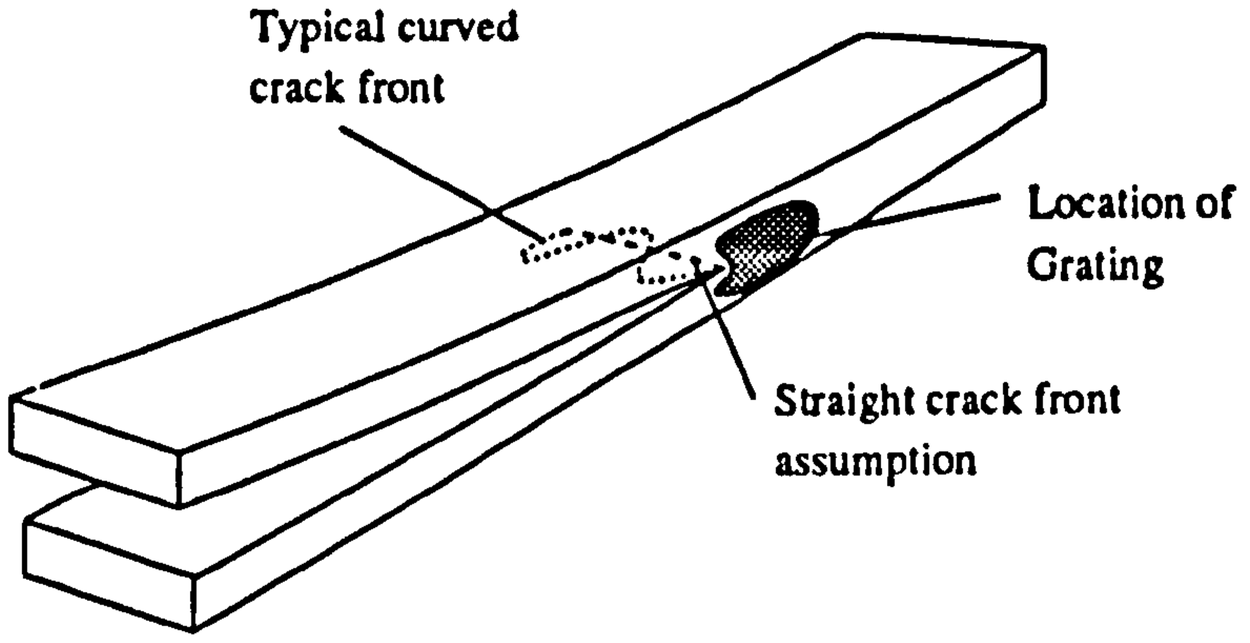


Figure 1.6: Location of diffraction gratings for moiré interferometry study of delamination fracture in composite laminates.

Delamination specimens are typically long narrow beams. There are two possible locations for the diffraction grating—on the top surface of the specimen, and on the thin edge. A grating on the top of the specimen would not provide the information necessary for computing the energy release rate, since it is not in direct contact with the fracture surface. However, a diffraction grating located on the thin edge of the specimen, as illustrated in Figure 1.6, can provide the relevant information.

Due to the three-dimensional nature of the stresses typical of cross-ply laminates, the surface strains will not reflect the average strains throughout the width of the specimen. In fact, the actual deformation is expected to change continuously to some extent along the entire length of the delamination front. The stress state at the edge of the specimen will reflect the true mechanical behaviour of the laminate at the edge, so the data obtained from the moiré fringe patterns can still be confidently used to calculate the energy release rate there. The limitation, however, is that the measured value of G at initiation may not represent the *critical* value for the specimen, for two important reasons—

- First of all, the actual delamination will not necessarily consist of a straight crack front, as shown by the dashed line of Figure 1.6. Rather, a curved, fully three-dimensional crack front may develop during delamination growth. This implies that the value of G for any particular delamination shape may depend on the location along the crack front. The moiré technique can only provide data at the observation edge, and there is no purely experimental way to determine the relationship between the energy release rate at the edge and, say, the average value for the entire delamination.

- The second difficulty with obtaining the *critical* value of the energy release rate with moiré interferometry, is that a subjective decision must be made on the part of the experimenter to determine when to record the fringe pattern. The information available through moiré interferometry merely represents the deformation of the specimen at the time of data acquisition. It is the responsibility of the experimenter to determine whether or not the data is to be associated with a *critical* condition, such as the point of incipient fracture. This judgement will inevitably be based on qualitative information. Since fracture is an inherently precarious situation, the exact moment of criticality is difficult to capture in practice, and the tendency is for the experimental determination of critical fracture parameters to underestimate the true value.

These difficulties are a result of the generality of the moiré technique, and are not altogether discouraging. The moiré method makes no assumptions regarding the behaviour of the specimen, thus arbitrarily complex specimen geometries subjected to general loading conditions can be analysed with confidence, and without the need to make any simplifying assumptions.

Chapter 2

Delamination Fracture Testing of Composite Materials

This chapter provides a summary of the available literature on the experimental characterization of delamination in composites. The discussion is broken down into three sections for convenience, according to opening (mode I), sliding (mode II) and mixed mode fracture behaviour. A final section is included to summarize various practical aspects of delamination testing.

2.1 Mode I

Mode I delamination is characterized by loading normal to the delaminating plane. The most convenient test to simulate such a fracture is the double cantilever beam (DCB) specimen. It is a relatively simple test, dating back at least to the early 1930's [16], and is capable of providing quite accurate results for a wide range of materials. The data reduction procedures for the DCB test range from the simple 'areas method', to more complex methods that provide empirical correction factors to simple beam theory results.

2.1.1 Nonlinear Beam Theory

In one of the earlier investigations of delamination testing of composites, Devitt, Schapery and Bradley [17] developed a method for reducing mode I delamination fracture toughness data for both elastic and viscoelastic double cantilever beam

specimens. They began by considering the elastic strain energy stored in a beam under the action of a bending load for small deflections given by

$$U_e = \frac{1}{2}\int_0^L \frac{M^2}{EI}\,ds,$$

where E is the axial modulus of elasticity, I is the area moment of inertia and M is the local bending moment of a beam of length L and width b, as illustrated in Figure 2.1. As a result of the low flexural rigidity of the thin composite specimens

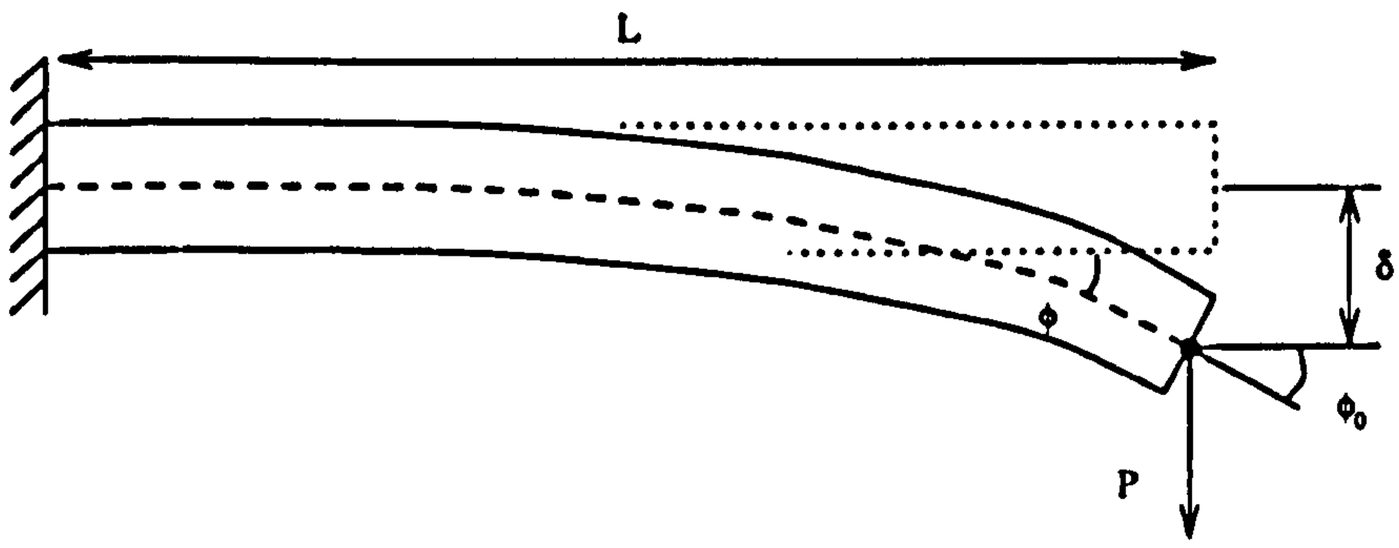


Figure 2.1: Notation for nonlinear beam theory analysis.

and long crack length, large deflections and rotations may be present during the test and the use of linear beam theory was considered inappropriate.

They therefore provided a correction to linear beam theory which was based on work done by Bisshopp and Drucker [18], in which arbitrarily large rotations and bending deflections were accounted for in the analysis of a cantilever beam. The principal nondimensional results for the load and deflection of the cantilever beam are

$$\sqrt{\frac{PL^2}{EI}} = \frac{1}{\sqrt{2}}\int_0^{\phi_0} \frac{d\phi}{\sqrt{\sin\phi_0 - \sin\phi}}$$

and

$$\frac{\delta}{L} = \frac{1}{\sqrt{2}}\sqrt{\frac{EI}{PL^2}}\int_0^{\phi_0} \frac{\sin\phi\,d\phi}{\sqrt{\sin\phi_0 - \sin\phi}}$$

where ϕ is the continuously varying angle between the centroidal axis of the beam and the horizontal, ϕ_0 is the angle of the tangent at the loaded end, and P is the load applied at the end of the beam. After an appropriate change of variables and transformations, the above equations were rearranged to yield

$$\sqrt{\frac{PL^2}{EI}} = F(k) - F(k, \theta_1)$$

ϕ_0	δ/L	PL^2/EI	$GbL^2/2EI$	$Gb/2P$
	$\frac{\delta}{L}$	$3\frac{\delta}{L}$	$\frac{9}{2}\left(\frac{\delta}{L}\right)^2$	$\frac{3}{2}\left(\frac{\delta}{L}\right)$
10	0.1160	0.3530	0.0613	0.1737
20	0.2302	0.7306	0.2499	0.3421
30	0.3406	1.1626	0.5814	0.5000
40	0.4455	1.6923	1.0880	0.6429
50	0.5437	2.3922	1.8330	0.7662
60	0.6340	3.4054	2.9502	0.8663
70	0.7167	5.0812	4.7758	0.9399
80	0.7948	8.6787	8.5558	0.9858
86	0.8472	14.9058	14.9293	1.0002
88	0.8711	20.7282	20.9704	1.0117
90	1.0000	∞	—	—

Table 2.1: Nonlinear beam variables (after [17]).

and

$$\frac{\delta}{L} = 1 - 2\left[\frac{E(k) - E(k,\theta_1)}{F(k) - F(k,\theta_1)}\right]$$

where $F(k)$ and $E(k)$ are the complete elliptic integrals of the first and second kind, respectively, and $F(k,\theta_1)$ and $E(k,\theta_1)$ are the corresponding incomplete integrals. Since the elliptic parameters k and θ_1 are related to ϕ_0 via the expressions

$$k \equiv \frac{1}{\sqrt{2}}\sqrt{1+\sin\phi_0}, \qquad \theta_1 \equiv \sin^{-1}\left(\frac{1}{\sqrt{2}k}\right),$$

the above results for $\sqrt{PL^2/EI}$ and δ/L were expressed as functions of only one parameter, ϕ_0. This implicit relationship was presented in tabular form, Table 2.1, illustrating the numerical relationships between the various nonlinear beam variables and also as a nondimensional load-deflection curve, Figure 2.2. For comparison, the load deflection result from linear beam theory are also included in Figure 2.2, given by

$$\frac{PL^2}{EI} = 3\frac{\delta}{L}.$$

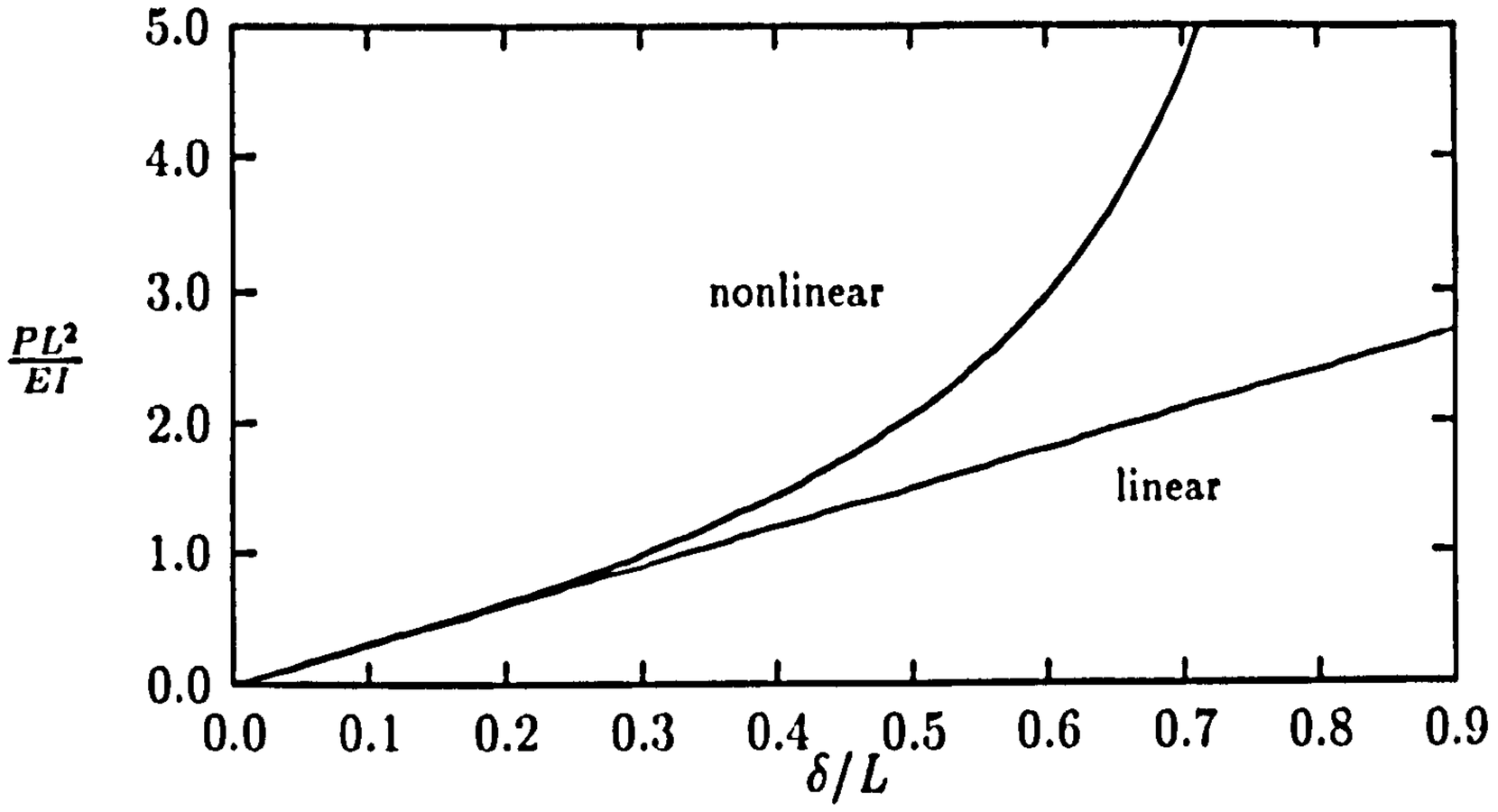


Figure 2.2: Nondimensional load-deflection curves for a cantilever beam (after [17]).

The strain energy release rate for the nonlinear beam was arrived at by introducing $f(\delta/L)$, such that

$$f(\delta/L) = \frac{U_e L}{2EI} = 2[F(k) - F(k, \theta_1)]^2.$$

Following the definition of the energy release rate

$$G = -\frac{1}{b}\frac{dU_e}{da}$$

and performing the differentiation with respect to crack length (in this case, L), the following nondimensional result was obtained for G, where W is the strain energy due to bending of the beam:

$$\frac{GbL^2}{2EI} = \frac{WL}{2EI}(1 + S),$$

where S is introduced for convenience, and equals 2 for linear beam theory. Since WL/EI and S are both functions of δ/L, $GbL^2/2EI$ is an implicit function of δ/L, and is also tabulated in Table 2.1 as column three. In Figure 2.3, the linear

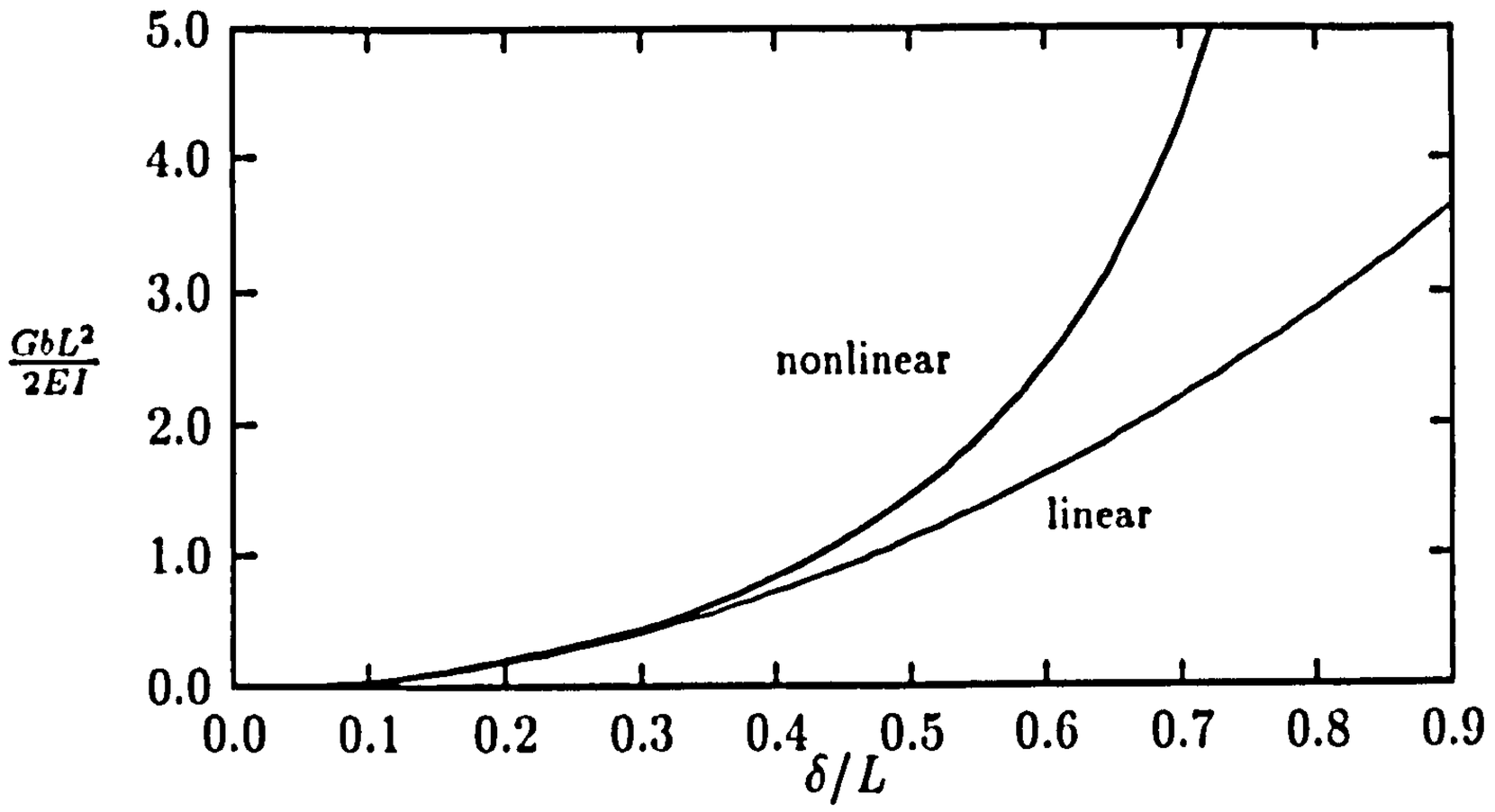


Figure 2.3: Nondimensional energy release rates for a cantilever beam (after [17]).

beam theory result for the energy release rate

$$\frac{GbL^2}{2EI} = \frac{9}{2}\left(\frac{\delta}{L}\right)^2$$

and the nonlinear result are shown graphically for comparison. The last column in Table 2.1, $Gb/2P$ was obtained by dividing the fourth column by the third, and allows for G to be determined without knowledge of the flexural rigidity of the beam EI.

During an experiment, the authors measured the instantaneous load P, grip displacement δ and crack length $2L$. After determining δ/L, Lagrangian interpolation was used to calculate both PL^2/EI and $Gb/2P$, and subsequently, G. They performed their fracture experiments under quasi-static crack growth—that is, they maintained a slow and steady rate of delamination progression. For any particular pair of δ-L measurements, the corresponding energy release rate would represent a critical value, or G_c.

Devitt *et al.* used this approach to reduce experimental data obtained from a range of unidirectional DCB experiments and found reasonable agreement. However, in their discussion, the authors clarified that several simplifying assumptions were implicit in their analysis. The most notable of these was that beam theory

presupposes a 'built-in' beam, implying that there is no rotation of the cross-section of the beams at the crack-tip.

2.1.2 Width-tapered DCB Specimens

Another early study of opening mode delamination in composite materials was performed by Bascom, Bitner, Moulton and Siebert [19]. In their experimental approach, width-tapered double cantilever beam specimens (WTDCB) were used to asses the delamination toughness of organic-matrix woven reinforcement composites. Their main interest as material developers was to assess the effectiveness of the addition of various matrix toughening agents on the observed delamination behaviour.

WTDCB specimens were not new, having been widely used in determining adhesive bond strengths. This specimen type provides a convenient and stable testing configuration which facilitates the reduction of the experimental load-deflection data. Recall the general expression for relating the energy release rate to the change in compliance

$$G = \frac{P^2}{2b}\frac{dC}{da}.$$

For an ideal elastic beam,

$$\frac{dC}{da} = \frac{24}{Eh^3}\frac{a^2}{b}$$

so that

$$G_I = \frac{12P^2}{Eh^3}\frac{a^2}{b^2}. \tag{2.1}$$

If the specimen width is tapered for a constant ratio of a/b, G_{Ic} can be determined from P_c and E alone, and is independent of a.

A correction to simple beam theory was included in their analysis to correct for the rotational displacement at the beam root. The expression for dC/da was modified to account for rotational displacement by including a constant crack length increment, βh, where h is the thickness of the specimen

$$\frac{dC}{da} = \frac{24}{Eh^3}\frac{(a+\beta h)^2}{b}.$$

In order to preserve the crack length independence of Equation 2.1, the loading

holes were moved an additional distance X_0 closer to the crack tip such that

$$b = (a + X_0)\frac{b}{a}$$

where X_0 is made to equal βh so that Equation 2.1 remained unchanged. β was determined to be 0.37 from previous experimental compliance measurements, and so in the study performed by Bascom *et al.*, the loading holes were moved a distance of $0.37h$ to correct for the rotational displacements.

The results of Bascom *et al.* clearly demonstrated the advantages of modifying the epoxy matrix with elastomeric additives in the form of a dispersed phase of soft inclusions in the matrix. In a discussion of their results and fracture surfaces, Bascom *et al.* concluded that the interlaminar fracture process involved considerable plastic deformation and yielding of the matrix resin. They argued that the size and shape of this process zone played a significant role in determining the nature and extent of toughening provided by elastomeric additives.

Another issue with which they were concerned was the discrepancy between fracture toughness data for laminates and 'neat' matrix material properties. The energy release rates obtained from their WTDCB specimens exceeded the expected value given for the matrix material alone. The larger fracture toughness of laminates was attributed to the restraint of the fibre layers on the crack tip deformation, but it was noted that this restraint would not be as great as encountered in thin bond lines between metallic adherands. Thus, they concluded that more effort should be focussed on understanding the size and scale of the process zone, and how it effects the delamination of laminates.

2.1.3 Experimental Compliance Method

Experimental compliance methods are based on performing a regression analysis of the measured load-displacement data using a general model of the compliance derived from conventional beam theory. The first analysis to utilize such an approach was that of Wilkins *et al.* [20].

Their analysis of mode I DCB specimens was based on a linear regression procedure designed to exploit several facts from conventional beam theory. First,

they noticed that the compliance of the specimen

$$C = \delta/P = \frac{2a^3}{3EI} = A_1 a^3$$

is a function of the cube of the crack length, so a log-log plot of the compliance versus crack length should yield a linear slope of 3. A least squares fit routine was used to calculate the intercept that gave the best fit of the data with a slope of 3, which provided A_1. Also, the beam theory expression for the critical load for a specimen of width w

$$P_c = \frac{\sqrt{G_c w E I}}{a} = \frac{A_2}{a}$$

is inversely proportional to the crack length, so a similar log-log least squares regression of critical load versus crack length (assuming a slope of -1) yielded A_2. The least squares parameters A_1 and A_2 were used to compute an average value of G_c for each specimen using

$$G_c = \frac{3A_1 A_2^2}{2w}.$$

Also in 1982, Whitney, Browning, and Hoogsteden [21] presented another study of the various data reduction methods for the DCB specimen. Their investigation included the area method, simple beam theory, and a generalized empirical method, known as Berry's method [16].

The energy release rate from conventional beam theory for the dcb specimen can be written

$$G_I = \frac{3P\delta}{2ba}.$$

Berry introduced an empirical generalization of the above relationship, that provided a correction for large displacement and beam root rotation. In particular, the load point displacement for a typical DCB test is written

$$\delta = RPa^n,$$

where R and n are constants determined experimentally from the relationship

$$\log(P/\delta) = -\log R - n \log a.$$

A least squares fit to the above equation for a series of loading and unloading curves allows R and n to be determined. The conventional beam theory result is recovered if $n = 3$. Given a measured P_c and δ_c, the critical energy release rate is determined from

$$G_{Ic} = \frac{nP_c\delta_c}{2ba}. \tag{2.2}$$

Whitney*et al.* also considered the inclusion of shear deformation effects in the analysis of DCB data. Their general result derived from Timeshenko beam theory [22] was

$$G_{Ic} = \frac{nP_c\delta_c}{2ba}\left(\frac{1+S/3}{1+S}\right),$$

where the shear deflection is given by the relationship involving the effective shear modulus, G_{12}, and the effective bending modulus, E_{11}^b

$$S = \frac{3E_{11}^b h^2}{32G_{12}a^2}.$$

It was noted that the precise experimental determination of G_{12} and E_{11}^b is not always an easy task. However, for the materials, specimen geometries and crack sizes tested in their experimental programme, shear deformation corrections were estimated to be on the order of 3%, and thus could be safely neglected.

The experimental results of Whitney *et al.* indicated that the DCB test method was capable of accurately measuring mode I fracture resistance in composite laminates, and that all of the data reduction methods under investigation were able to discriminate between materials of significantly different interlaminar toughness. They suggested that further work should focus on DCB testing of multidirectional laminates, and that the results should be compared to those obtained with the free-edge delamination specimen.

2.1.4 The J-integral Approach

The advantages of an energy approach in the study of a delamination fracture in composites is evident in the very use of the energy release rate, G. The

interpretation of G as the energy released during the delamination event, however, is only valid when the material behaviour is linear elastic. Keary *et al.* [23] recognized this limitation, and investigated the application of the J-integral to the analysis of DCB data.

In elastic-plastic fracture mechanics, the J-integral was introduced as a parameter to account for the energy dissipated in the nonlinear deformation of metals. Although J is defined via an integral, standard methods have evolved for computing J from load-displacement data. In the case of linear elastic fracture, J is identical to G.

Keary *et al.* used the conventional J-testing method for reducing load displacement for several DCB experiments. They compared these results to similar values computed from the area method, linear beam theory and the compliance calibration method. All of the data reduction methods produced comparable results, although the authors noted that the linear beam theory and compliance method led to overestimates of the energy release rate.

These overestimates were attributed to several nonlinear deformation mechanisms observed in their specimens, such as permanent deformation in the arms of the DCB specimens and hackled fracture surface areas, indicating significant plastic deformation of the matrix material during the delamination process.

2.1.5 Corrected Beam Theory

In all DCB tests, some means must be provided to transfer load to the specimen. This can be accomplished by using hinges or by using end blocks bonded to the specimen. In the case of end blocks, a correction to the simple beam theory analysis was presented by Williams [24]. In the same work, Williams also presented an analysis accounting for large displacement effects. Williams further developed his corrections to conventional beam theory in [25] and [26]. The three main correction factors proposed by Williams comprise what is known as corrected beam theory and take the following form:

- a large displacement/end block effect parameter (F)
- a correction for shear deformation/beam root rotation (χh), and a
- correction for stiffening due to the end blocks (N).

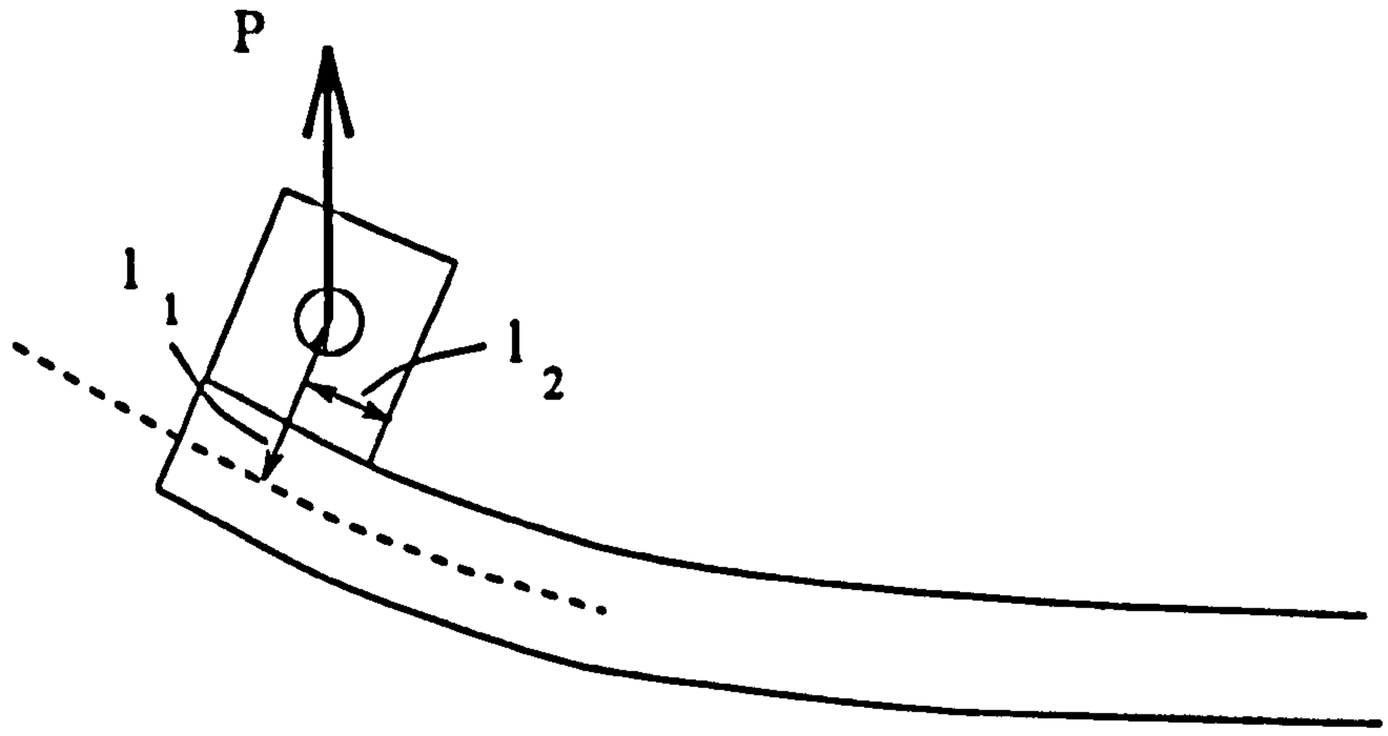


Figure 2.4: Notation used for the end block correction.

The large displacement/end block analysis performed by Williams was similar to that of Devitt et al. [17] discussed above. The resulting correction factor for a symmetric DCB specimen due to large displacements and end block effects is given by

$$\frac{G}{G_0} = F = 1 - \frac{3}{10}\left(\frac{\delta}{a}\right)^2 - \frac{3}{2}\left(\frac{\delta}{a}\frac{l_1}{a}\right),$$

where G_0 is the uncorrected linear beam theory result, and l_1 is as shown in Figure 2.4.

Because the shear modulus in laminates is lower than the axial modulus, shear deformation of the arms of the DCB specimen require yet another correction factor. This effect is particularly marked in the built-in assumption used in the analysis of the cantilever beams. After a lengthy analysis based on the elastic foundation model of Kanninen [27], Williams concluded that both shear deformation and root rotation in DCB specimens could be corrected for by adding a length χh to a where

$$\backslash = \sqrt{\frac{E_{11}}{E_{11}G_{12}}}\left[3 - 2\left(\frac{\Gamma}{1+\Gamma}\right)^2\right] \tag{2.3}$$

and

$$\Gamma = 1.18\sqrt{\frac{E_{11}E_{22}}{G_{12}}}.$$

For unidirectional laminates, $E_{11} \gg E_{22}$ and G_{12} so that $\Gamma \gg 1$ and

$$\chi = 0.3\sqrt{\frac{E_{11}}{G_{12}} + 3.4\sqrt{\frac{E_{11}}{E_{22}}}}.$$

For isotropic materials $\chi \approx 0.7$, while for carbon-fibre/epoxy laminates typical values are $E_{11} = 130\,\text{GPa}$, $E_{22} = 8\,\text{GPa}$ and $G_{12} = 3.5\,\text{GPa}$, giving $\chi = 2.4$; a much larger effect. This correction to a is intended to be used in all of the preceding equations involving the crack length—thus $(a + \chi h)$ should replace a in the expression used for G_0 and F.

It is necessary to draw attention to the similarity between the χh correction factor and the experimental compliance method originally proposed by Berry [16], discussed above. Recall in the experimental compliance method that the load point displacement and crack length are related through $\delta = RPa^n$, and the parameters R and n are determined by curve fitting to compliance data. The best fit value of n to the $(a + \chi h)$ form is given when

$$n = \frac{3}{1 + (\chi h/a)}$$

and typically, since $\chi h \approx 5\,\text{mm}$ for $h = 2\,\text{mm}$ and $a \approx 50\text{–}100\,\text{mm}$, n will be about 2.7–2.8, as is observed.

The χh correction factor can also be directly determined from the experimental load-deflection data, in much the same way as the experimental compliance calibration factor, n, of section 2.1.3. This is accomplished by plotting the cube root of the compliance versus a, where χh is interpreted as the x-axis intercept. Furthermore, E_{11} will be recovered as the slope of this plot.

Williams argues that the χh correction is a better representation than the experimental compliance method since it models the actual mechanism involved in the deformation of the DCB. In practice, both the experimental compliance method, and the χh method yield effectively the same results.

The final correction factor considered by Williams corrects for the additional stiffening due to the bonding of end blocks to the specimen. The result, again using the notation of Figure 2.4, is

$$N = 1 - \left(\frac{l_2}{a}\right)^2 - \frac{9}{18}\frac{\delta l_1}{a^2}\left[1 - \left(\frac{l_2}{a}\right)^2\right] - \frac{9}{35}\left(\frac{\delta}{a}\right)^2.$$

Thus, including the above three correction factors, the corrected beam theory energy release rate for a DCB specimen is given by

$$G_I = \frac{3}{2}\frac{P\delta}{b(a+\chi h)}\left(\frac{F}{N}\right). \tag{2.4}$$

In practice, an empirical approach is generally adopted, whereby the correction factor χh is obtained empirically from the experimental data. By plotting the cube root of the compliance versus the crack length, a correction factor Δ can be extracted as the intercept of the curve with the x-axis. Equation 2.4 is generally re-written as

$$G_I = \frac{3}{2}\frac{P\delta}{b(a+\Delta)} \tag{2.5}$$

so that G_{Ic} can be obtained without actually computing χh. In fact, the relationship between Δ and χh can be used to independently determine an estimate of E_{11}. This relationship is given by

$$E_{11} = \frac{8(a+\Delta)^3}{Cbh^3}.$$

2.2 Mode II

The mode II or 'sliding' failure mode is characterized by a state of pure shear at the crack tip. In general, G_{IIc} is significantly larger than G_{Ic} for composite laminates. Mode II tests are more difficult to perform and the results are more prone to scatter than the DCB test.

2.2.1 Short-Beam and Four-Point Shear Tests

One of the earliest tests adapted from isotropic material testing for the purpose of characterizing composite delamination behaviour was the short-beam shear (SBS) test. The SBS specimen involves only a limited amount of material in its fabrication, and leads to a relatively simple experiment to determine G_{IIc} from the measured applied load P.

As discussed by Browning, Abrams and Whitney [28], however, the SBS test suffers from a significant limitation that seriously impairs its ability to yield meaningful results. Since the thickness of composite SBS specimens tends to

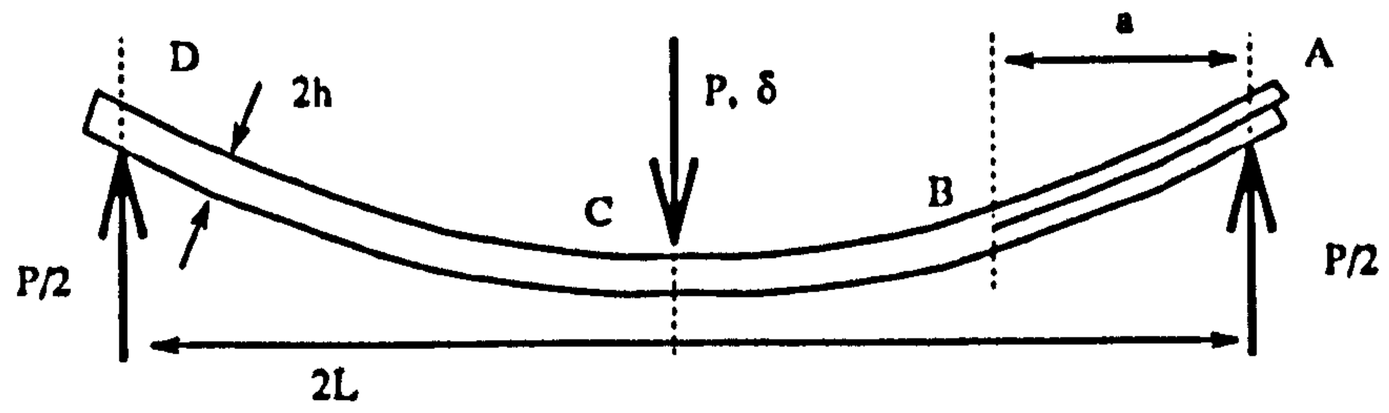


Figure 2.5: The end notched flexure specimen.

be small (< 16 plies), the thin specimens are completely dominated by stress concentrations induced by the load pins. This destroys the assumed parabolic shear stress distribution used to calculate the apparent shear stress at failure. One straightforward method to improve the SBS specimen when testing composites is to use thicker specimens (> 50 plies), but this results in expensive test specimens.

Instead, Browning *et al.* suggested the use of a four-point shear (FPS) test. This type of test would allow for the use of a longer, as opposed to a thicker specimen. Furthermore, the shear load in the FPS specimen is twice that of the SBS specimen, so the desired shear failure mode is more readily obtained in the FPS test.

Browing *et al.* performed an extensive investigation of shear failure in carbon-fibre/epoxy using a conventional SBS specimen (16-ply), a thick SBS specimen (50-ply), and two FPS specimens (16-ply and 24-ply). A careful investigation of the post mortem fracture surfaces indicated the characteristic mode II hackle patterns for all of the specimens except the thin SBS specimens. These thin SBS specimens also grossly overestimated the G_{IIc} values obtained using the other specimens.

2.2.2 End Notched Flexure Tests

The end notched flexure (ENF) fracture specimen is essentially a three point flexure specimen with an embedded delamination at one end, shown schematically in Figure 2.5. For a symmetric or asymmetric delamination, the ENF specimen produces a condition of pure mode II loading. This specimen appears to have been first used by Russell and Street [29].

The analysis offered by Russell and Street was very straightforward—they simply derived an expression for the compliance of the ENF specimen from con-

ventional beam theory obtaining

$$G = \frac{9P^2a^2C}{2b(2L^3 + 3a^3)}.$$

The effect of friction was estimated from the extent of hysteresis observed during the loading and unloading of the specimen and found to cause an overestimate in G of about 2%.

A more thorough study of the ENF specimen was performed by Carlsson, Gillespie and Pipes [30]. These authors were specifically concerned with three aspects of the ENF specimen:

- the effects of interlaminar shear deformation,
- minimizing geometric nonlinearity and avoiding nonlinear material behaviour, and,
- assessing the amount of energy dissipated through frictional sliding along the initially delaminated portions of the beam.

They introduced a modified expression for the energy release rate to include the effects of shear deformation and an analysis on sizing the ENF specimen to obtain crack growth within the linear elastic regime. A detailed consideration of the influence of friction was also presented.

Their derivation of the energy release rate for the ENF specimen was based on the change of compliance with crack extension. The compliance of the ENF specimen is defined as the displacement, δ, at the central loading pin (see Figure 2.5) divided by the applied load measured at the same point. Following the notation in Figure 2.6, δ can be calculated from

$$\delta = \frac{\Delta_{AB} + \Delta_{BC} + \Delta_{CD}}{2}.$$

The beams BC and CD were modelled as cantilever beams with elastic modulus E_{11}, shear modulus G_{12} and thickness $2h$. Using Timeshenko beam theory, Δ_{BC} and Δ_{CD} can be written

$$\Delta_{CD} = \frac{PL^3}{4E_{11}bh^3} + 0.3\left(\frac{PL}{G_{12}bh}\right)$$

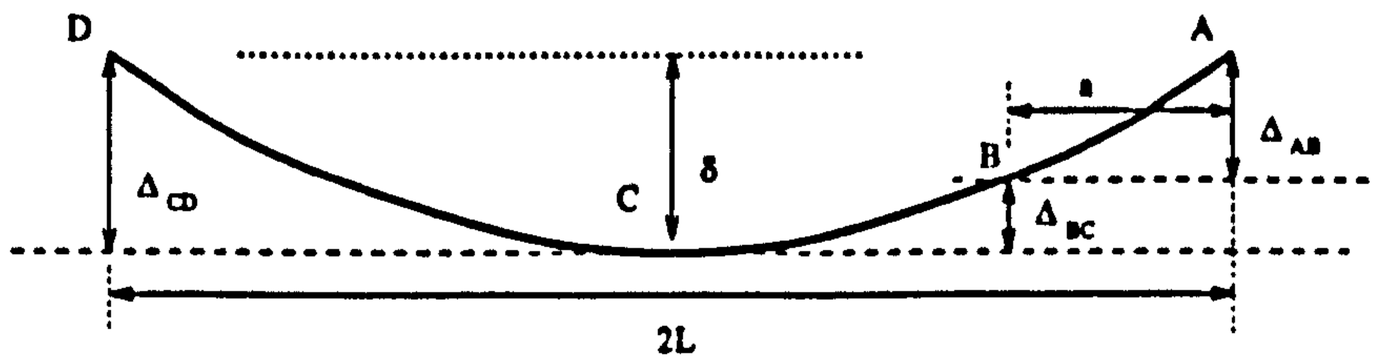


Figure 2.6: Definition of vertical displacements for the ENF specimen.

and

$$\Delta_{BC} = \frac{P(2L^3 + 3aL^2 + a^3)}{8E_{11}bh^3} + 0.3\left(\frac{P(L-a)}{G_{12}bh}\right).$$

The displacement Δ_{AB} has two components—one due to bending and shearing deformations of the beam and the other due to the rotation of the cross section at point B in Figure 2.6.

Assuming that the ends of the beam are allowed to rotate, and that each of the split components of the beam carries the same load, $P/4$, the displacement component due to bending and shear can be calculated from

$$\Delta_{AB}^{\text{deformation}} = \frac{Pa^3}{E_{11}bh^3}\left(1 + \frac{3h^2E_{11}}{8a^2G_{12}}\right).$$

The displacement due to the rotation of the cross section can be approximated by the slope at B times the length of the delaminated region

$$\Delta_{AB}^{\text{rotation}} \approx \frac{3P}{8E_{11}bh^3}\left(aL^2 - a^3 + \frac{ah^2E_{11}}{G_{12}}\right).$$

Adding all of the contributions to the displacement and substituting into the expression for the compliance yields

$$C = \frac{2L^3 + 3a^3}{8E_{11}bh^3}\left(1 + \frac{2(1.2L + 0.9a)h^2E_{11}}{(2L^3 + 3a^3)G_{12}}\right).$$

Using the definition of the energy release rate,

$$G_{II} = \frac{9a^2P^2}{16E_{11}b^2h^3}\left[1 + 0.2\frac{E_{11}}{G_{12}}\left(\frac{h}{a}\right)^2\right].$$

When the quantity $(E_{11}/G_{12})(h/a)^2$ is small, the above equation reduces to the conventional beam theory solution. The error introduced by not including the

shear deformation effects was shown to vary between 5–18%, being worse for specimens with large thickness-to-span ratios.

The geometric nonlinear behaviour of the ENF was evaluated by noting that according to linear beam theory, the curvature of a beam

$$\frac{1}{R} = \frac{d^2y/dx^2}{(1 + (dy/dx)^2)^{3/2}}$$

is generally approximated with

$$\frac{1}{R} \approx \frac{d^2y}{dx^2}$$

since the square of the slope $(dy/dx)^2$ is assumed to be much less than unity. To obtain an upper bound on the allowable displacement, δ_{max}, in the ENF specimen, Carlsson *et al.* used the following expression for the maximum slope, (y'_{max}), at the central loading pin

$$(y'_{max}) = \frac{3(L^2 + 3a^2)(\delta_{max})}{2L^3 + 3a^3}$$

to obtain the relationship between (y'_{max}) and (δ_{max}). Writing the critical displacement as

$$\delta_c = \frac{(2L^3 + 3a^3)}{6a}\sqrt{\frac{G_{IIc}}{E_{11}h^3}}$$

and combining the above two equations in light of the constraint $\delta_c \leq \delta_{max}$ yields the following criteria that must be met by the thickness of the specimen

$$h \geq \sqrt{\frac{G_{IIc}(L^2 + 3a^2)^2}{4(y'_{max})^2 a^2 E_{11}}}.$$

The authors also showed that material nonlinearities and flexural failure can be avoided by proper sizing of the ENF specimen. The maximum bending strain, ϵ, at the center of the beam (disregarding the local stress singularity at the crack tip) can be written

$$\epsilon = \frac{6Lh\delta}{2L^3 + 3a^3}.$$

By a similar reasoning as above, a thickness requirement can be formulated from the condition $\delta_c \leq \delta_{max}$ where δ_{max} is the maximum allowable displacement related to the maximum allowable strain (ϵ_{max}) to maintain linear elastic behaviour.

Their calculations yielded

$$h \geq \frac{L^2 G_{IIc}}{a^2(\varepsilon_{max})^2 E_{11}}.$$

An approximate expression for the energy dissipated due to friction in the ENF specimen was arrived at by Carlsson *et al.* by integrating a distributed normal force, $N(x)$, acting over a contact region of $\pm 2h$ about the center of the loading pin between the two cracked portions of the beam at point A. During crack growth, the relative sliding of the cracked beams, $\Delta u(x)$ results in work done, W_f, proportional to the coefficient of sliding friction, μ, given by

$$W_f = \int_{-2h}^{2h} \mu N(x) \Delta u(x)\, dx.$$

After some further simplifications, the final approximate expression for the quantity of work due to friction was determined to be

$$W_f < \frac{3P^2\mu}{8E_{11}bh^3}[a^2 + h^2(E_{11}/g_{12})/12].$$

Evaluation of the above expression for a range of μ values, material properties and loading conditions, revealed, as Russel and Street found, that the energy dissipated through friction in the ENF specimen was < 2–4%.

Following their initial work, Carlsson, Gillespie and Tretheway [31] conducted a further investigation into the ENF test specimen as a means to provide G_{IIc} values from composite laminates. Specifically, they were concerned with the influence of precracking, the sensitivity of the data reduction scheme and the observed nonlinear load-deflection curves.

The nonlinear load-deflection curves were not expected, since the specimen dimensions had been to chosen so that they would remain well within the linear elastic regime. Closer inspection of the fracture surfaces of the ENF specimens revealed that the nonlinearities were most likely arising from a combination of slow stable crack growth and microcrack coalescence preceding unstable fracture. These effects were found to be highly rate dependent—higher rates of loading producing smaller damage zones (more brittle material response) resulting in lower G_{IIc} values, and lower loading rates allowing for larger damage zone development and correspondingly larger G_{IIc} values.

The energy release rates determined from three types of precracks were also

presented. Specimens with a mode I precrack (produced by arresting a delamination driven by the introduction of a wedge at one end of the specimen) resulted in slightly lower values for G_{IIc} than specimens precracked under a stable mode II arrangement. However, mode I precracking provided the advantage of a more precise determination of the initial crack length, since a post mortem fracture surface analysis would clearly distinguish the transition from mode I to mode II crack growth. Specimens were also tested from the blunt Teflon insert, resulting in grossly overestimated G_{IIc} values.

A more robust data reduction was also introduced and verified. The technique is very similar to Berry's experimental compliance method discussed above for the DCB specimen. By writing the compliance of the ENF specimen as

$$C/C_0 = 1 + m(a/L)^3$$

where m is a parameter to be determined through a regression analysis of the compliance data, the energy release rate can be written

$$G_{II} = \frac{3ma^2P^2}{2b}. \tag{2.6}$$

Since the ENF is an unstable fracture specimen, it is necessary to perform an experimental compliance procedure before the actual fracture experiment. The parameter m is determined by plotting the normalized compliance versus crack length cubed data, as illustrated in Figure 2.7.

As with all test methods, empirical correction factors have been developed to further improve the reliability of the data reduction process. Following Williams [32, 26], the results of such a consideration are presented. The most significant inadequacy of conventional beam theory is the assumption of the built-in boundary condition at the crack root.

Conventional beam theory expressions for the energy release rate incorrectly assume that there is no rotation or deflection of the cross section of the beam at the built-in end (at the crack tip), but in reality there is both some rotation and deflection. This effect can be modelled by adding a length χh to the real crack length, where χ is a constant given by the elastic properties of the material. This is the same correction discussed above for the DCB test, and Equation 2.3

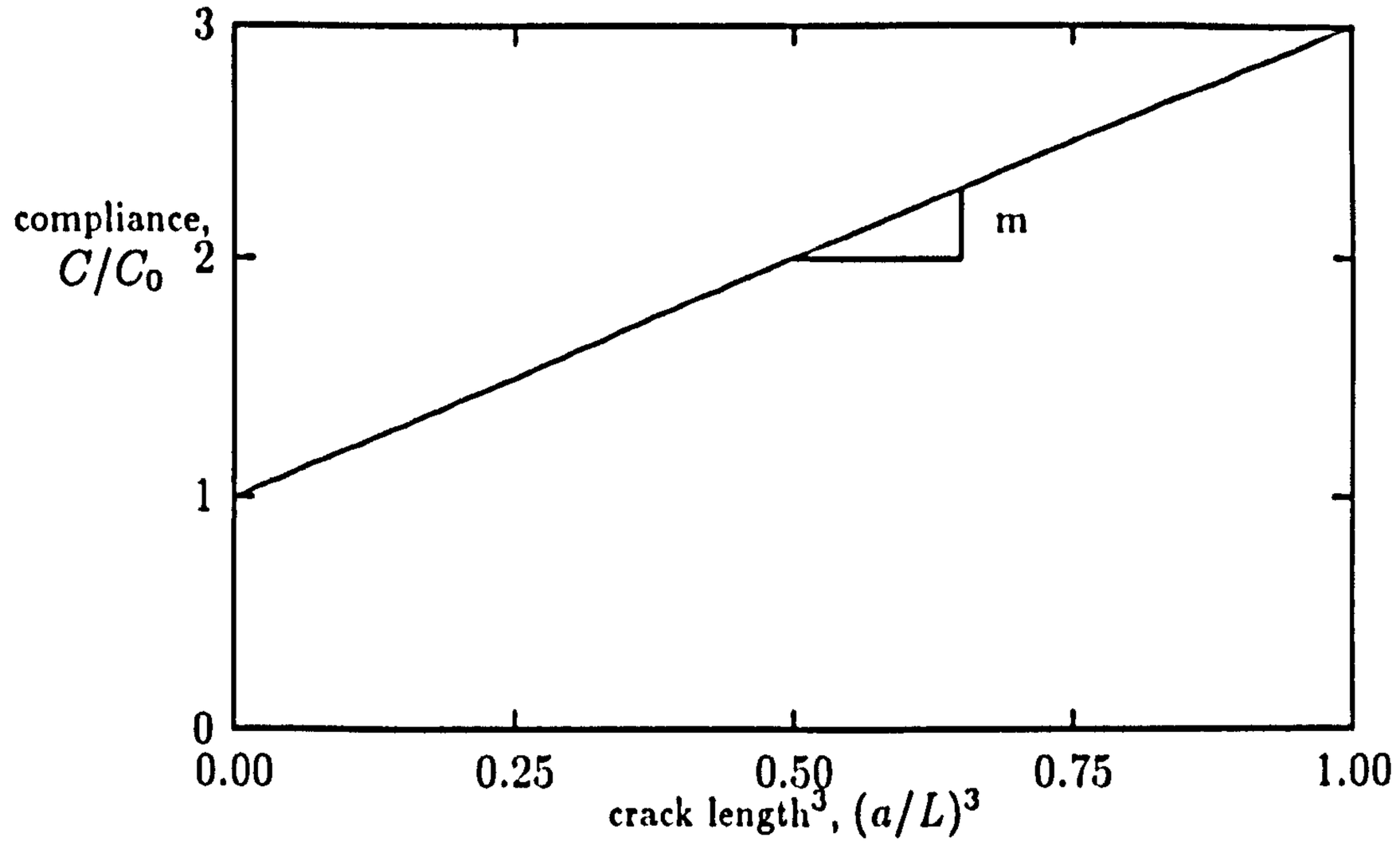


Figure 2.7: Normalized compliance versus normalized crack length for determination of m.

can also be used to compute χ for the ENF specimen. However, noting again that the expression for the compliance will also include this correction, it is more reasonable to compute χh from a regression analysis of experimental data and, then

$$G_0 = \frac{9P^2(a+\chi h)^2}{16b^2h^3E_{11}}$$

where G_0 is the uncorrected beam theory result.

An additional multiplicative factor, F, can then be introduced to correct for large displacement effects, such that $G_{II} = FG_0$, and

$$F = 1 - \Theta\left(\frac{\delta}{L}\right)^2$$

where

$$\Theta = \frac{3[15 + 50(a/L)^2 - 30(a/L)^3 + 63(a/L)^4 - 34(a/L)^5 + 15(a/L)^6]}{5[2 + 3(a/L)^3]^2}$$

2.2.3 End Loaded Split Tests

The end loaded split (ELS) test configuration shown in Figure 2.8 uses a split laminate beam specimen loaded at the cracked end and fixed at the opposite

Figure 2.8: The end loaded split test specimen.

end. The measured values of P_c and δ_c from an ELS test can be used to calculate corresponding values of G_{IIc} using linear beam theory. The ELS has been used extensively for the determination of G_{IIc}, for example by Corleto and Bradley [33] and Yang and Williams [32].

Again, following Williams [32, 26], the results of applying correction factors to the conventional beam theory equations are presented. The conventional beam theory result for the energy release rate in the ELS test is given by

$$G_{II} = \frac{9P^2a^2}{4b^2E_{11}h^3}.$$

The corrected expression involving the χh factor is

$$G^0_{II} = \frac{9P\delta}{2B(a+\chi h)}\left[\frac{(a+\chi h)^3}{3(a+\chi h)^3+(L+2\chi h)^3}\right]$$

The additional factor F can then be introduced as before to correct for large displacement and end block effects, where

$$F = 1 - \frac{15+50(a/L)^2+63(a/L)^4}{[1+3(a/L)^3]^2}\frac{3}{20}\left(\frac{\delta}{L}\right)^2 + \frac{3}{h}\frac{\delta}{L}\frac{l_1}{a}\left[\frac{1+3(a/L)^2}{1+3(a/L)^3}\right].$$

The case of a asymmetric loaded ELS specimen, where the direction of the applied load is reversed was also analyzed, and the resulting expression for F can be found in [26].

2.3 Mixed Mode

Mixed mode fracture tests are inherently more difficult to perform and analyze. Especially in the case of laminated composites, complications arise because of

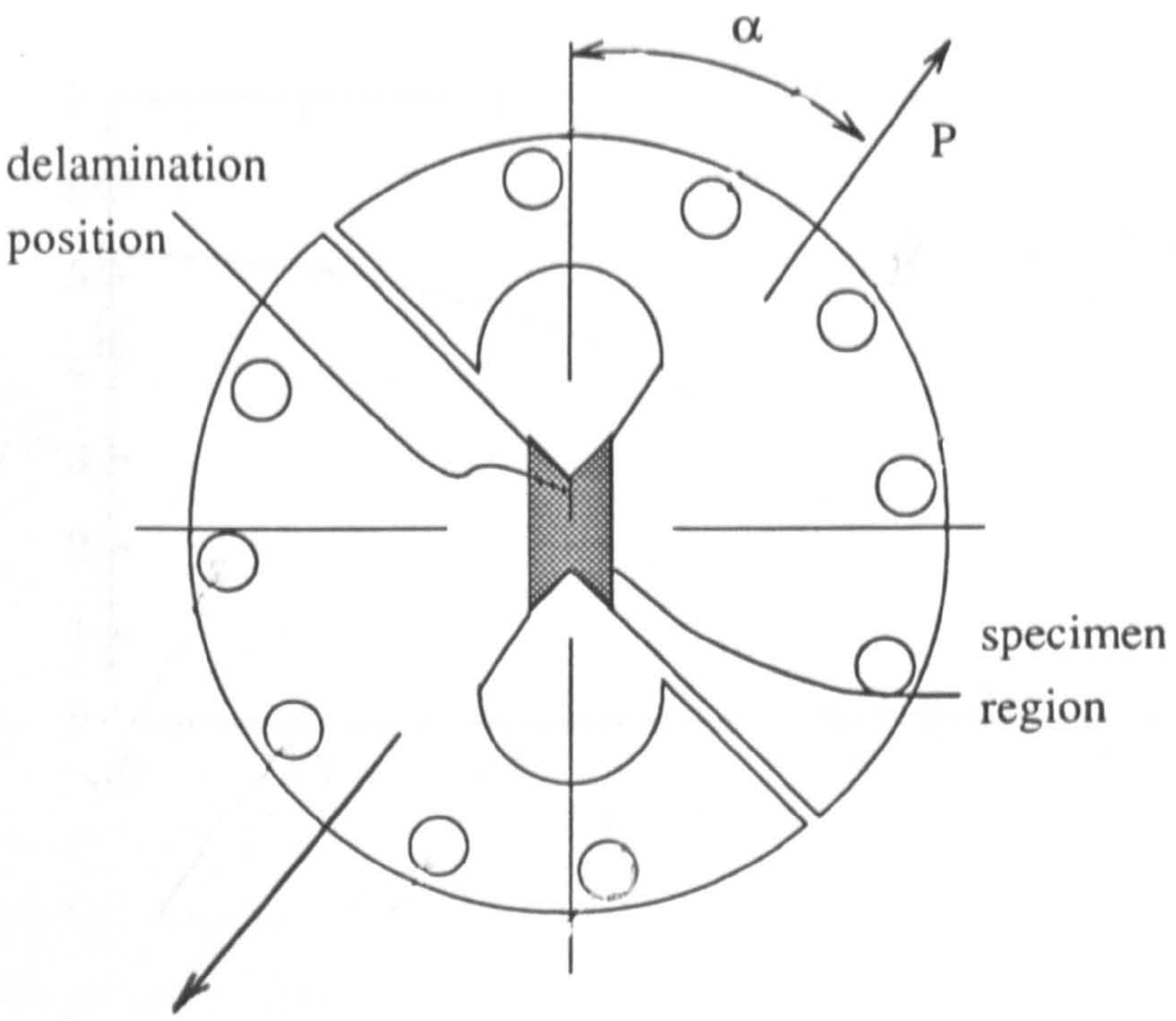


Figure 2.9: The modified Arcan test rig.

the sensitive dependence of fracture behaviour on the type of loading. Further difficulties arise because it is usually necessary to perform several types of different experiments to obtain a range of mixed mode ratios, and this introduces additional uncertainties when comparing experimental results. Another problem involves defining the amount of 'mixity' or the ratio of mode I (opening) to mode II (sliding) during a test, and insuring this ratio remains constant.

2.3.1 The Arcan Test

One of the earlier attempts at measuring mixed mode fracture parameters was performed by Jurf and Pipes [34]. These authors modified a test method introduced by Arcan [35] for producing an arbitrarily loaded single edge notch specimen. The test rig is illustrated in Figure 2.9. The Arcan test requires only a small sample, and is easily performed in a conventional testing machine. Furthermore, the data from the test is easy to reduce, and the same specimen design and test rig can be used over a wide range of mixed mode loadings.

The experimental results presented by Jurf and Pipes for a carbon-fibre/epoxy composite system are reproduced in Figure 2.10, where $\alpha = 0°$ corresponds to

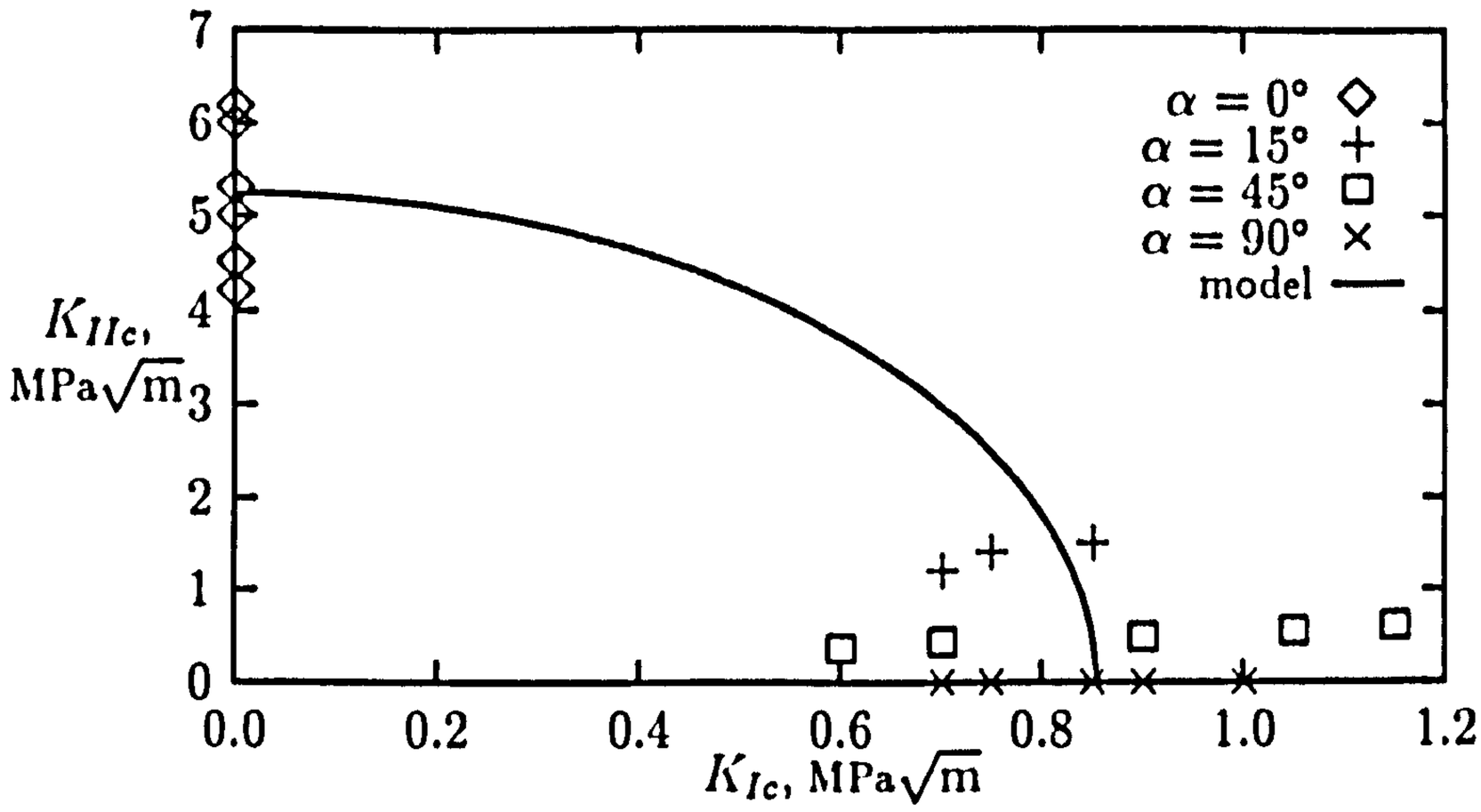


Figure 2.10: Experimental results of Jurf and Pipes.

a pure mode II test, and $\alpha = 90°$ refers to a pure mode I test. The results are shown in terms of the stress intensity factors, K_{Ic} and K_{IIc}, which can be related directly to G_{Ic} and G_{IIc}. When converted, the G_{IIc} were found to be approximately 9 times larger than the pure G_{Ic} results. Considerable scatter plagued all but the pure mode I results. Inspection of the fracture surfaces revealed distinctly different morphologies—matrix cleavage fracture dominated the mode I specimens, while fibre-matrix debonding and matrix fragmentation provided the more tortuous fracture path in the pure mode II specimens.

The authors used their results to determine a failure locus assuming the simple model

$$\left(\frac{K_{Ic}}{877}\right)^2 + \left(\frac{K_{IIc}}{4980}\right)^2 = 1.$$

Two data points from the $\alpha = 15°$ group were excluded in the plot of Figure 2.10 by the authors because they significantly deviated from the general trend of this simplified failure model. In their discussion, they noted that the fracture surfaces of the deviant specimens were more similar to the pure mode II surfaces, whereas for the remaining specimens ($\alpha = 15$–$90°$), the fracture surfaces were more or less identical to each other.

Although the data in Figure 2.10 seems to fit the simple proposed model, there

is no reason to expect that there should exist some smooth function representing a failure locus for delamination. The large scatter in the K_{IIc} values, and the wild deviation with the $\alpha = 15°$ specimens clearly illustrates the sensitivity of the mixed mode loading ratio on the results of fracture tests.

2.3.2 Free Edge Delamination Tests

A more comprehensive study of delamination failure in multi-layered composites was performed by O'Brien [36]. He studied the delamination behaviour in unnotched cross-ply laminates using the free-edge delamination test. Several specimens of $[\pm 30/\pm 30/90/\overline{90}]_s$ were tested under static and fatigue loading conditions. This particular layup was chosen because of its innate tendency to delaminate under a tensile load. Delamination growth and stiffness loss were monitored nondestructively, and the laminate stiffness was found to decrease linearly with delamination size. Post mortem studies revealed significant matrix cracking within the 90° plies, as well as clear evidence of delamination at the −30/90 interface. The delamination would occasionally 'jump' through the midplane of the laminate via the 90° matrix cracks.

Values of the critical energy release rate were determined from the measured critical strain in the $[\pm 30/\pm 30/90/\overline{90}]_s$ composites. The presence of both interlaminar shear and normal stresses at the free edge of the specimen required a numerical treatment to partition the energy release rate into G_I and G_{II} components. The approach used by O'Brien for this and subsequent work was based on the finite element virtual crack closure technique. The critical G_c for the $[\pm 30/\pm 30/90/\overline{90}]_s$ laminate was determined from the experimental data and used to predict the nominal strain at the onset of delamination in a $[+45_n/-45_n/0_n/90_n]_s$ laminate where $n = 1, 2, 3$. Delaminations were anticipated along the 0/90 interfaces, and the experimental measure of critical strain was used to predict the applied load necessary to initiate delaminations along this plane. O'Brien appears to have been the first investigator to successfully apply experimental delamination results in a predictive manner.

In his concluding remarks, O'Brien observed that the preliminary prediction of delamination onset in $[+45_n/-45_n/0_n/90_n]_s$ laminates using a critical G_c determined from tests on $[\pm 30/\pm 30/90/\overline{90}]_s$ laminates indicated that G_c may be

independent of ply orientation that make up the delaminating surfaces. However, a more systematic study of mixed mode effects by O'Brien [37] further clarified the situation.

To explore the role of the two contributions G_{Ic} and G_{IIc} on composite delamination, O'Brien used three nearly identical specimens, designated as A ($[\pm 45/0/90]_s$), B ($0/\pm 45/90]_s$) and C ($[45/0/-45/90]_s$). Since these laminates all contain the same plies, they will all have the same total G. Thus, if G_c governs the onset of delamination, ϵ_c for each specimen should be the same. If G_I governs delamination onset, then ϵ_c will be lowest for the layup with the highest percentage of G_I contribution. A finite element analysis indicated that the contribution of G_I to the total G was 85%, 57% and 28%, for specimens A, C and B, respectively. The experimental results confirmed that the laminates with the larger mode I contributions tended to delaminate at lower critical strains.

A similar result was obtained by Rybicki *et al.* [38]. In their study they employed DCB and cracked lap shear (CLS) tests to evaluate mode I and mixed mode energy release rates in carbon-fibre/epoxy laminates. Using several different specimens they tested specimens with G_{II}/G_I ratios from 0–5. Reduction of their experimental data, including mode partitioning was handled in a similar manner as O'Brien—that is via the virtual crack closure technique. For the cases considered by Rybicki *et al.*, the data suggested that the mode II component of the energy release rate did not appear to have as strong an influence on the crack extension event as the mode I component.

2.3.3 A Closer Look At Mixed Mode Partitioning

As can be seen from the results of O'Brien [37] and Rybicki et al. [38], the individual components sometimes play a more important role in controlling the onset of delamination than the total G. Thus, it is important to provide a means for separating or partitioning the total G values from mixed mode experiments.

In many ways, the global perspective provides the most reasonable way to partition the various components of the energy release rate. Following Williams [39], the general beam theory expression for the energy release rate for a laminate as

illustrated in Figure 2.11 is

$$G = \frac{3}{4b^2 h E_{11}} \left[\left(\frac{2h}{h_1}\right)^3 M_1^2 + \left(\frac{2h}{h_2}\right)^3 M_2^2 - (M_1 + M_2)^2 \right].$$

This result can be partitioned into individual expression for the opening and

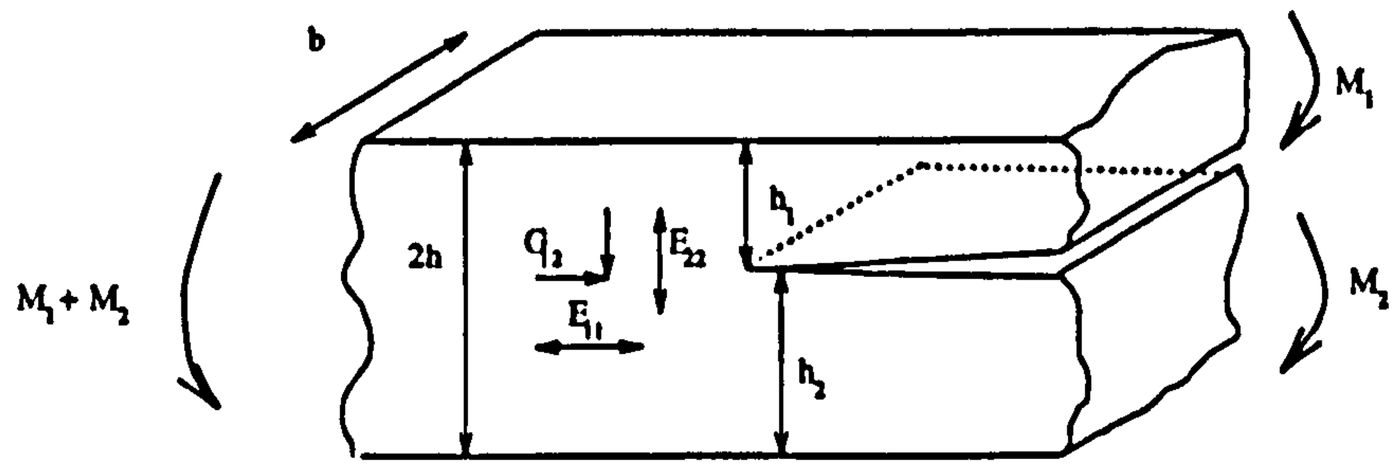


Figure 2.11: A beam under general bending moments.

sliding modes, respectively, as

$$G_I = \frac{6}{b^2 E_{11}} \frac{h_1^3}{h_2^3(h_1^3 + h_2^3)} \left[M_2 - \left(\frac{h_2}{h_1}\right)^3 M_1 \right]^2 \tag{2.7}$$

and

$$G_{II} = \frac{6}{b^2 E_{11}} \frac{3h_1 h_2}{(h_1 + h_2)^2 (h_1^3 + h_2^3)} (M_2 + M_1)^2. \tag{2.8}$$

The above expressions are somewhat simplified by introducing

$$M_I = M_2 - \left(\frac{h_2}{h_1}\right)^3 M_1$$

and

$$M_{II} = M_2 + M_1.$$

Noting that for a symmetric DCB specimen $h_1 = h_2 = h$ and $M_2 = -M_1 = Pa$, we can write Equation 2.7 as

$$G_I = \frac{12P^2a^2}{b^2h^3E_{11}} = \frac{3P\delta}{2ba}.$$

In [40], Charalambides *et al.* discussed and compared two schemes for analyzing mixed mode delamination data. The first method was based upon a consideration of the local singular stress field ahead of the crack tip, and the

second was a global method based upon a consideration of the applied energy release rate. In many instances, these two different schemes give the same predicted mode mix, but in a few critical cases they predict very different degrees of mode mix.

By reviewing the available experimental data for such cases it was concluded that the global scheme resulted in a more consistent interpretation of the data. In several cases the global method clearly gave better agreement between the theoretical predictions and the observed results.

For example, they considered in detail the asymmetric double cantilever beam test. Using the local field analysis, a mixed mode situation is expected, with G_I/G ranging from approximately 63–100% for h_2/h_1 ranging from 0–1. Globally, however, we have

$$M_{II} = M_1 + M_2 = 0$$

so the equations predict only a G_I component. Using the local partitioning method, the authors expected to see a 60 J/m^2 difference between the case of a laminate with $(h_2/h_1) = 1$ and one with $(h_2/h_1) \approx 0.25$. However, with $G_{Ic} =$ 270 J/m^2 and $G_{IIc} = 600$ J/m^2 and a standard deviation of 20 J/m^2, they only measured a difference of 20 J/m^2 between the two cases—a result clearly dominated by the scatter of the data. Thus, the authors concluded that the global partitioning method more accurately reflected reality.

They suggested that the difficulties with the local method were a result of the very localized nature of the singular-dominated region ahead of the crack tip, the relatively large damage zones in the materials studied and the surface roughness interactions that obscured the true singularity. These conclusions were justified by considering measured damage zone sizes from scanning electron microscopy.

Another interesting observation discussed by Charalambides *et al.*, consistent with the findings of O'Brien [37] and Rybicki *et al.* [38], is that mixed mode delamination is predominantly controlled by G_I, with a secondary effect arising from the G_{II} component. They proposed that fracture occurred when the total energy release rate reached a critical value, i.e. when

$$G_c = G_I + G_{II} \sin^2 \omega, \tag{2.9}$$

where ω can be regarded as the slope of the fracture surface roughness. When the fracture surfaces are smooth then $\omega = 0$ and $G_c = G_{Ic}$. The other part of G_{II}, $(G_{II}\cos^2\omega)$ is assumed to be lost in friction.

2.3.4 Mixed Mode Bending Test

A convenient mixed mode delamination experiment was proposed by Reeder and Crews [41], designated as the mixed mode bend (MMB) test. Their rig, illustrated in Figure 2.12, allows for a single load to simultaneously produce mode I and mode II bending loads on the specimen. By varying the position of the load, different fixed ratios of mode I to mode II behaviour can be obtained. The primary advantage of the MMB test rig is that it allows for the testing of identical specimens under a wide range of mixed mode loadings.

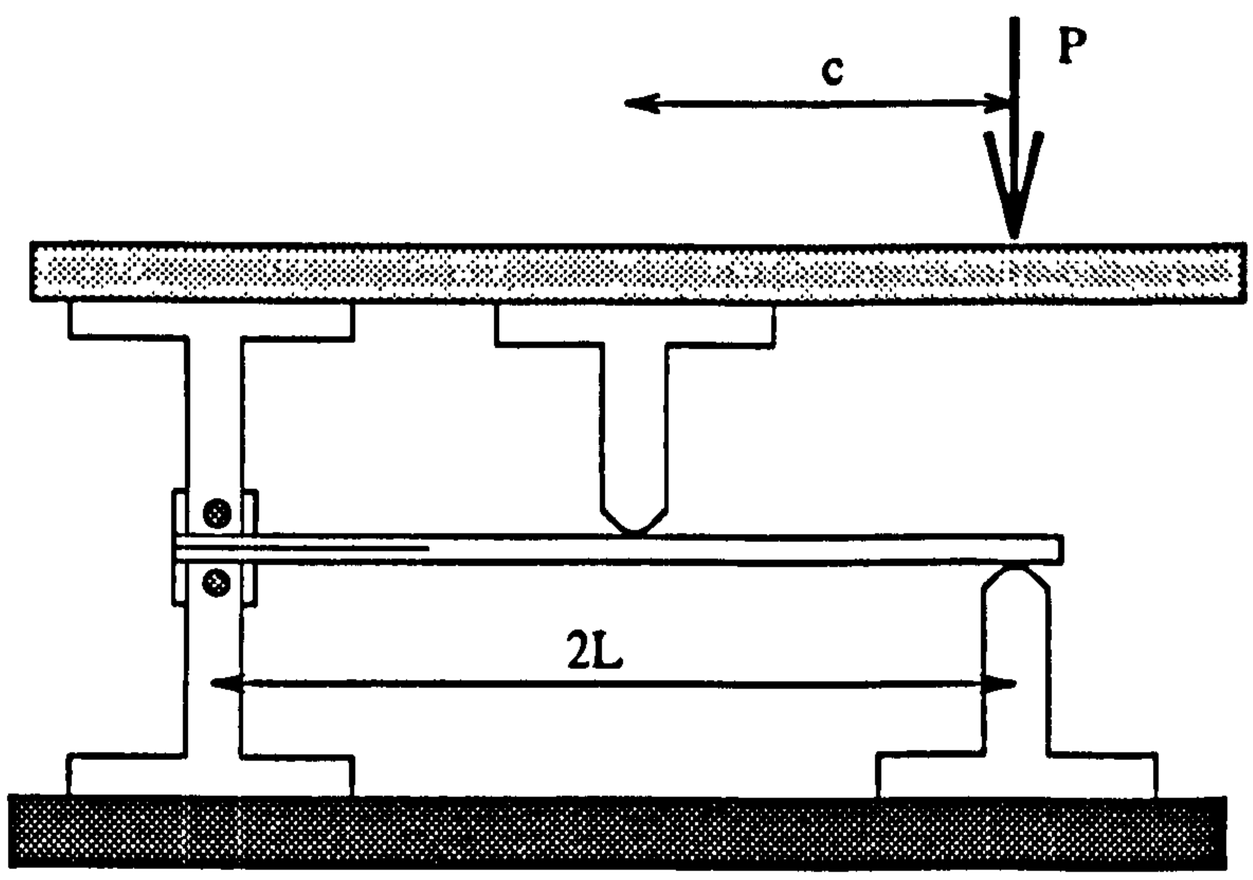


Figure 2.12: The mixed mode bend test rig.

Using a simple beam theory analysis, Reeder and Crews derived expressions for both the mode I and mode II components of the energy release rate. The mode I component of the applied load P, is given by

$$P_I = \left(\frac{3c - L}{4L}\right) P,$$

where c and L are defined in Figure 2.12. Using the conventional beam theory expression,

$$G_I = \frac{12a^2 P_I^2}{b^2 h^3 E_{11}},$$

they obtained

$$G_I = \frac{3a^2P^2}{4b^2h^3L^2E_{11}}(3c - L)^2.$$

Similarly for mode the mode II component of the load

$$P_{II} = \left(\frac{c + L}{L}\right) P.$$

Using the conventional beam theory expression,

$$G_{II} = \frac{9a^2P_{II}^2}{16b^2h^3E_{11}},$$

they obtained

$$G_{II} = \frac{9a^2P^2}{16b^2h^3L^2E_{11}}(c + L)^2.$$

An expression for the ratio of mode I to mode II was determined from the individual expressions given above for G_I and G_{II} as

$$\frac{G_I}{G_{II}} = \frac{4}{3}\left[\frac{(3c - L)}{(c + L)}\right]^2 \qquad c \geq \frac{L}{3}. \tag{2.10}$$

Notice that G_I/G_{II} is only a function of load position c and half-span length L. The G_I/G_{II} ratio is zero for $c = L/3$ and is invalid for smaller c values since this model does not take into account contact between the two arms of the specimens.

The total energy release rate given by

$$G = \frac{3a^2P^2}{16b^2h^3L^2E_{11}}[4(3c - L)^2 + 3(c + L)^2] \tag{2.11}$$

was compared by the authors to a finite element result, and they found that the beam theory equations resulted in an 15% underestimate of the finite element solution. Closer investigation revealed that the G_I values were in underestimate by 18%, while the G_{II} values were only 6% too low. Thus, the error was attributed to the beam theory equation for the DCB test.

Following the elastic foundation correction of Kanninen [27] and the shear deformation correction of Aliyu and Daniels [42], they wrote their expression for G_I as

$$G_I = \frac{3P^2(3c - L)^2}{4b^2h^3L^2E_{11}}\left[a^2 + \frac{2a}{\lambda} + \frac{1}{\lambda^2} + \frac{h^2E_{11}}{10G_{12}}\right], \tag{2.12}$$

where

$$\lambda = (3k/bh^3E_{11})^{1/4}$$

and

$$k = 2bE_{22}/h.$$

Adding the shear deformation term into the expression for G_{II} yielded

$$G_{II} = \frac{9P^2(c+L)^2}{16b^2h^3L^2E_{11}}\left[a^2 + \frac{0.2h^2E_{11}}{G_{12}}\right]. \tag{2.13}$$

After discovering difficulties with the original design when testing tough composite laminates, the original rig was redesigned by the same authors. The problem with the original design was that large rotations of the loading lever produced nonlinear geometric behaviour resulting in errors of up to 30% in calculating G_I and G_{II}. A simple nonlinear iterative analysis was presented by the authors [43] for correcting for this problem. However, the most straightforward solution involved a simple modification of the rig. The modification employed the addition of a saddle and bearing assembly to eliminate the horizontal component of the applied load P.

2.4 Some Practical Comments

The use of interlaminar fracture tests to measure the delamination resistance of unidirectional composites is now widespread. Standards have been adopted for testing and interpreting the results of both mode I and mode II experiments. However, some practical difficulties will inevitably complicate delamination fracture testing of multidirectional composite laminates. In this section, an attempt is made to summarize several important issues.

2.4.1 Stability of Fracture Tests

The stability of a fracture experiment refers to the ability to grow stable cracks in a given specimen geometry. Some fracture experiments, like the DCB test are inherently stable, and always provide stable crack growth. In such cases, several values of the energy release rate can be determined during a single experiment.

In this section, a simple analysis is discussed for determining what conditions allow for stable crack growth in the ENF fracture experiment.

The stability of crack growth may be judged from the sign of dG_{II}/da—if dG_{II}/da is positive, unstable crack growth will occur, while stable crack growth occurs if dG_{II}/da is zero or negative. Following Carlsson et al. [30], the 'fixed grip' (displacement control) expression for the energy release rate is

$$G_{II} = \frac{\delta^2}{2bC^2}\frac{dC}{da}.$$

Differentiation of this expression yields

$$\frac{dG_{II}}{da} = \frac{\delta^2}{2bC^2}\left[\frac{d^2C}{da^2} - \frac{2}{C}\left(\frac{dC}{da}\right)^2\right].$$

Neglecting the influence of shear, the above expression can be written

$$\frac{dG_{II}}{da} = \frac{9\delta^2 a}{8E_{11}b^2h^3C^2}\left[1 - \frac{9a^3}{2L^3 + 3a^3}\right],$$

and for stable crack growth,

$$a \geq 0.7L.$$

A similar analysis can be applied to other potentially unstable fracture experiments, like the MMB test to determine appropriate expressions to guide in the design of specimens.

2.4.2 Rate Effects

A general tendency for polymeric materials is to exhibit rate dependent mechanical properties. In the case of fracture toughness values, rate dependence is closely linked to the development of damage zones at the delamination front. The actual damage that occurs before delamination will strongly influence the subsequent fracture.

By carefully monitoring DCB tests with acoustic emission transducers, de Charentenay, Harry, Prel, and Benzeggagh [44] were able to study in detail the damage development in the DCB test. The monitoring of acoustic emission events during their experiments provided a unique means to detect the onset of delam-

ination, and to qualitatively explore the nature of stable damage development. It was clear from their acoustic emission results that the mechanisms involved in the initiation and further stable growth of delaminations were different.

Unique acoustic signatures were correlated to the fracture surface morphology of laminates of various materials. A close comparison of load-displacement and fracture resistance curves to the acoustic emission data clearly revealed the complicated development of damage during delamination, however, two specific mechanisms were identified. Fracture initiation was identified with the onset of microscopic matrix/fibre debonding, and subsequent damage development was controlled by multiple cracking and large scale interfacial debonding. They concluded that material development investigations should take into account these two distinct modes of failure—a good material must not only have a high initiation fracture energy, but also must develop extensive damage before and during delamination propagation.

The concerns alluded to by de Charentany regarding the different stages of delamination propagation were further echoed by Aliyu and Daniel [42]. Their objective was to evaluate the various experimental techniques and analysis methods for the characterization of interlaminar fracture toughness and to determine the effects of strain rates on such measurements for a graphite/epoxy composite. These authors developed an expression for the energy release rate in the DCB specimen that accounted for the kinetic energy of the beam. In addition, they included shear deformation effects and an additional term to account for the rotation of the built in end of the beam.

Their results with graphite/epoxy composites indicated that the energy release rate increases with crack velocity by up to 28% for the ranges of crack velocities studied (8.5–51 mm/s).

2.4.3 Determining the Critical Energy Release Rate

Another issue related to the stability of a fracture test is the determination of the *critical* energy release rate from the experimental data. Indeed, much concern has been expressed regarding this point, for both the DCB and particularly, the ENF specimen.

In his early investigation of the DCB fracture test, Berry [16] observed that

there was always a certain degree of arbitrariness involved in defining the 'equilibrium' condition from which the *critical* energy release rate should be calculated. A true steady state, at constant deflection could not be realized and the crack continued to increase in size at a slow, but measurable rate. His conclusion for DCB tests of crack growth in polymeric materials was that so long as the specimen was in a strained condition, a true equilibrium condition may never be realized.

The problem of defining the *critical* energy release rate in composite laminate testing translates into identifying correctly the *critical* load, P_c and the *critical* displacement, δ_c from the load-displacement plot of an experiment. For DCB experiments on composite laminates, this is not as difficult as Berry's conclusions suggest.

Because the DCB test is a stable one, a steady crosshead rate can drive the cracked laminate to a critical condition, and then, as the displacement is increased further, subsequently extend the crack in a stable manner. The specimen can be unloaded and the process repeated several times, to provide enough data to generate an R-curve. Figure 2.13 illustrates a typical load-deflection plot for a DCB specimen. The main problem that arises in extracting the critical load and displacement from such plots stem from occasional nonlinear material behaviour just prior to crack growth. This problems can be circumvented by adopting a (somewhat) arbitrary definition of the critical load, for example, the maximum load, the load determined at a 5% increase of the compliance, or simply the load at the visual onset of crack growth [45]. In the case of the DCB test, there is not too much difference between energy release rates determined using these various methods, providing a consistent method is used.

The ENF specimen, due to its inherent instability and subcritical crack growth, presents a more difficult task for determining the *critical* condition. As with the DCB test, the critical load can be arbitrarily taken as the maximum load, the load at 5% increase in compliance, etc., but for the ENF specimen, there is a larger difference between these definitions.

A discussion regarding the various definitions of the critical load by Davies *et al.* [46], however, provides a reliable framework. A lower bound value can be deduced by noting the first point of nonlinearity of the load-deflection plot, while an upper bound will be given by the maximum load. The logical compromise is

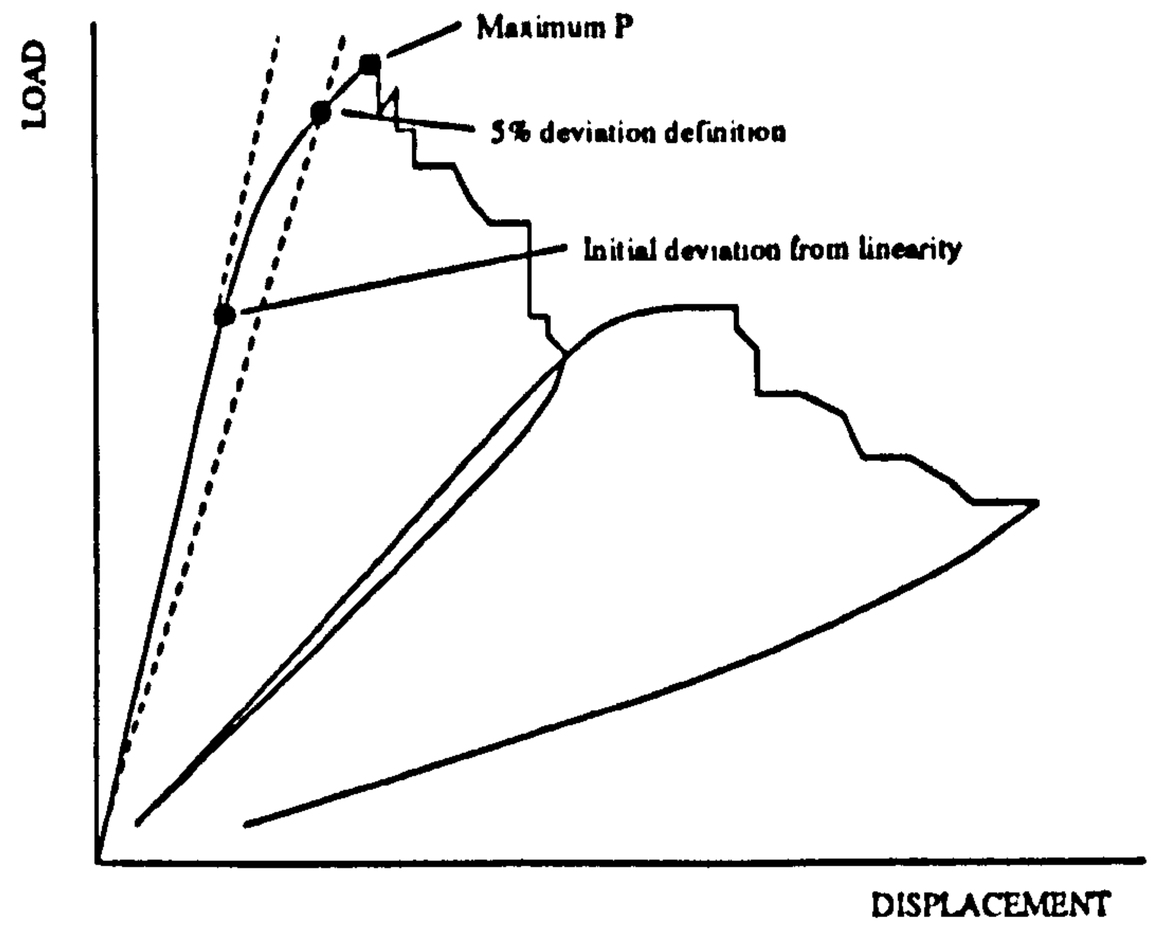


Figure 2.13: A typical load-displacement plot from a DCB experiment.

to use the 5% offset value.

2.4.4 The European Group on Fracture Protocol

The European Group on Fracture (EGF) has devoted considerable attention to establishing the reproducibility of energy release rate values for composite laminates. They produced a protocol [45] summarizing the various data reduction schemes for both the DCB and ENF tests. They also included recommendations about the correct design of specimens, end blocks etc., and specific procedures for loading the specimens and reducing the data. Several aspects of the protocol were evaluated in a round robin exercise [46], and a summary of their recommendations is presented below.

For all tests, the EGF recommended a specimen width of 20 mm, a thickness of 3 mm and a nominal length of 125 mm or 165 mm for DCB and ENF specimens, respectively. Starter defects were found to greatly improve the repeatability of the tests, provided the thickness of the inserts was kept below 15 μm. Load application via hinges or end blocks was found to be adequate, but the appropriate correction factors presented by Williams [26] should be used if necessary. In an attempt to provide a similar conditioning for all specimens, a drying procedure was recommended. Specifically, the specimens should be stored for a minimum of 10 days at 77°C.

For the stable DCB tests, the specimens should be loaded in displacement

control at a crosshead rate of 0.5 mm/minute. This rate was chosen so that the crack growth could be followed with the help of a ×70 microscope. A thin layer of white ink was suggested to help identify the position of the crack. An observer should note the first sign of deviation from nonlinearity in the load-displacement plot, and then follow the crack for a growth of approximately 5 mm. After each 5 mm growth increment, the specimen should be unloaded and reloaded to allow for determination of the compliance at each successive crack length. The data reduction procedures recommended by the EGF for the DCB test are the corrected beam theory method and Berry's experimental compliance method.

For the ENF mode II specimens, a fixed span ($2L$) of 100 mm was recommended, with a minimum crack length to half span ratio (a/L) of 0.5. The EGF was more specific with the precracking of ENF specimens, since there is a greater sensitivity in these specimens to the initial nature of the delamination. Testing directly from the Teflon inserts produced a large scatter in the results, so two methods for introducing precracks were studied. Specifically, tests were conducted from mode I and mode II precracks. The mode I precracks seemed to provide the most reliable method, with a scatter of G_{IIc} values on the order of 20%. Again the corrected beam theory method and Berry's experimental compliance technique were recommended to reduce the experimental data.

Chapter 3

Experimental Details

The purpose of this work was to develop an alternative method of measuring the energy release rate for delaminations in composite laminates, based on an experimental photomechanics approach using moiré interferometry. The first experimental objective was to validate this new method, by using, in the first place, conventional delamination toughness test methods as a benchmark.

Three types of experiments were used to validate the moiré method:

- pure mode I, double cantilever beam (DCB)
- pure mode II, end notched flexure (ENF)
- mixed mode-mixed mode bending (MMB)

During the validation phase, the limitations and advantages of the moiré technique relative to the conventional delamination test methods were discovered. Specifically, the problem of determining the critical energy release rate was encountered. The approach used to validate the moiré method, therefore, was based on a wider interpretation of the conventional data analysis.

The standard DCB and ENF tests and their associated data reduction procedures are described in this chapter to provide relevant background. Details about how the moiré measurements were integrated into the standard test procedures are specifically discussed.

3.1 Specimen Geometry and Material Properties

A single specimen geometry was used in all types of the tests performed for this work. Figure 3.1 shows the relevant dimensions. This specimen provided a useful range of approximately 35–40 mm of crack growth with an initial crack length of approximately 40 mm.

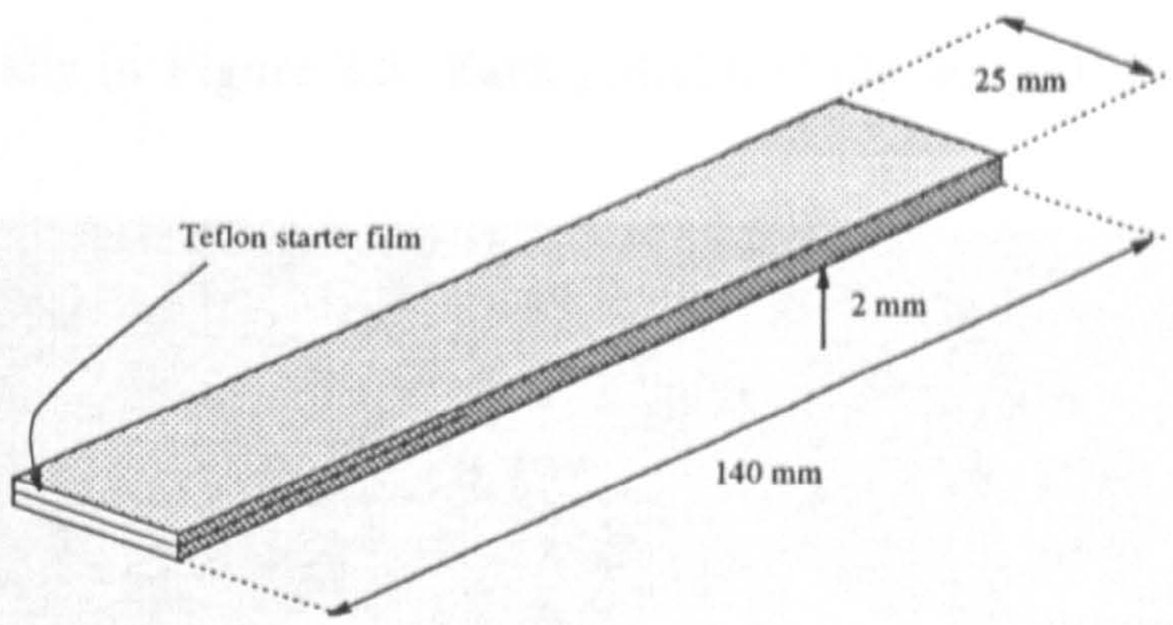


Figure 3.1: Standard specimen geometry.

The unidirectional specimens were cut from plaques composed of 16 layers of T300/914C pre-preg. The fibres were oriented along the length of the specimens. Artificial delaminations in the form of double teflon film were introduced prior to curing at two different depths within the plies. On one side of the plaque, the teflon film was located at the midplane (between plies 8 and 9) to provide symmetrically delaminated specimens. Asymmetric specimens were fabricated on the other half of the plaque by placing the teflon starter film between plies 13 and 14.

Laminate properties as provided by the manufacturer for the 0° material are summarized below in Table 3.1. The notation used is as follows: E_x is the elastic modulus in the fibre direction, E_y is the elastic modulus perpendicular to the fibres (effectively the same for both the thickness and the width direction), G_{xy} is the shear modulus, and ν_{xy} and ν_{yx} are the two Poisson ratios.

In addition to the unidirectional specimens, several plaques of multidirectional specimens were also fabricated. The multidirectional layup chosen for this work was designated as $[0_2/+45/0_2/-45/0/90]_s$. The sixteen layer laminate is illus-

E_x (MPa)	128615
E_y (MPa)	9672
G_{xy} (MPa)	5443
ν_{xy}	0.476
ν_{yx}	0.036

Table 3.1: 0° Unidirectional Laminate Properties.

trated graphically in Figure 3.2. Each individual ply was 0.125 mm thick, so the

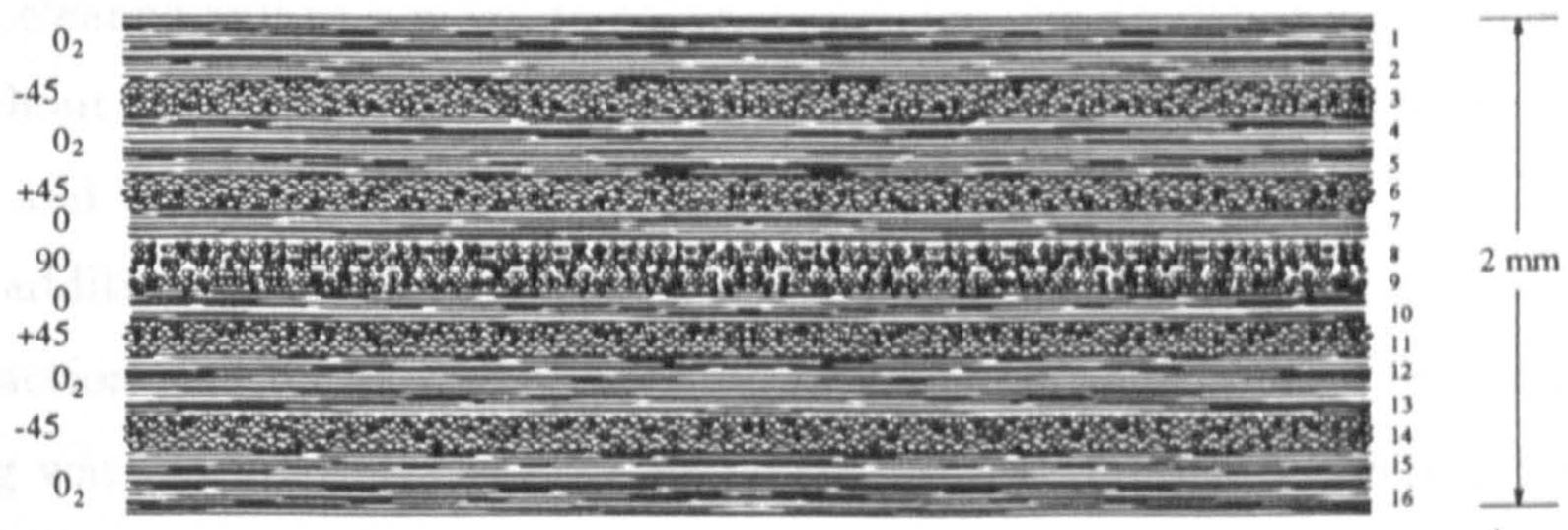


Figure 3.2: Multidirectional laminate used in this work.

total thickness of the cured laminate was 2 mm. Multidirectional plaques with both symmetric and asymmetric delaminations were fabricated and tested.

3.1.1 Specimen Preparation and Conditioning

The DCB and MMB tests required the bonding of aluminium end blocks on to the specimens as illustrated in Figure 3.3. The end blocks were necessary to mount the specimens in the testing rig. Since the ENF test utilized a simple a three point bend arrangement, no end blocks were required for these tests.

A special jig was made to insure that the specimen end blocks were mounted correctly. The key issue was to mount the end blocks squarely with respect to the specimen, so that when the specimen was placed in the testing rig, the loading pins would fit with a reasonable tolerance.

The procedure for mounting the end blocks was straightforward. A specimen was placed in the jig, aligned, and the end blocks were bonded to the specimen surfaces using epoxy. After testing, the end blocks could be snapped off the spec-

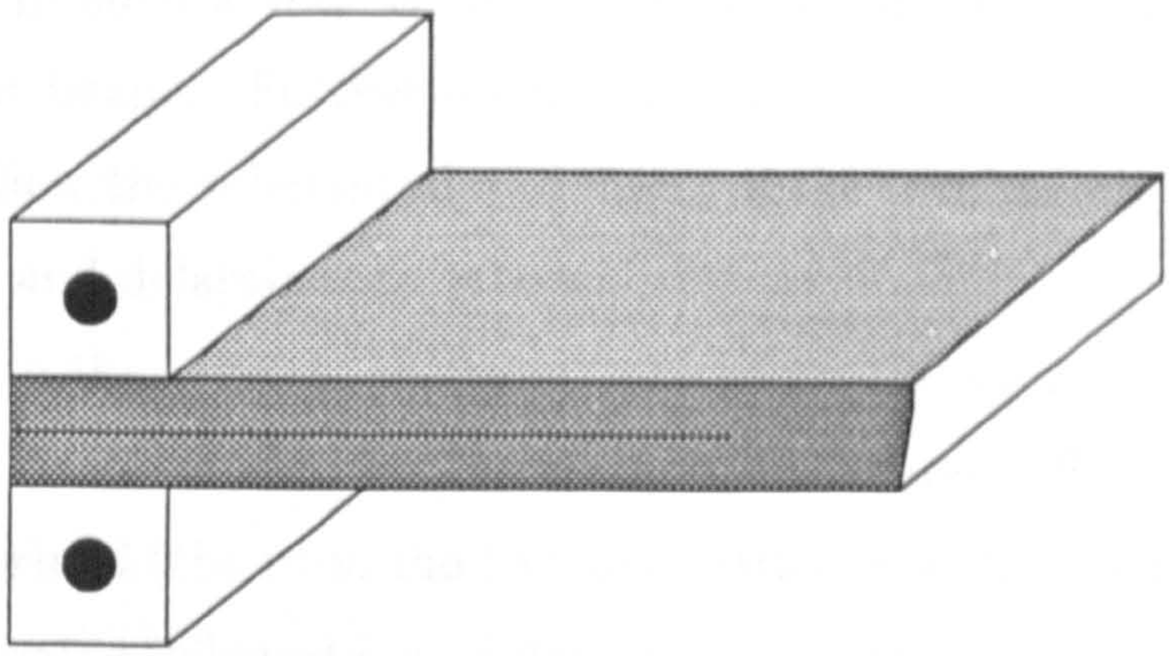


Figure 3.3: A schematic of the end block/specimen configuration.

imens, cleaned and re-bonded to other specimens. No problems were encountered throughout all of the experiments with the strength of the bond between the end blocks and the specimens.

In addition to bonding of the end blocks, it was also necessary to introduce a diffraction grating on one of the thin edges of the specimens. The diffraction grating was required for the experimental optical technique of moiré interferometry. The types of gratings used, and the necessary specimen preparation will be discussed in more detail in Chapter 4, and moiré interferometry is dealt with in Chapter 5.

It was also necessary to precrack the specimens before the experiments. To insure the introduction of suitable precracks, non-destructive scanning (NDE) was used to map the delaminated surfaces of the specimens during and after the precracking procedure. These procedures will be summarized in Section 3.5.

Prior to testing, the specimens were stored in an airtight container containing a fresh supply of silica gel, in order to insure a consistent moisture content from specimen to specimen. Specimens were tested within minutes of being removed from this controlled environment.

3.2 Mechanical Testing Rig

Since the experiments involved simultaneously capturing optical fringe patterns and standard load/displacement records, several special considerations regarding the design of the load apparatus were required. It was important to design

the load frame in such a way as to provide sufficient clearance for the optical components and beams. Furthermore, it was necessary to provide a means for the optics to follow the substantial movement of the specimen in response to the applied loading and delamination growth.

In addition to the optical constraints placed on the load apparatus, as mentioned above, it was also necessary to consider that three different types of tests were to be performed (the DCB, the ENF and MMB). For these reasons, the design of the load apparatus followed a modular approach, involving a solid base utilized by all three tests, and various components which were fitted for the individual tests.

Two photographs of the experimental apparatus are shown at the end of this chapter in Figure 3.19.

3.2.1 The Base of the Load Frame

A solid base was bolted directly to the optical table, and a platform mounted on top of the base to accommodate the various fixtures for each test. The function of the base was to provide a stable yet adjustable support for the platform to ride on, so that the necessary positioning adjustments could be made during the experiments to ensure that the crack-tip was centred and in focus with respect to the optics. The base, illustrating the necessary degrees of motion is schematically depicted in Figure 3.4.

The entire load frame assembly was small and compact and it was a relatively simple matter to provide the necessary translation and inclination of the specimen during an experiment. The base allowed for three degrees of freedom—one horizontal translation, vertical translation, and rotation in the plane containing those two. As with all interferometric techniques, vibration isolation and stability were of extreme importance. Thus, it was necessary to provide these movements in such a way as to not degrade the stability of the entire platform.

On top of the base, a 12 mm thick steel platform was mounted. By changing the various fixtures and load cells, it was possible to perform the DCB, ENF and MMB tests on this platform.

All of the tests utilized a hand driven load screw, which is detailed in Figure 3.5. This device was constructed by modifying an already existing positioning

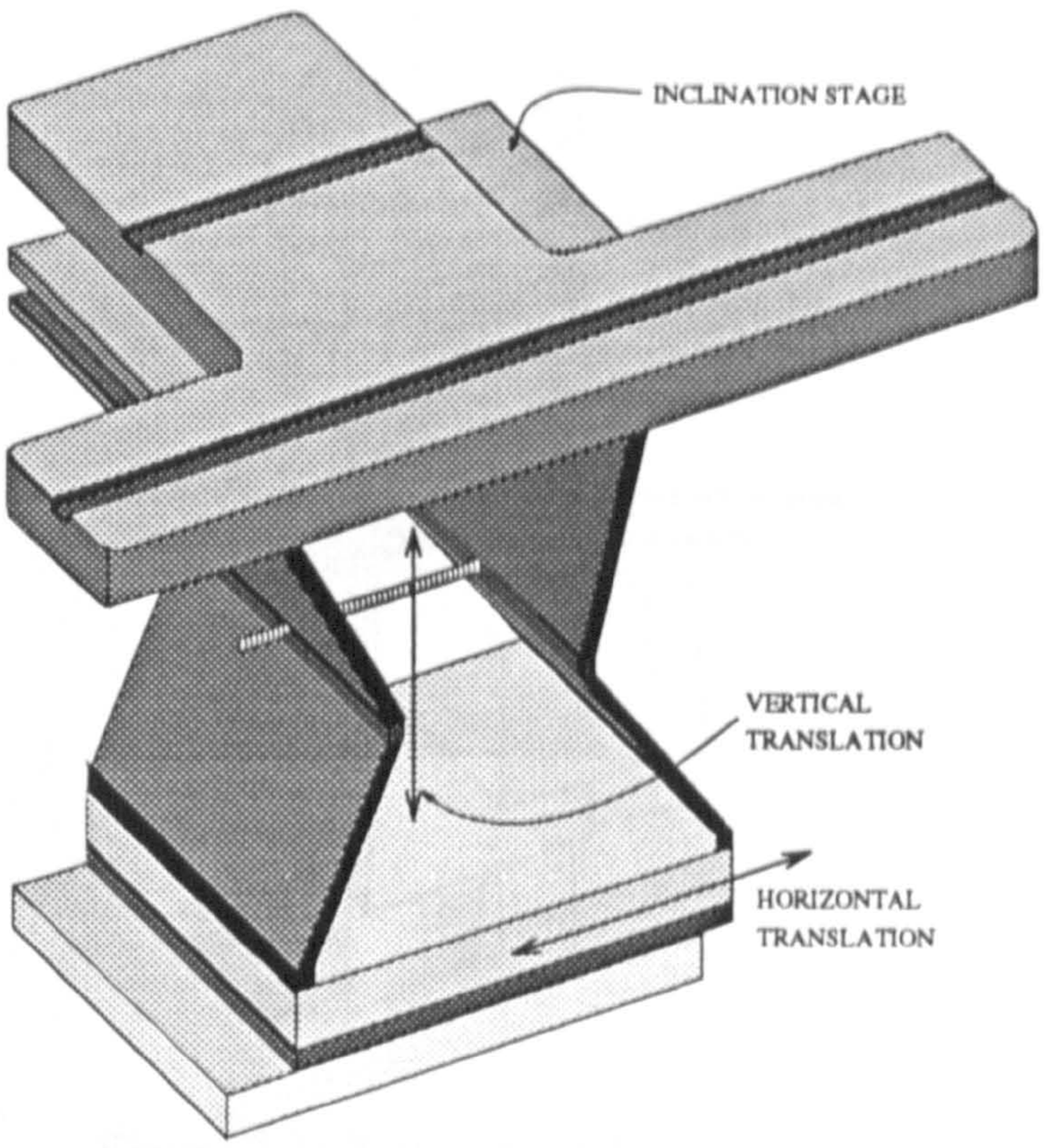


Figure 3.4: The base of the load frame

device. A carriage was driven by the screw, upon which a block was mounted to provide a location for the load cells to be secured. Teflon pads were introduced between the carriage and the face of the screw mount to allow smooth, backlash-free translation of the carriage as the screw was turned. One revolution of the screw yielded a 0.5 mm linear displacement of the mounted carriage, which provideded both sufficient resolution and mechanical advantage for all of the tests.

Vertical displacement of the carriage was monitored using an LVDT, mounted as shown in Figure 3.5. This transducer had a linear range of 30 mm. Several pairs of mounting holes on the carriage allow for the LVDT to be located appropriately for each type of test.

The adjustable base, the platform and the load screw constituted the main components of the load frame. Each of the three tests, however, required additional components, which are described in the sections below.

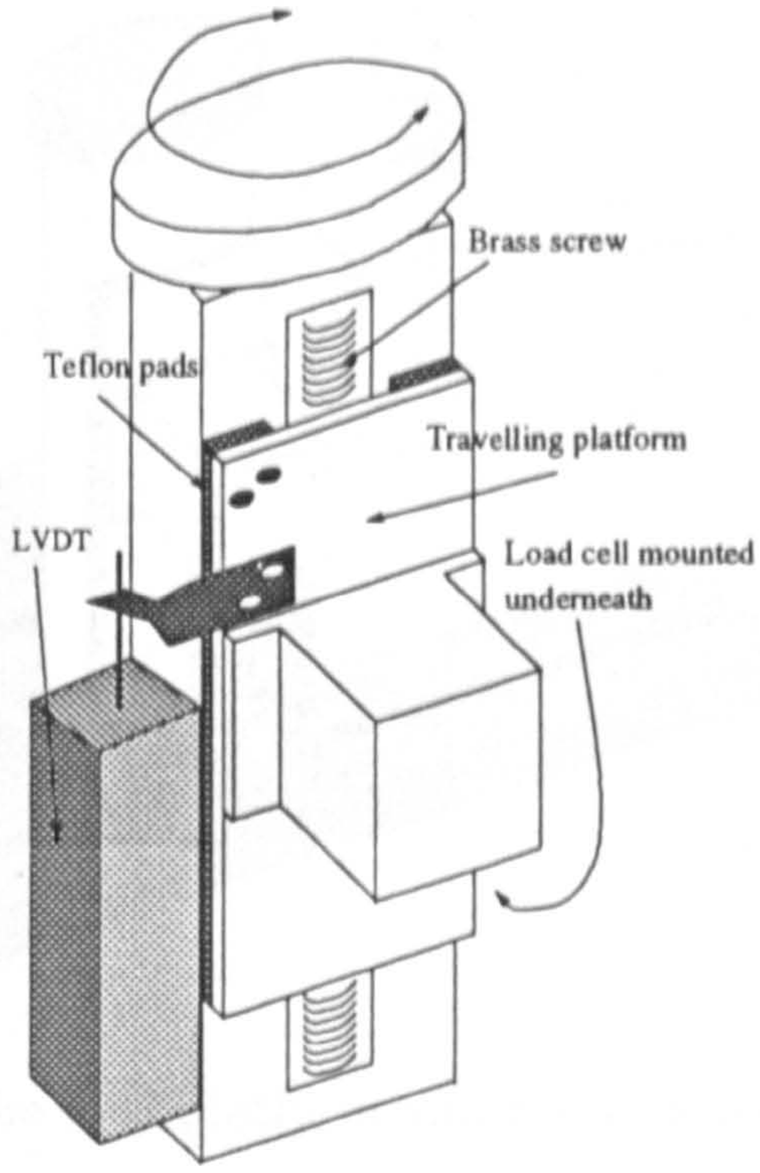


Figure 3.5: The hand-driven load screw

3.2.2 Double Cantilever Beam Test Fixtures

In the DCB test, the objective was to separate the two halves of the specimen in a symmetric manner, creating a pure mode I opening of the delamination. Apart from the equal and opposite forces on each of the arms of the specimens, no other applied forces or moments were desired.

This symmetric type of loading was most readily accomplished by simply pulling the two arms of the specimen away from each other, as illustrated in Figure 3.6. The DCB tests required two additional mounts that were connected to the specimen end blocks via smooth pins.

The tensile loads were measured using a miniature, 50 N full-scale load cell, mounted as shown in Figure 3.6. This load cell provided a sensitivity of 5 mV/N. The typical maximum load during a DCB test was approximately 35 N.

Due to the construction of the load cell, it was not possible to screw the device firmly into the load cell mounting block. It was necessary to use hard rubber washers between the load cell and the two surfaces that it was connected to. The rubber washers allowed the load cell to be tightened sufficiently, without introducing stress in the device.

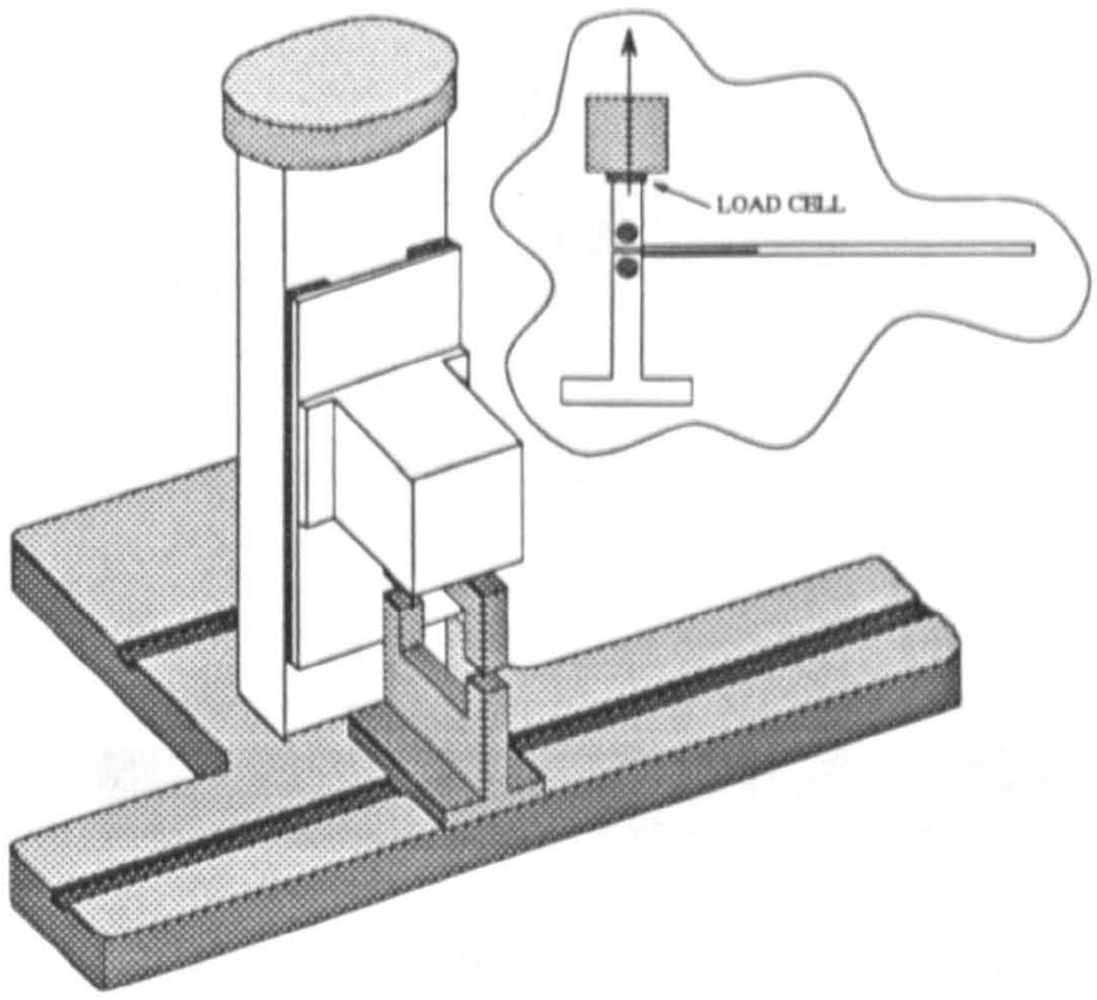


Figure 3.6: The load platform and the DCB test arrangement

Special care was taken to insure that the various loading fixtures were properly aligned. This was necessary since any mis-alignment would place an undesired load or moment on the specimen during loading, and cause the delamination to grow in an unsymmetric manner.

3.2.3 End Notched Flexure Test Fixtures

The ENF test was based on a simple three point bend loading arrangement, the objective being to produce a state of pure mode II shear load on the delamination front. Figure 3.7 illustrates the ENF test.

The loading fixtures were made of steel and their edges were carefully rounded to provide flat and smooth contact with the specimen. Again, the parallelism between the fixtures was the most crucial alignment concern.

For the ENF tests, loads were measured using a more heavy duty tension and compression load cell, with a range of 2500 N full-scale, (with a a sensitivity 0.08 mV/N). The maximum load during a typical test was only approximately 500 N, so the range of this load cell was not exploited as successfully as with the DCB load cell. Regardless, suitable precision was obtained with the ENF load cell.

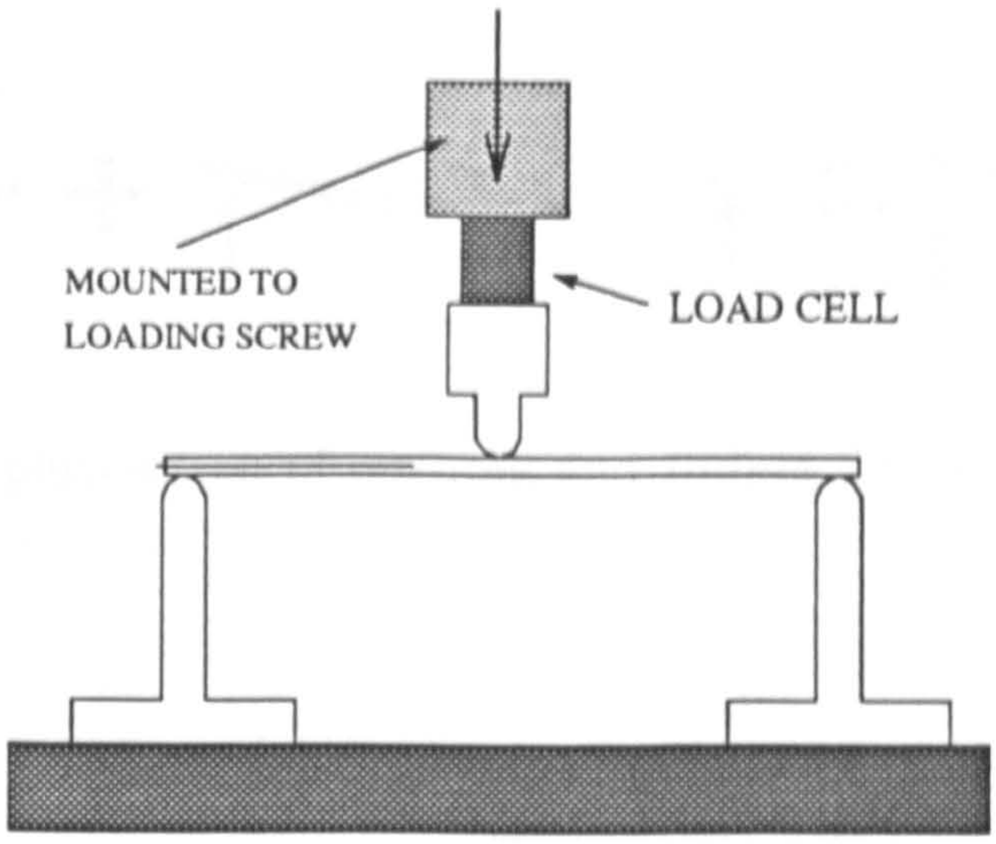


Figure 3.7: The ENF test arrangement

3.2.4 Mixed Mode Bending Test Fixtures

The DCB and ENF tests were based on simple load applications that could be accomplished in relatively straightforward ways. The MMB test, on the other hand, was more complicated, since it involved providing a superposition of the DCB and ENF tests. Thus, both mode I and mode II loading were applied to the specimen, simultaneously.

The opening and sliding modes associated with the mixed mode bend test were applied to the specimen through a single load acting on a lever as illustrated in Figure 3.8. By varying the position of the fulcrum, different amount of mode I versus mode II were achieved.

The relative amount of mode I to mode II will be called the mixed mode ratio, and defined as the ratio of the independently determined energy release rates as follows:

$$\Psi \equiv \frac{G_I}{G_{II}}$$

Thus for pure mode I loading, $\Psi = \infty$, and for pure mode II, $\Psi = 0$.

A straightforward implementation of the fulcrum/lever idea of Figure 3.8 was first proposed by Reeder and Crews [41], and is schematically illustrated in Figure 3.9. Using this design, a range of mixed mode ratios could be obtained, by simply changing c, the distance between the applied load and the fulcrum.

The details of the analysis of this original rig, as carried out by Reeder and Crews, were presented in Section 2.3.4. An important result of this analysis

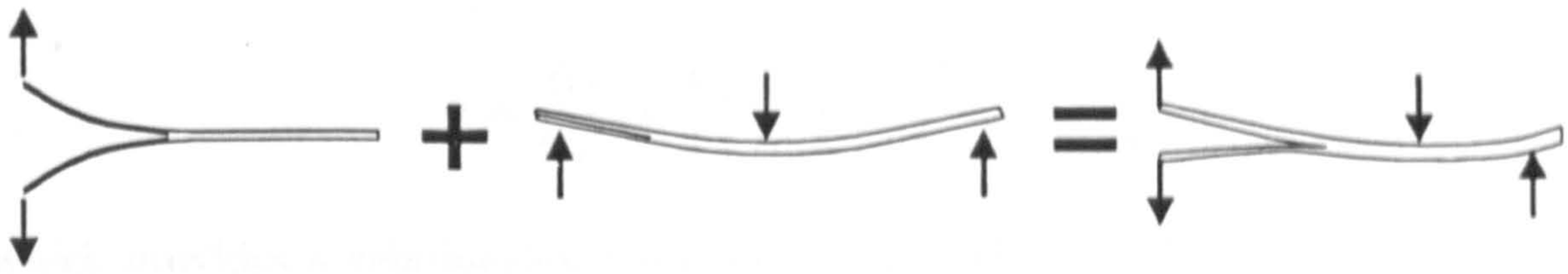


Figure 3.8: The superposition of opening and sliding deformation modes in the MMB test.

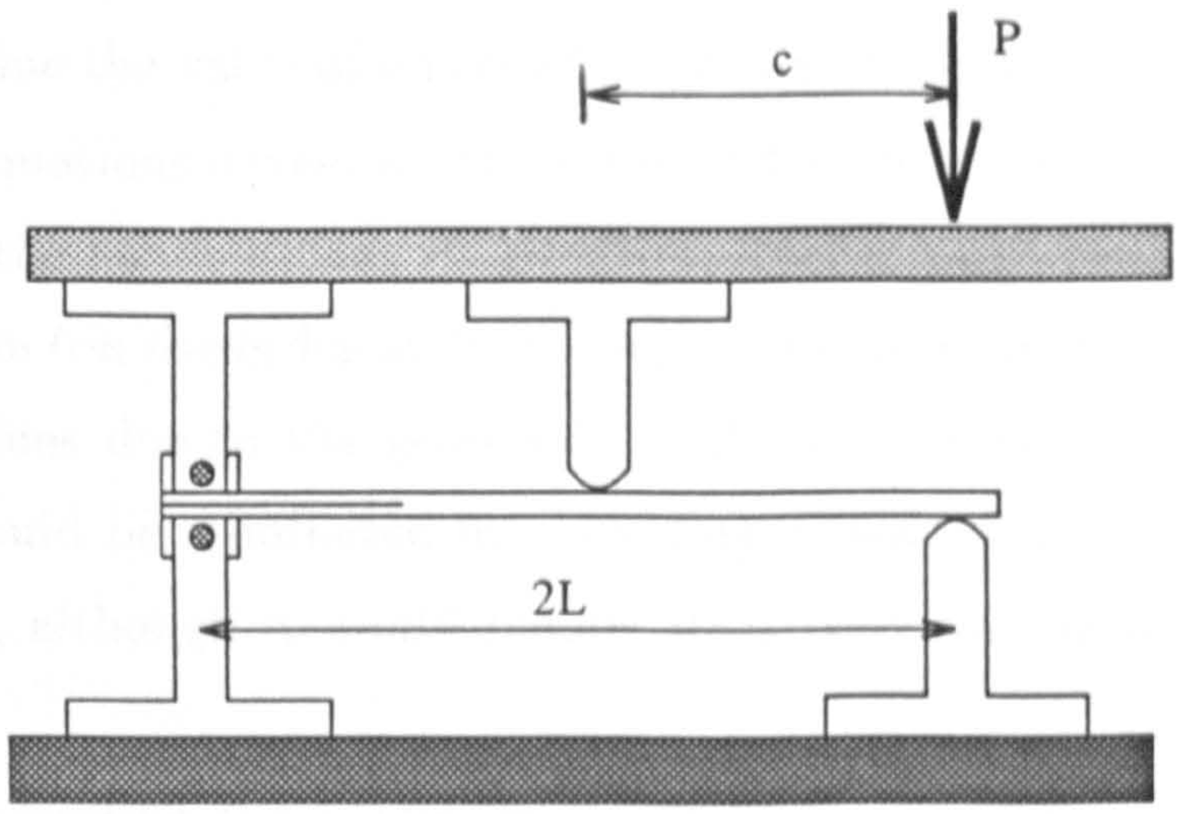


Figure 3.9: The original MMB test rig.

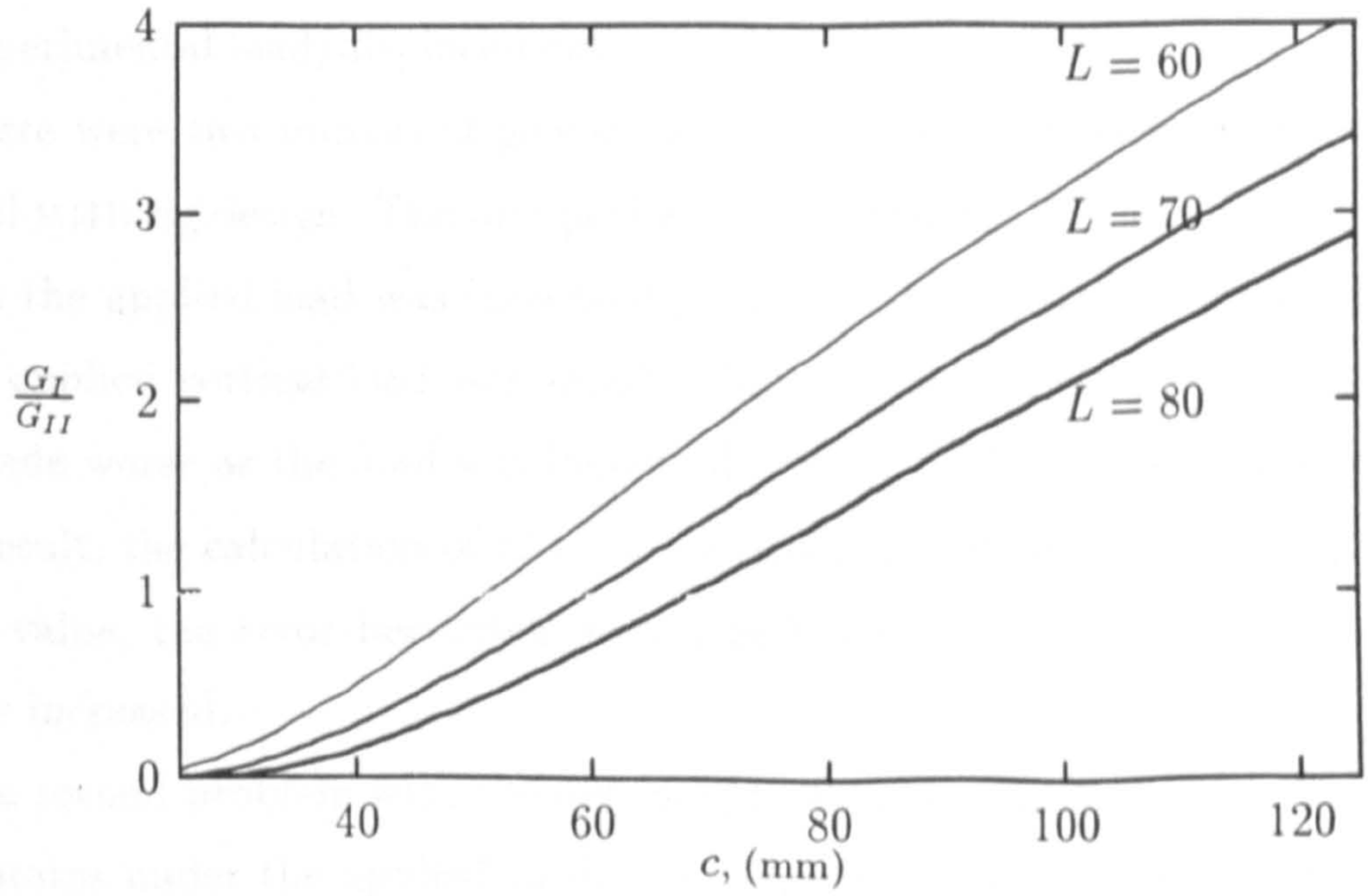


Figure 3.10: Mixed mode ratio ,Ψ, as a function of c and L.

relevant to the design and use of this rig was Equation 2.10,

$$\Psi = \frac{G_I}{G_{II}} = \frac{4}{3}\left[\frac{(3c - L)}{(c + L)}\right]^2 \qquad c \geq \frac{L}{3}$$

which provides a relationship between c, L and the mixed ratio, Ψ. Using the nominal laminate properties as given above, a plot of the expected mixed mode ratio versus c for three values of L is shown in Figure 3.10.

For the above equation to be valid, $c \geq L/3$, because the model does not account for contact between the two arms of the specimens. Figure 3.10 can be used to determine the value of c needed to provide for a desired value of Ψ.

Using the equations developed in Section 2.3.4, values of G, G_I and G_{II} were obtained from the applied load, P. However, Reeder and Crews discovered that significant errors (on the order of 30 %) were introduced in the calculated energy release rate values due to the geometric nonlinearities associated with the rig. These errors could be eliminated by providing a more complicated data reduction procedure, although it would require iteratively solving a set of nonlinear equations.

Instead of adopting the nonlinear equation solving as the standard method for reducing their experimental test data, they used their nonlinear analysis to redesign the MMB rig, so as to minimize the geometric nonlinearity. In this way, the linear equations (those presented in Section 2.3.4) could still be used to reduce the experimental load/displacement data.

There were two sources of geometric nonlinearity that were identified in the original MMB rig design. The first problem was a result of the rotation of the lever arm as the applied load was increased (or as the delamination grew). Thus, the actual applied vertical load was smaller than the measured load, and this effect was made worse as the load was increased. Figure 3.11 illustrates this situation. As a result, the calculation of G from the applied load would underestimate the actual value, the error becoming more significant as the total deflection of the lever is increased.

The second problem with the original MMB rig was more subtle. As the lever arm rotates under the applied load, the roller applying the load will translate a distance of Δc as illustrated in Figure 3.12. The changing load location causes the moment arm to become shorter, thus reducing the effective moment actually

opening the specimen. Again this will result in an underestimate of the energy release rate. This problem was minimized by effectively lowering the lever height, so that the magnitude of Δc was reduced, as shown in Figure 3.12.

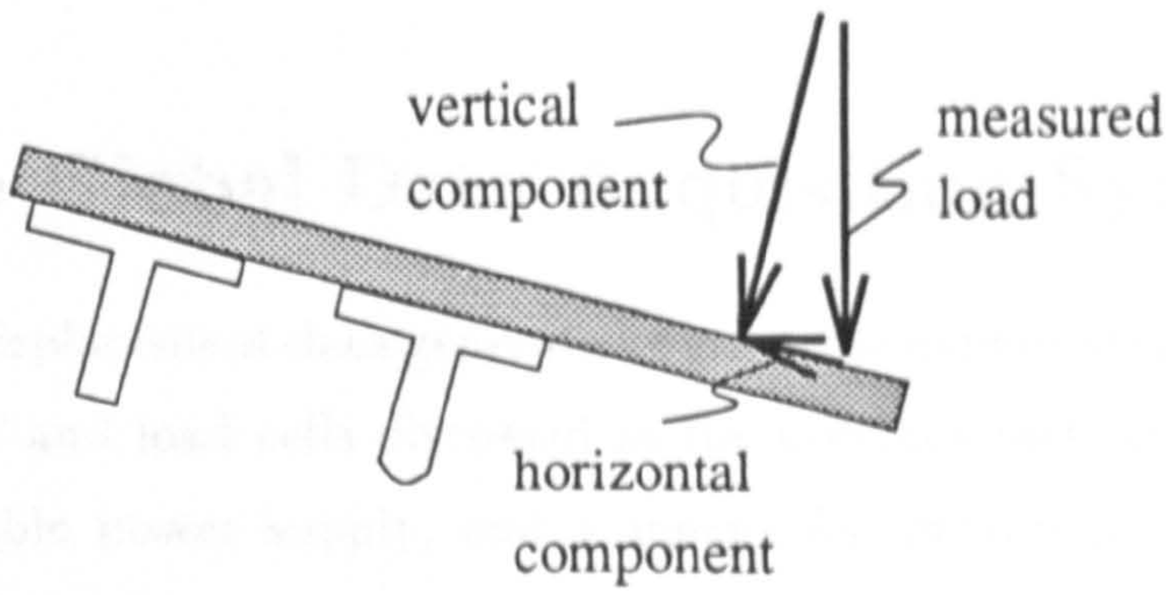


Figure 3.11: The discrepancy between the measured and applied loads in the original design of the MMB rig.

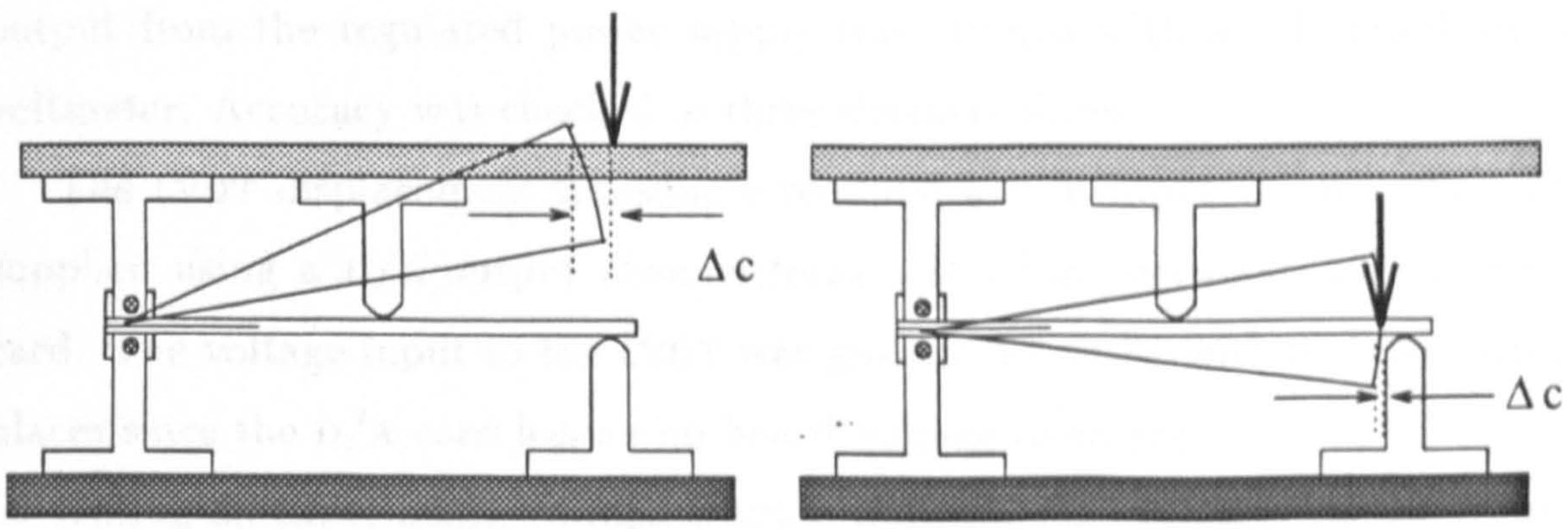


Figure 3.12: Effect of lever height on the load point translation.

The modification proposed by Reeder and Crews were adopted for the design of the MMB rig used for this work. This design includes a more sophisticated saddle/yoke assembly to apply the load, that extends down the sides of the lever and the specimen. The rollers of the saddle are in contact with a yoke that transmits the load to the top lever in a way that minimizes Δc. A schematic of the saddle/yoke system used to apply the load is shown in Figure 3.13.

Other improvements were also incorporated in the design of the MMB rig. The smooth, flat surfaces of the three point bend contacts were replaced with bearing-mounted rollers. The pins that secured the end blocks to the rig were also replaced with roller and bearing assemblies. The entire MMB fixture was made from aluminium to reduce the weight and to facilitate tighter design tolerances.

For the MMB tests, loads were measured using a similar tension and compression load cell, with a range of 250 N full-scale, (0.859 mV/N). The maximum load encountered during a typical test was approximately 150 N, so the range of this load cell was appropriate.

3.3 The Global Data Acquisition System

The load and displacement data generated during the experiments were measured using the LVDT and load cells discussed in the previous sections. These devices require a suitable power supply, and a means for converting the raw voltage outputs into displacements and loads. In this section, the system responsible for both of these aspects is discussed.

All three load cells required a stable 15 V input voltage. A conventional regulated dc power supply was used for this task. Before each test, the voltage output from the regulated power supply was verified with a calibrated digital voltmeter. Accuracy was checked to three decimal places.

The LVDT displacement transducer required a 10 V input voltage. This was supplied using a D/A output channel from a standard personal computer lab-card. The voltage input to the LVDT was guaranteed to be within three decimal places since the D/A card has an on-board voltage reference.

During an experiment, voltage outputs from the load cell and the LVDT were input to two A/D converter channels also on the PC lab-card. Digitization of the input signals occurred at a rate of 1 kHz. The resulting digital voltage values were conditioned with Finite Impulse Response (FIR) digital filters. These filters were designed to provide low pass filtering of the incoming signals, for the purpose of removing spurious spikes from the signals. The filters used a sampling window of 64 digitized values and were designed to pass only frequencies less than 20 Hz (events with a duration of $> 1/20$ s). The digital filters also facilitated the implementation of a simple digital data compression routine that reduced the volume of information that needed to be stored on the PC's hard drive.

A real-time load versus displacement curve was constructed from these filtered signals and plotted to the screen during an experiment to facilitate careful loading/unloading of the specimen. The data was simultaneously stored in a

(compressed) data file for post-processing, to extract compliances, displacements and critical loads.

The PC lab-card was only one of three plug-in devices installed in the laboratory personal computer. The other two cards, a frame grabber and an array processor, will be discussed in more detail in Chapter 5. All three of these cards were under the control of one computer programme during an experiment. This arrangement allowed for coordinated data acquisition from the various sources (load, displacement and moiré fringe image data).

The programme responsible for the management of these cards was called XMELT. Written in the C language, XMELT provided a uniform interface to the software libraries supplied with the lab-card, the frame grabber and array processor. XMELT was based on a fairly simple macro language, so that the user could define commonly performed activities, thus saving time during an experiment.

Although it was developed for this work, XMELT represents a general purpose, user programmable data acquisition system. Since most of the complexity of XMELT is concerned with the moiré fringe acquisition, it will be described in more detail in Chapter 5. A schematic of the complete data acquisition system (XMELT) is shown in Figure 3.14 for reference.

3.4 Global Data Reduction and Testing Details

The global data reduction procedures used for the DCB and ENF tests were based on commonly accepted methods, similar to those adopted by the European Group on Fracture [45]. The MMB data reduction was based on the work of Reeder and Crews [41].

3.4.1 Double Cantilever Beam Tests

During the DCB experiments, the continuous load versus displacement record was generated and stored in a data file on the personal computer as discussed above. In addition to this data, it was also necessary to measure the crack length at various times during the experiment. The load/displacement data and the crack length measurements were then used to determine G_I.

Because the DCB test provided for stable fracture, several G_{Ic} values could be obtained for a test, each associated with a different crack length. In fact, most DCB global data reduction procedures utilize the compliance versus crack length trend to provide a correction to the G_I values.

During an experiment, the displacement between the end blocks was slowly increased, until noticeable crack growth occurred. At this point, the displacement was increased further, in such a way so that the delamination would extend approximately 5–7 mm. After the delamination had been extended, it was necessary to wait several minutes for the delamination to stop growing. Because the specimen was simultaneously interrogated via the moiré interferometry fringes, it was possible to detect very slight delamination growth.

Once the delamination growth was complete, the specimen was unloaded. At this time, the new crack length was determined by using a magnifying glass and marking the surface of the specimen with a sharp pencil. A thin layer of typewriter correction fluid painted on the edge of the specimen before the test facilitated the location of the delamination fronts.

This procedure of loading and unloading was repeated a number of times, producing a load deflection plot similar to the one shown in Figure 3.15. After the completed test, a hardcopy of this plot would be generated, so that compliance measurements could be made for each of the loading slopes. Other relevant specimen dimensions such as the thickness and the width were also measured at the end of a test.

Three different methods were used to convert the compliance, crack length measurements, critical load and the corresponding load point displacement into G_{Ic} values. These methods were the area method, Berry's method and the modified beam theory method. The latter two provided the most reliable results. The experimental compliance, or Berry's method was chosen as the benchmark global measurement.

The theory of Berry's method and the modified beam theory method were discussed in detail in Sections 2.1.3 and 2.1.5. Below only the actual implementation of these methods is described.

Area Method

Following the definition of the energy release rate (Equation 1.1), G_I can be obtained directly from a graphical interpretation of the load displacement plot. The change of potential energy during crack propagation is represented by the shaded area OAA' of Figure 3.16.

The energy release rate was recovered as

$$G_{Ic} = \frac{P_{ci}\delta_{fi} - P_{fi}\delta_{ci}}{2b(a_{fi} - a_{ci})}$$

where a_{ci} is the crack length at the critical point (just before propagation), a_{fi} is the crack length at the end of propagation, and b is the specimen width. The values of P_{ci}, δ_{fi}, P_{fi} and δ_{ci} were obtained directly from the hardcopy of the load/deflection plot and used with the crack lengths to compute G_I for each loading/unloading cycle. The area method has the disadvantage of being sensitive to errors in the crack length measurements.

Berry's Method

Berry's method provides a more accurate analysis of the experimental data, and involves a correction factor based on the trend of compliance versus crack length.

The first step with Berry's method is to determine the compliance as a function of crack length. This was done by simply fitting a straight line to each loading cycle of the load deflection plot, and extracting the slope. These values were then used in a log-log plot of the compliance versus the crack length. A least squares fit to these data then provided the correction factor n, as discussed in Section 2.1.3.

The final step was to tabulate G_{Ic} values as a function of crack length using Equation 2.2. As with the area method, P_c and δ_c were extracted from the hardcopy of the load displacement plot.

Corrected Beam Theory Method

In a similar way, the modified beam theory method utilizes a correction factor to correct the raw beam theory expressions in determining the energy release rate.

The modified beam theory method makes use of a plot of the cube root of

the compliance versus the crack length to determine its correction factor. The correction factor, Δ, (using the notation of Section 2.1.5) was obtained as the intercept of this curve with the x-axis. Equation 2.5 was then used to tabulate values of G_{Ic} versus crack length.

Experimental values for E_{xx} were also determined at each crack length using the expression

$$E_{xx} = \frac{8(a+\Delta)^3}{Cbh^3}$$

where b is the specimen width, h is half the height of the specimen and C is the measured specimen compliance. These values provided a cross-check on the results.

3.4.2 End Notched Flexure Tests

Unlike the DCB test, the ENF test does not provide stable crack growth. Thus when the delamination begins to grow, it will continue until it reaches the middle of the specimen. For this reason, conventional ENF tests produce only one critical energy release rate, for one crack length.

Another difference between the ENF test and the DCB test is that the correction factor used for reducing the ENF test data is based on an independent set of experiments performed before the actual delamination test. Prior to the actual delamination experiment, a compliance calibration/crack length calibration was performed.

The compliance calibration requires the compliance of the specimen to be measured for a range of crack lengths. This was accomplished by first recording the load/displacement plot at one crack length, and then sliding the specimen in the fixture, as illustrated in Figure 3.17, and then recording a new load/displacement plot. After repeating this procedure several times, a plot of the compliance versus the cube of the crack length yielded the desired correction factor, m, as the slope. The compliance calibration procedure was performed *prior* to the precracking of the specimen, so that the crack length could be known more precisely.

After performing the compliance calibration, the specimen would be precracked and repositioned in the test rig at the desired crack length. The delamination test would then be performed in much the same way as the DCB test, except that there would be no unloading of the specimen.

After the experiment, the critical load was determined from a hardcopy of the load versus displacement plot, and used with the measured specimen dimensions in the equation $G_{IIc} = 3ma^2P_c^2/2b$ (Equation 2.6), to determine G_{IIc}.

3.4.3 Mixed Mode Bending Tests

The testing procedure used for the MMB experiments was very similar to those discussed above for the DCB and ENF experiments. However, due to the complicated nature of the MMB test, there do not exist any reliable correction factors that can be determined empirically. Both the correction factors and the energy release rate values are based on expressions involving specimen material properties.

The simple beam theory equations for G_I and G_{II} were given in Section 2.3.4 as

$$G_I = \frac{3a^2P^2}{4b^2h^3L^2E_{11}}(3c - L)^2$$

and

$$G_{II} = \frac{9a^2P^2}{16b^2h^3L^2E_{11}}(c + L)^2.$$

When the correction factors are included, the following expressions are obtained

$$G_I = \frac{3P^2(3c - L)^2}{4b^2h^3L^2E_{11}}\left[a^2 + \frac{2a}{\lambda} + \frac{1}{\lambda^2} + \frac{h^2E_{11}}{10G_{12}}\right]$$

and

$$G_{II} = \frac{9P^2(c + L)^2}{16b^2h^3L^2E_{11}}\left[a^2 + \frac{0.2h^2E_{11}}{G_{12}}\right].$$

where λ and k are defined as before in Section 2.3.4.

During the MMB tests, load/displacement plots were recorded in the same way as the DCB tests. Hardcopies of these plots were then used to determine the critical loads and by using the above equations, values of G, G_I and G_{II} were obtained.

3.5 Some Practical Experimental Considerations

3.5.1 Precise Crack Length Determination

During the experiments and the precracking procedure, it was necessary to monitor the precise location of the delamination front. This was accomplished by two methods.

The first method was based on applying a thin layer of typewriter correction fluid on the thin edge of the specimen to increase the visibility of the delamination front. A magnifying glass (×10) was then used to locate the delamination front at the edge of the specimen.

The main advantage of this method was that it could be performed during an experiment, while moiré measurements were made on the opposite edge. However, this method could only be used to determine the delamination position *at the edge*, so no information about the shape of the delamination was available.

Ultrasonic

NDE Scans of Crack Front Profile

A more sophisticated method for determining the location and shape of the delamination was to perform nondestructive examination (NDE) of the specimen. This method provided a highly reliable two-dimensional map of the delaminated surface within minutes.

However, the NDE scans required that the specimens be submerged in water during the mapping process, so this method could not be used during an experiment, or after the grating had been introduced on the specimen. Regardless, the NDE scans played a crucial role in the development and verification of the fatigue precracking procedure discussed below.

3.5.2 Precracking Procedures

Due to the relatively thick starter films used in the manufacture of these specimens, and the sensitivity of delamination to crack front geometry, it was necessary to precrack all types of specimens prior to testing. The precracking involved two steps—first, mode I precracks were introduced by forcing the arms of the specimen open, then mode II precracks were propagated through fatigue loading in three point bend.

Forced Wedge Precracking

The initial mode I precrack was used as a starter crack, since the teflon inserts left a significant notch after moulding. To prevent unstable delamination growth, the specimens were clamped across their width, ~1 mm ahead of the artificial delamination, prior to the insertion of the wedge.

To insure that the initial mode I precracks were normal to the edges of the specimens, typewriter correction fluid was applied to both edges of the specimen, so that the delamination front could be monitored at the two free edge positions. The clamping and wedging procedure was was repeated until the detected delamination fronts were within approximately 0.5 mm of being normal to the edges of the specimen.

Ultrasonic NDE of the wedge precracked specimens revealed that the mode I precracks were generally straight and normal to the edge of the specimen. Those that deviated significantly from this condition were rejected.

Once a relatively straight mode I delamination was obtained, the specimens were submitted for mode II fatigue crack growth.

Fatigue Precracking

Fatigue precracking of the specimens provided a stable procedure for growing delaminations with natural crack front shapes. This was important since the moiré method was dependent on the full displacement field surrounding the delamination at the free edge. These displacement fields are believed to be affected by the three-dimensional stress state *within* the material, although the precise nature and the extent of this relationship is unclear.

A three point bend rig was designed for precracking the specimens. A schematic illustration of the rig is shown in Figure 3.18. The rig was designed so that the specimen could be pushed (conventional three point bend arrangement) and pulled. Thus twice as many cycles could be performed, and the specimens experienced equal amount of compression and extension throughout.

The appropriate number of cycles and the maximum deflection during the fatigue procedure were determined iteratively by performing occasional NDE scans of the fatigued specimens. The 'typical' case involved a crack length of 25 mm, a deflection of ±7 mm and approximately 500 dual direction cycles.

Each specimen was fatigued, and NDE scanned to verify the quality of the initial delamination. Those that were found to deviate from an accepted standard, were rejected.

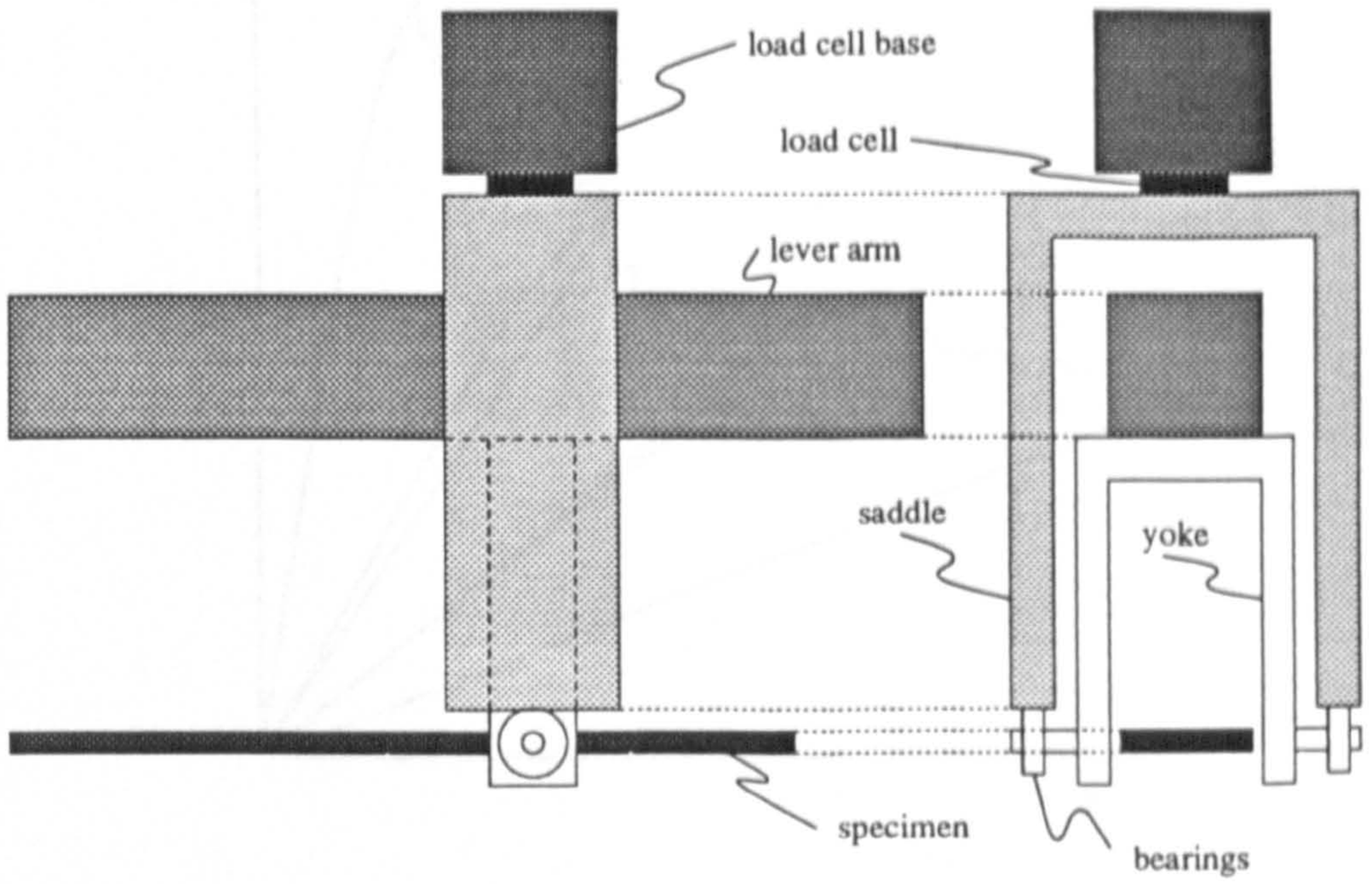


Figure 3.13: Schematic of the yoke load application used in this work.

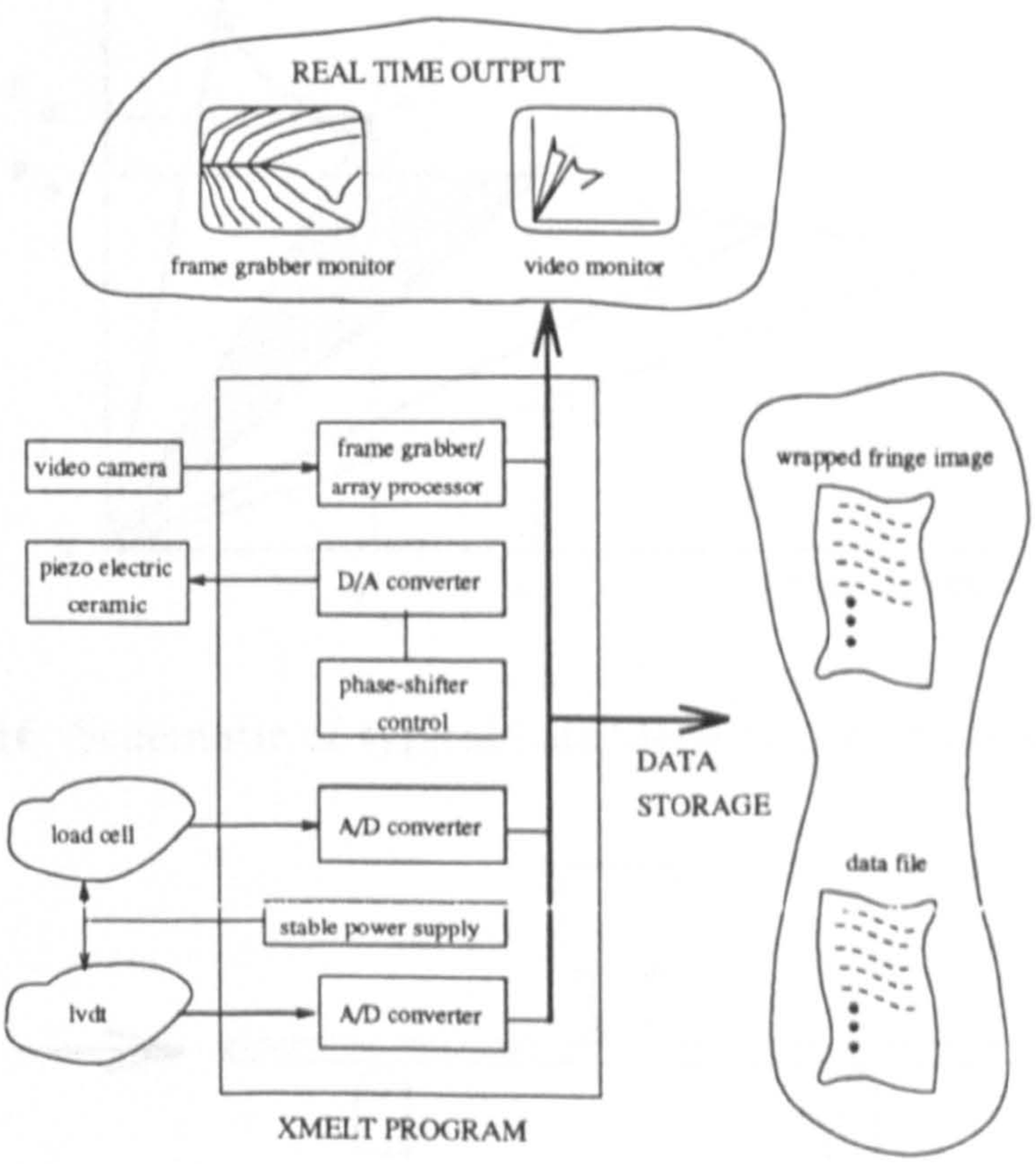


Figure 3.14: The XMELT data acquisition system

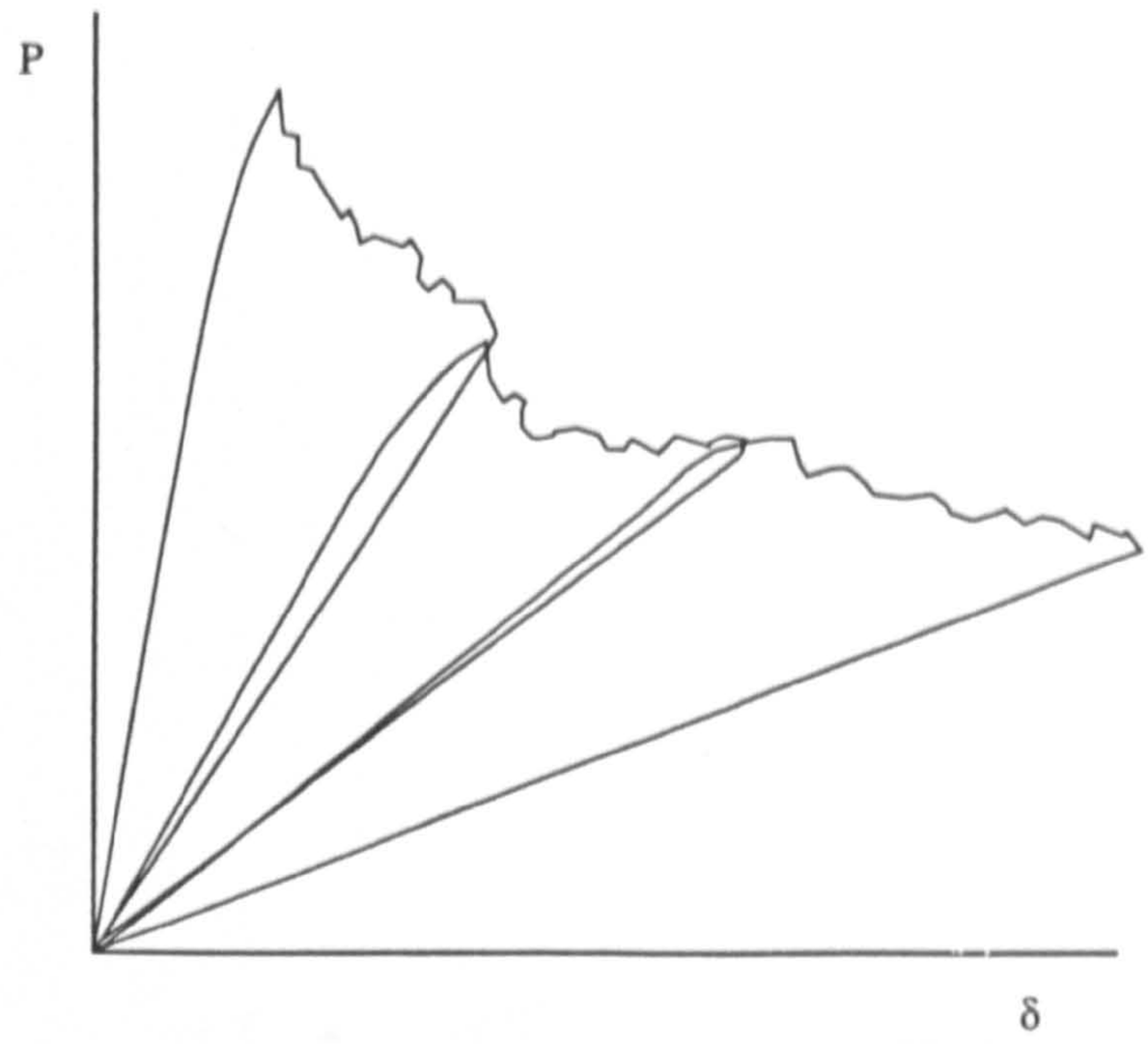


Figure 3.15: A typical load/deflection plot from a DCB test.

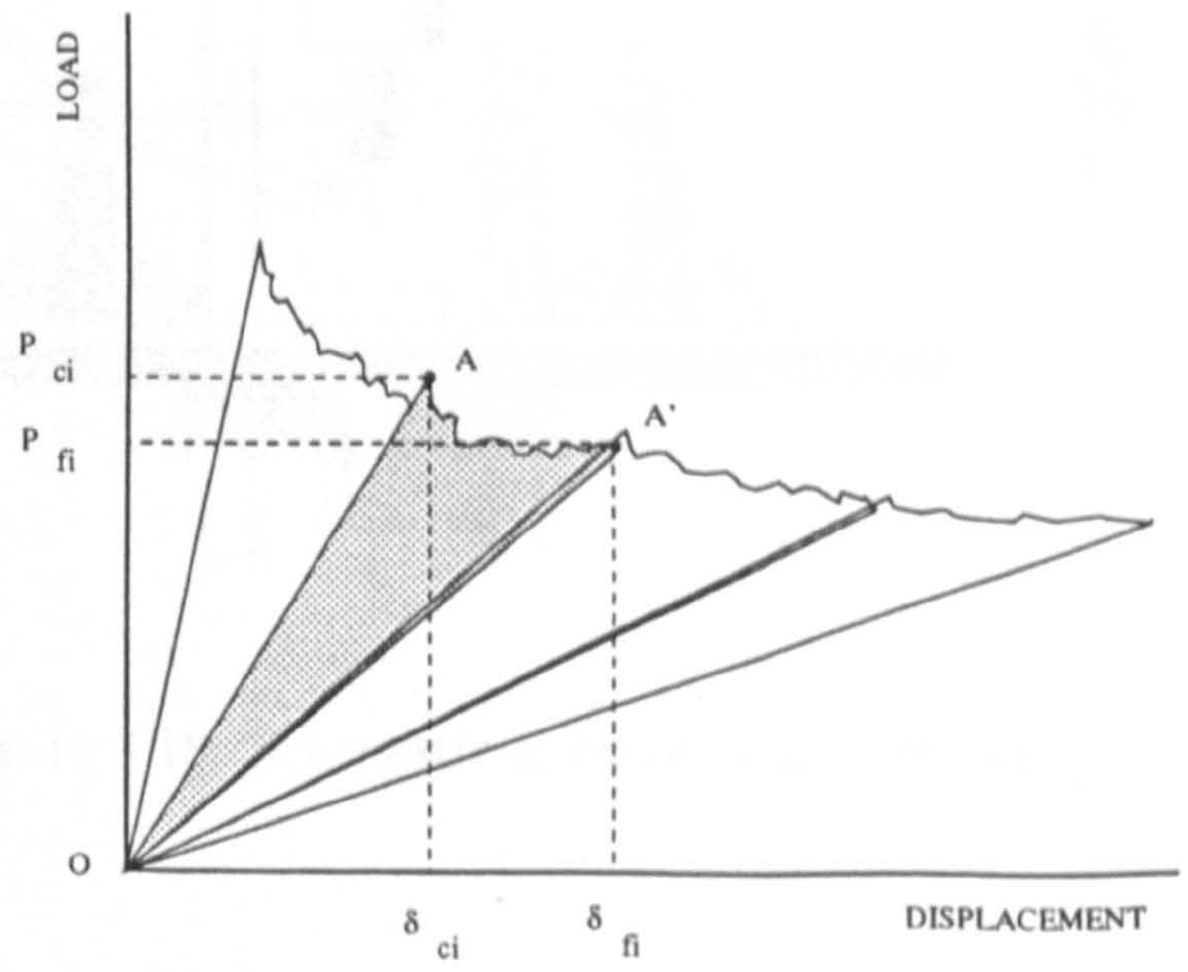


Figure 3.16: Schematic of typical load/deflection plot for a DCB test.

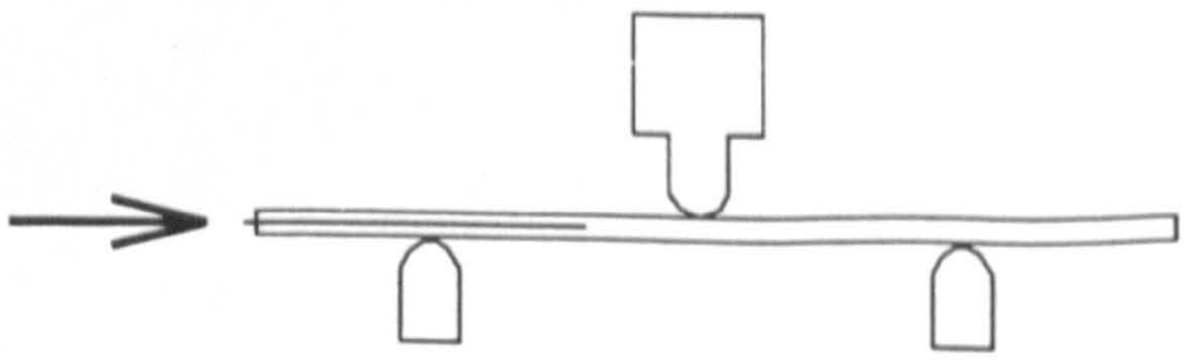


Figure 3.17: Procedure used to vary the effective crack length of an ENF specimen.

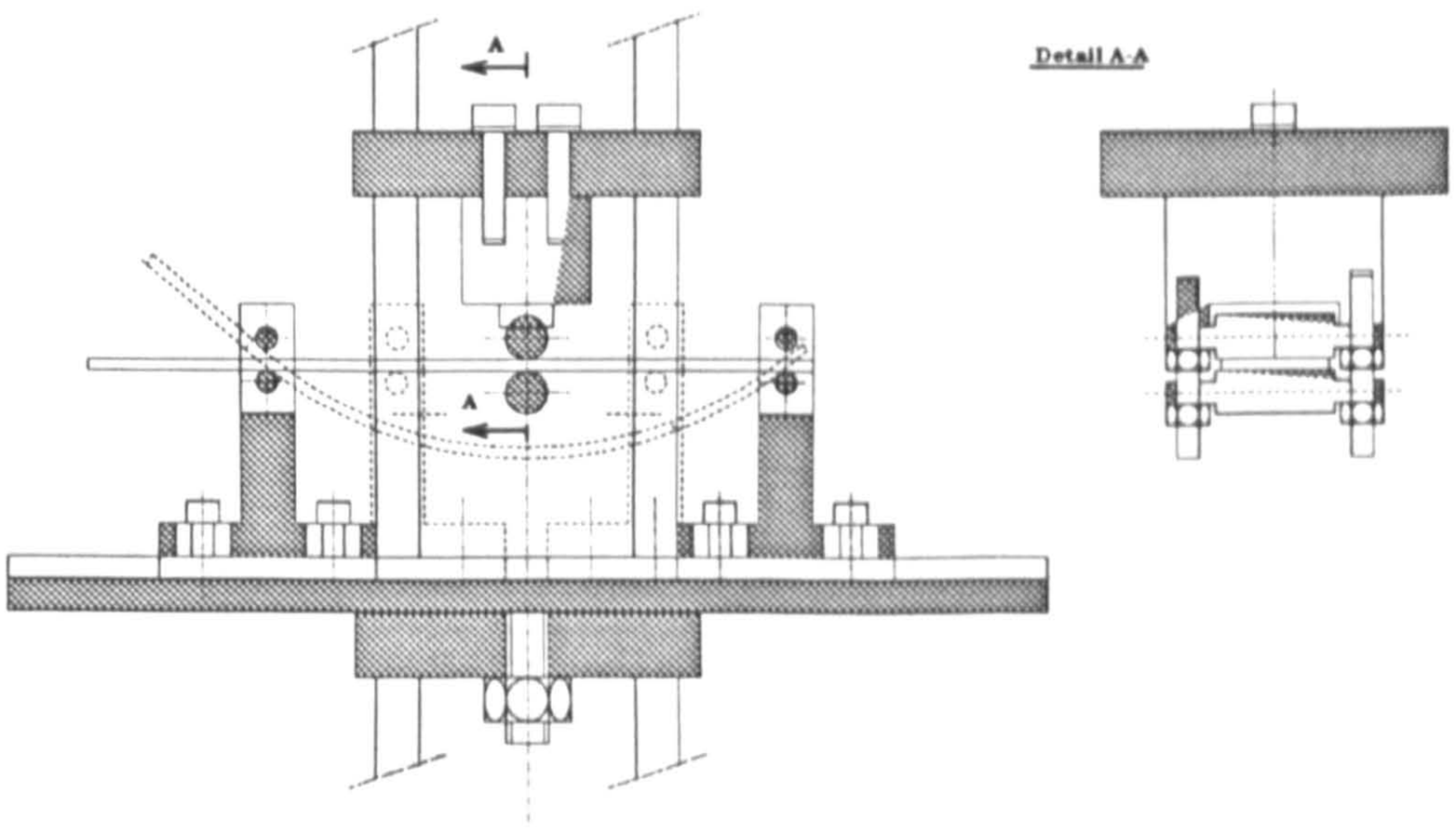


Figure 3.18: Schematic of three point bend fatigue rig.

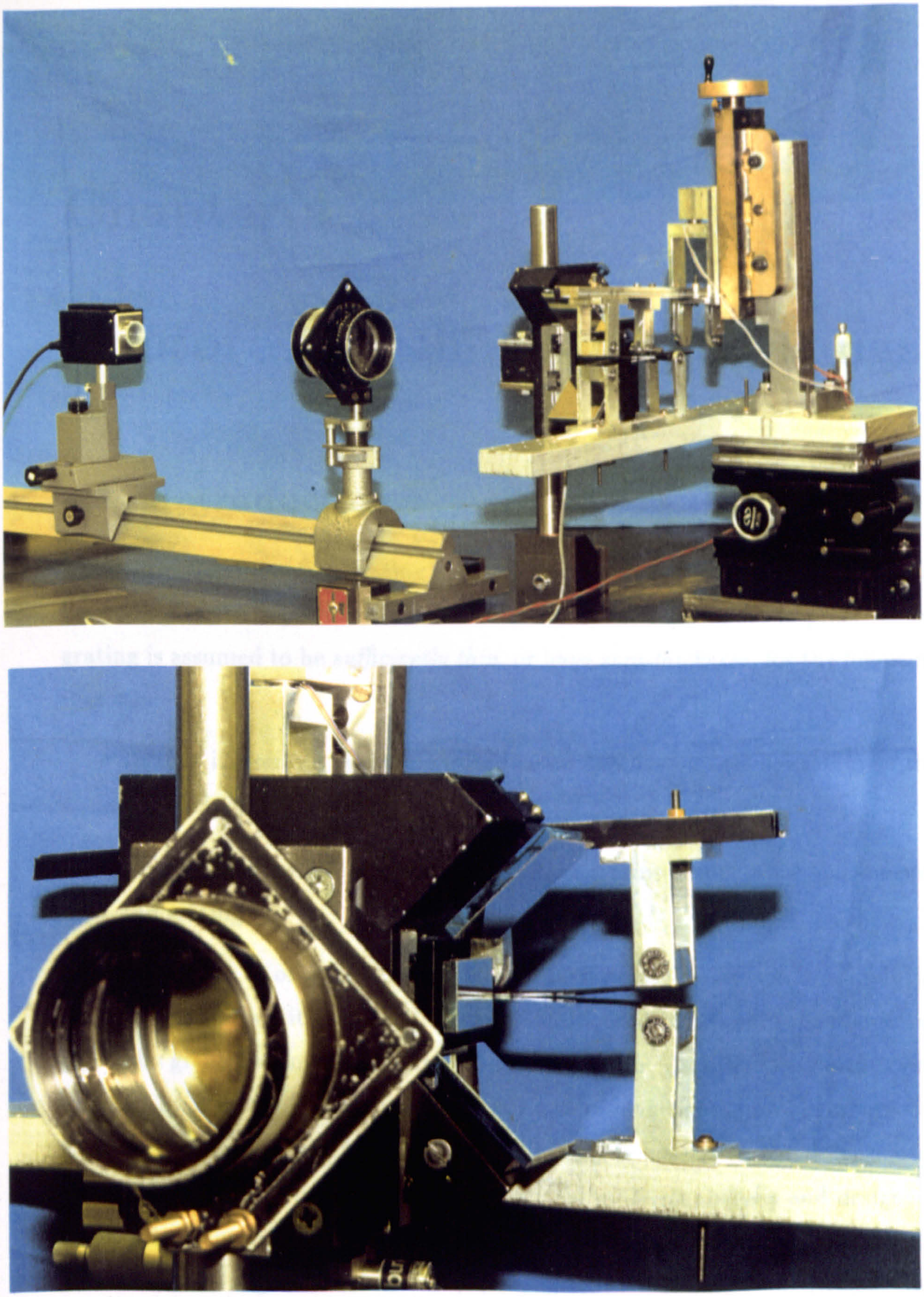


Figure 3.19: Photographs of the experimental apparatus.

Chapter 4

Photoresist Diffraction Gratings

4.1 Introduction

The role of the diffraction grating in moiré interferometry is to provide an optical 'en-coding' of the in-plane surface strains of a deformed specimen. In practice, the grating is assumed to be *sufficiently thin*, or have *zero-thickness*, for the following reasons:

- so that it exerts a negligible influence on the mechanical behaviour of the a specimen
- to eliminate an effective 'smearing-out' of strains through the thickness of the grating

The latter item is of most concern when investigating heterogeneous material systems on the constituent level. Conventionally used moiré diffraction gratings (typically greater than 20 μm thick), do not satisfy the 'zero-thickness' constraint imposed by a growing number of micromechanics investigations. Therefore, the need to explore new grating fabrication methods is crucial to the extension of moiré interferometry into the realm of advanced material analysis and design.

In this chapter, a technique for producing photoresist diffraction gratings directly onto the edges of carbon-fibre/epoxy specimens is described. Because the delamination behaviour of these specimens is to be investigated, the gratings are to be bonded to the thin edge of the specimens as shown in Figure 4.1. Many of the details and comments in this chapter are specific to the problems and difficulties associated with this type of specimen and geometry, but the information

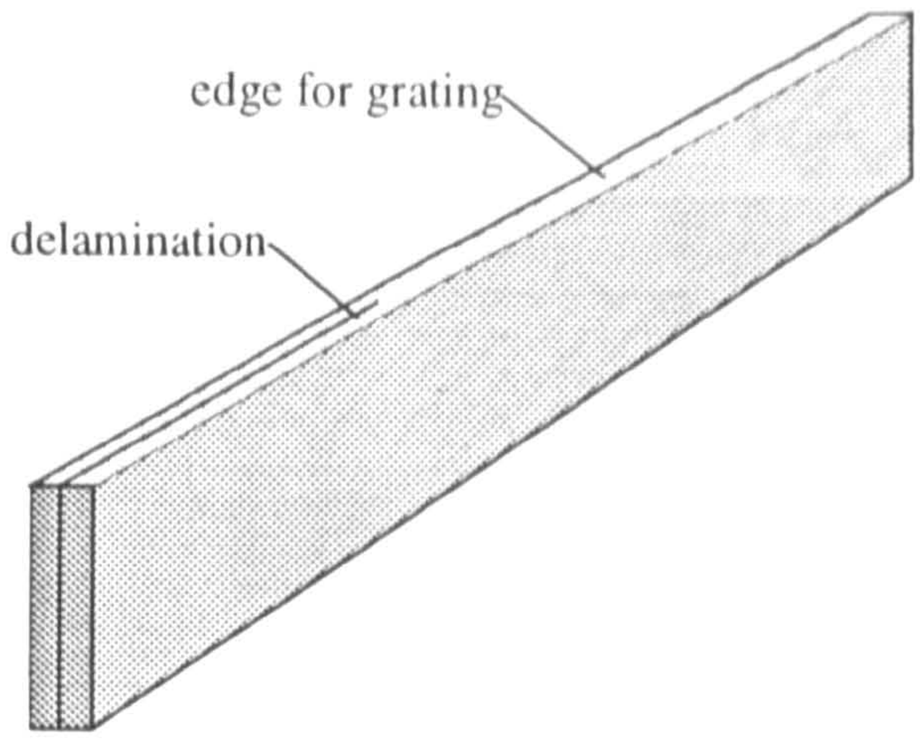


Figure 4.1: Geometry of carbon-fibre/epoxy specimens.

provided could still be useful for applying this technology to other materials and geometries.

Before discussing the photoresist diffraction grating technique, the conventional methods used to fabricate diffraction gratings is discussed.

4.2 Conventional Diffraction Grating Replication

Conventional diffraction gratings used in moiré interferometry investigations are generally transferred to the surface of the specimen by a moulding process. Uncured epoxy is prepared and poured onto the specimen surface, and the 'master' grating is placed in close contact with specimen. After curing of the epoxy , the 'master' grating is removed while the diffracting profile remains on the specimen surface. This process is illustrated in Figure 4.2

While the technique of grating replication is quite straightforward, there are several subtle aspects that should be mentioned. The preparation of the epoxy must included a means for removing air trapped in the liquid, so that small air bubbles do not remain in the final epoxy layer. Naturally, both the 'master' grating, and the specimen surface should be clean and free of grease and dust.

There are two advantages of the epoxy replication method:

- specimens can be prepared relatively quickly and reliably

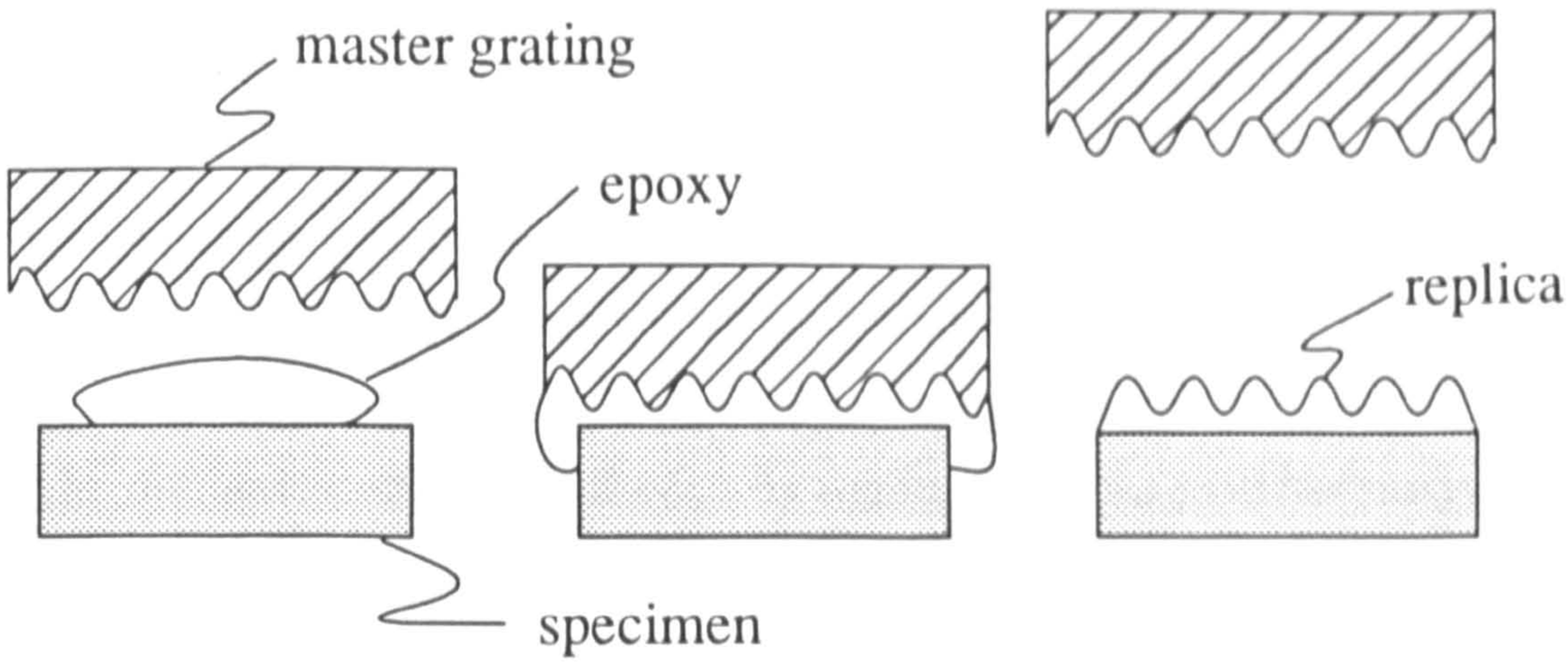


Figure 4.2: Conventional epoxy replication of diffraction gratings.

- the thin layer of epoxy acts to 'fill in' surface irregularities, such as scratches, etc.

Some types of specimens, such as carbon-fibre/epoxy composites, are very difficult to polish, and the last item above provides a means for eliminating surface irregularities that can not be eliminated by polishing.

The main disadvantage of the epoxy replication method is that the thin epoxy layer has a finite thickness. If care is not taken to insure that the layer is thin, then the thickness of the epoxy may obscure local variations in displacements. Usually, a small weight is used to press the 'master' grating in close contact with the specimen surface during the curing process. Care must be taken, however, because if too large a weight is used, the epoxy will be squeezed too thin, and the grating will not replicate properly.

The epoxy replication process can be used to replicate any type of 'master' grating onto a specimen surface. There are two common methods for producing the 'master' grating, however, and these are discussed below. Details about the replication procedures necessary for these types of diffraction gratings are also included.

4.2.1 Silicon Rubber Master Grating

One practical scheme for replicating diffraction gratings onto specimens involves using a 'master' grating made of silicon rubber. Provided the silicon rubber master grating does not become contaminated or damaged, it can be reused,

indefinitely, to replicate diffraction gratings onto new specimen surfaces.

Because of the nature of silicon rubber, no release agent is needed and the specimen can be removed easily after the epoxy has cured. However, it is not advisable to use even small weights to decrease the thickness of the epoxy layer when replicating a silicon rubber grating, since the weights will deform the 'master' grating profile.

The diffraction efficiency afforded by replicating silicon rubber 'master' gratings is generally a function of the quality of the 'master' grating. Because very efficient silicon rubber gratings can be used, this means that high diffraction efficiencies can be obtained with this technique.

The diffraction efficiency can be drastically improved by vacuum depositing a thin film of aluminium on the surface of the replicated epoxy layer.

4.2.2 Holographically Exposed Photographic Plates

Another method for providing 'master' gratings involves producing gratings on a one-by-one basis by holographically exposing photographic plates. The exposure process is similar to that described below for photoresist gratings.

Generally, a large number of photographic plates are exposed to a pair of collimated and interfering laser beams and developed according to the manufacturer's specification. By rotating the photographic plate through 90° during exposure, crossed diffraction gratings are produced.

After the normal photographic development procedure, it is important to introduce a release agent, so that the 'master' grating can be removed without tearing or lifting the epoxy from the surface of the specimen. The commercially available Kodak photoflo provides a good release agent when it is diluted and used to coat the 'master' grating prior to replication.

After the 'master' grating has been coated with a release agent, a thin film of aluminium should be vacuum deposited on the surface of the 'master' grating. The procedure discussed above for epoxy replication of the 'master' grating can then be used to transfer the grating profile to the specimen.

The primary advantage of holographically exposing photographic plates to produce diffraction gratings is that there is little chance of introducing contamination such as dust, etc., since each 'master' grating is used only one time.

Unfortunately, the diffraction efficiency obtainable with holographically exposed photographic plates is quite low, compared to other methods of producing diffraction gratings.

4.3 Photofabrication

Photofabrication is a technique generally concerned with replicating stencilled patterns on the surface of materials. Applications of this technique include electronic circuit fabrication, chemical milling of steel parts, and the making of printing cylinders.

Photoresist is a unique polymeric liquid that plays an important role in high precision photofabrication applications. Industrial procedures involving the manufacturing of micro-electronics devices generally utilize photoresist because of its unique photochemical properties.

The electronics industry is by far the most active in photofabrication technology, particularly in high resolution applications. Producers of integrated circuits and thin-film circuits demand reliable methods of producing extremely intricate, micron scale patterns on wafer substrates.

The generalized process used to fabricate micro-electronics circuits is illustrated in Figure 4.3, and is summarised below:

a) a mask is positioned above a substrate coated with photoresist, and illuminated

b) the exposed resist is removed during a development process

c) the metal substrate is etched with acid, leaving behind the desired image of the mask

It is not necessary to provide an etching process when making photoresist diffraction gratings—but in some circumstances, it may be desired. For instance, when working with metallic substrates, a post-etch performed in an ion mill can significantly enhance the diffraction efficiency.

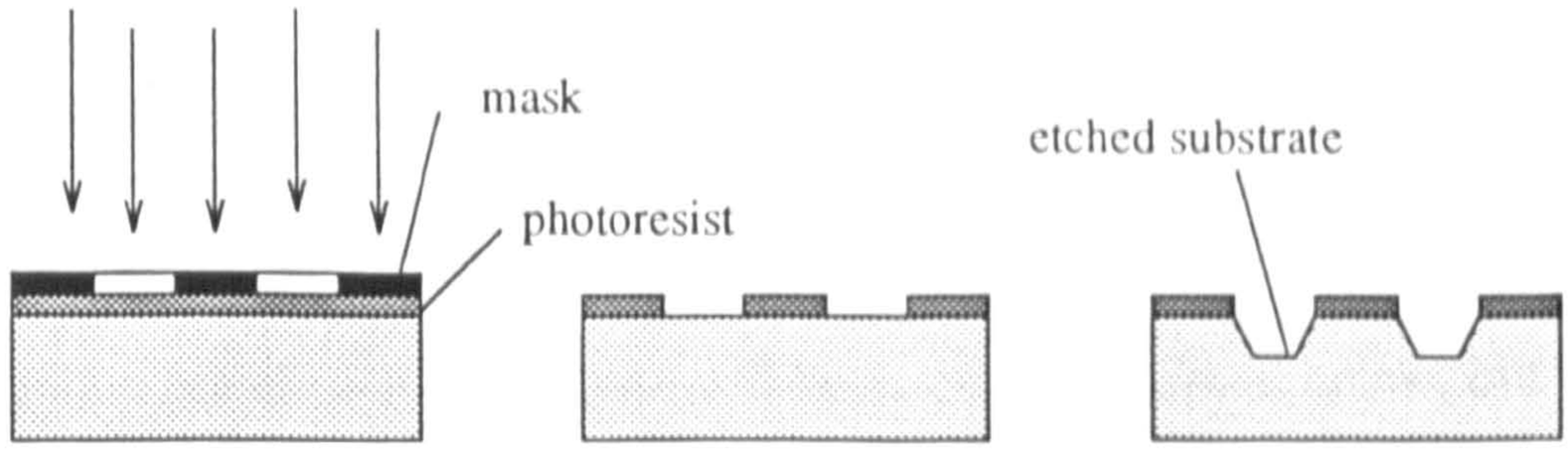


Figure 4.3: Micro-electronic circuit fabrication.

4.4 Making Photoresist Gratings

The making of photoresist diffraction gratings is not a difficult task, but like other grating fabrication techniques, it requires special equipment, experience and a bit of patience. The following discussion attempts to provide a general guide to establishing the facilities and procedures needed to implement the method.

4.4.1 Laser and Interferometer

The most important item needed to make resist gratings is obviously a laser. Several types of lasers can be used for exposing photoresist, but both the wavelength and the power of a laser must be considered. Photoresist is only sensitive to a particular region of the electromagnetic spectrum. The wavelength of a laser must agree with the exposure characteristics of the resist—not just any light can expose the resist. Furthermore, long exposure times may be necessary if one does not have a laser that provides a *sufficient* amount of *suitable* radiation.

This characteristic of photoresist translates into both a blessing and a curse. The 'blessing' is that one does not need to work *in the dark* with the resist, such as is necessary with standard photo-sensitive materials. Ordinary room light has little effect on exposing the resist (within reason, of course). At the same time, this 'lack of sensitivity' results in longer exposure times.

Photoresists vary, but most general purpose types have an exposure threshold of 100–150 $\mathrm{mJ/cm^2}$. However, several factors have a strong influence on actual exposure times, such as:

- thickness of photoresist layer

- relative concentration of solvent prior to exposure
- agitation during development
- optical absorbance/reflectance of specimen surface

The significance of each item above will be discussed in detail below, but the list should indicate that actual exposure times can be expected to vary. The energy threshold estimate above provides a useful guide for selecting a laser, but actual exposure times will have a wide range about this value.

For this work, research with photoresist diffraction gratings has been performed exclusively with a Helium Cadmium (HeCd) laser, which has a nominal power output of 14 mW, and a wavelength of 442 nm. This particular laser has interchangeable optics which allows for operation in the UV, ($\lambda = 325$ nm), where photoresist is most sensitive. One disadvantage of HeCd lasers, however, is that they have a fairly short coherence length (approximately 10 cm).

Coherence problems do not arise provided one carefully designs the exposure interferometer to minimize path length differences. In addition, an interferometer used in the exposure of photoresist must be maximized with regards to stability. During an exposure, the entire system should be completely enclosed to further reduce localized thermal fluctuations. Exposure times generally fall into the range of 15–20 minutes for a standard linear grating. The difficulty with such long exposure times is maintaining a vibration free optical table and designing an interferometer with negligible drift. Obviously, a more powerful laser would alleviate these problems. For future work, it is recommended to regard laser power as a top priority.

4.4.2 Photoresist

There are two types of photoresist—negative and positive. In negative resist systems, the exposed areas of the resist remain on the surface after development; and in positive systems, the exposed areas will be removed from the specimen during development. For diffraction grating fabrication, the positive resist proves the most convenient.

There are several specifications used by manufactures to describe a particular photoresist:

- spectral sensitivity
- viscosity
- contrast data or exposure characteristics

In addition to these, it is also common to see refractive index and purity data.

Photoresist is typically sensitive in the 'blue' region of the spectrum, but a plot of *absorbance* versus *wavelength* (or *spectral sensitivity*) is needed to insure compatibility between a resist and a laser. A typical spectral sensitivity plot is shown in Figure 4.4.

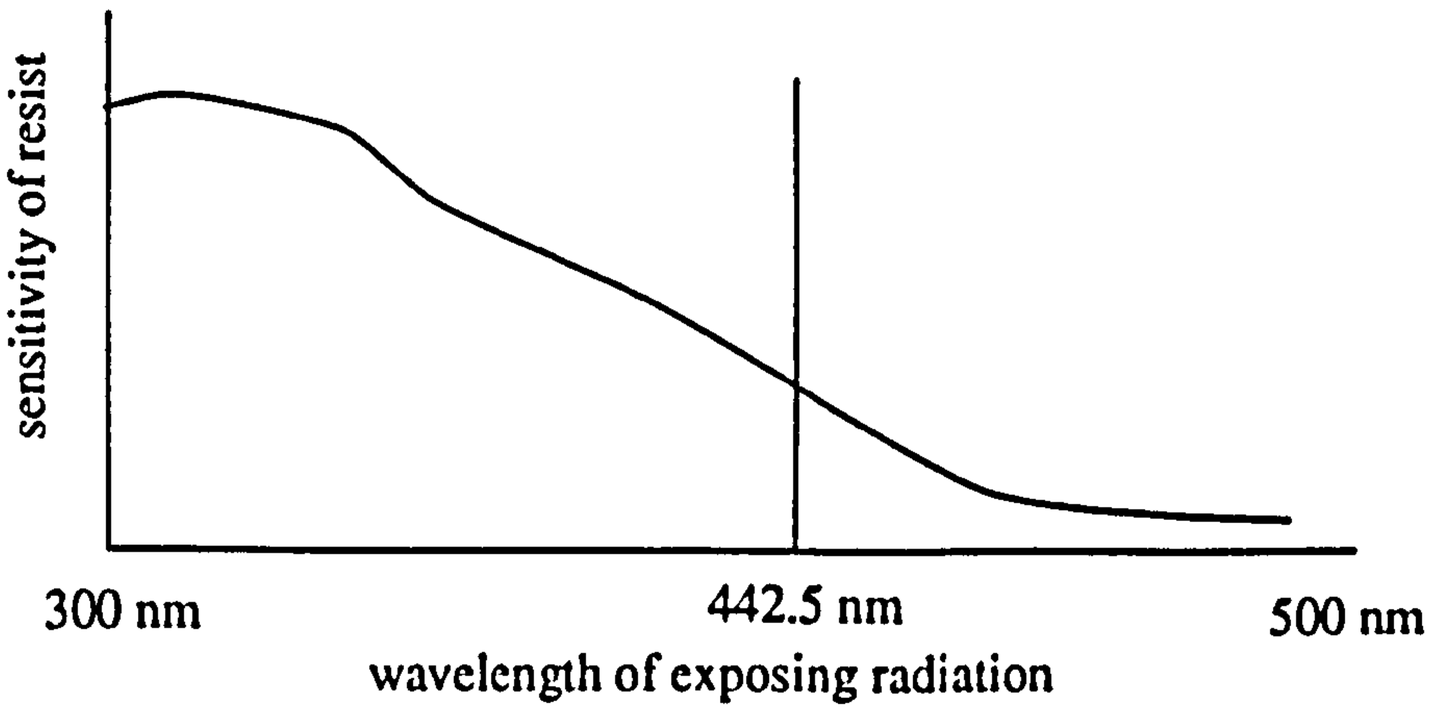


Figure 4.4: Generalized spectral sensitivity of photoresist.

Most photoresist suppliers offer a range of resists that are distinguished by their relative viscosities. This is due to the fact that viscosity plays an important role in determining the final resist thickness when applied by spin or dip coating.

An interesting and useful fact about photoresist is that the rate of exposure is distinctly non-linear. Figure 4.5 is a typical semi-log plot of *percentage exposed resist* versus *exposure energy*. This non-linear behaviour suggests that a pre-exposure process could be included to 'prime' the photoresist prior to interferometric exposure. The actual holographic exposure time can then be reduced without a corresponding loss of contrast.

Some suppliers market specific resists tailored for the different coating methods. In particular, an anti-orange-peel-effect resist is available for application by air-spray. A much improved coating can be obtained by using a resist tailored for a particular coating method.

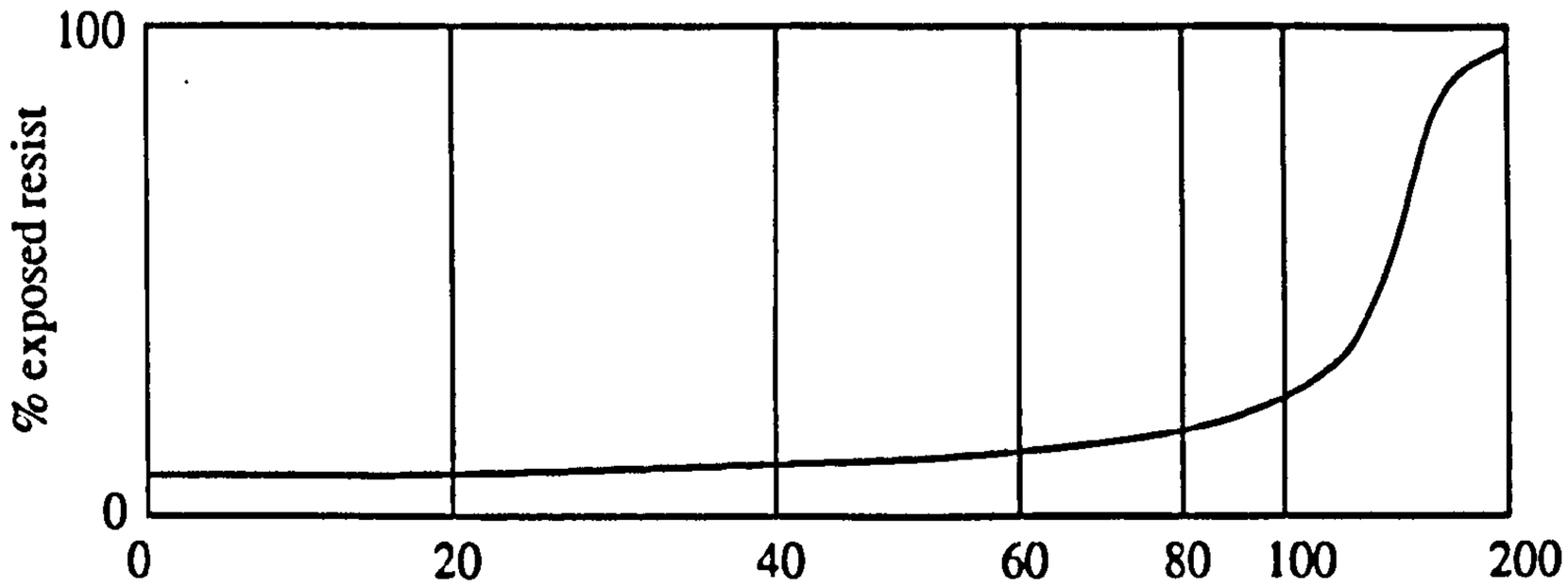


Figure 4.5: Non-linear exposure threshold or contrast data.

It is also necessary to consider the possibility of chemically damaging a specimen during grating fabrication. It should be verified that the mechanical properties of a specimen are not significantly changed during the photoresist process.

There is one chemical associated with photoresist products that is considered to be a health hazard: propylene glycol monoethyl ether acetate. Most conscientious suppliers have replaced the above demon with a relatively less toxic cousin, propylene glycol mono*m*ethyl ether acetate. This latter chemical is not associated with the adverse blood and reproductive effects as the former.

4.4.3 Specimen Surface Preparation

The most important quality variable in the entire photoresist process is cleanliness. If a clean room facility is available, this will significantly reduce contamination. If there is no such facility, then one must go to great lengths to insure grating quality.

Dust particles will inevitably find their way onto freshly sprayed specimen surfaces. Due to a surface tension effect, a potentially devastating thickness variation will result. Meanwhile, general contamination of the specimen surface can lead to adhesion problems. Scumming of the diffraction surface is another result of poor cleaning and insufficient rinsing. Traces of extraneous chemicals during any stage of the procedure, can lead to adverse surface reactions that will result in a 'cloudy' or 'scummy' film remaining on the resist.

Apart from cleanliness, there exists a more serious problem related to the specimen surface. If one is attempting to produce 'zero thickness' gratings, then

the surface on which the grating is to be applied must be absolutely smooth. As the resist dries, it attempts to maintain the same film thickness over the entire surface. Thus, the resist will follow the roughness of the specimen surface, and there will be no filling of gaps or rough spots.

Diffuse surface diffraction gratings can be used in moiré interferometry, but their efficiencies will be poor due to the scattering of light. The effect of a rough diffraction surface can be thought as a gross smearing of the grating's point spread function. Instead of finding crisp diffracted spots in the focal plane of a moiré interferometer, one finds a 'patch' of light scattered about the central spot.

The surface roughness problem does not exist when the specimen can be properly polished. Very efficient gratings have been produced on polished ceramics, aluminium and glass. But some specimens are virtually impossible to polish, such as porous materials, or two phase materials where the separate phases vary significantly in hardness.

The final issue relating to surface preparation is adhesion. Photoresist is an organic compound, and its ability to adhere to a specimen will depend on both the chemical and physical nature of the specimen surface. Organic materials (such as epoxy based composites) provide an ideal surface, but metals and some inorganic materials can form oxide layers which hinder adhesion.

The problem of adhesion is encountered in the micro-electronics industry because the photoresist is generally applied directly to metal substrates. Therefore, most suppliers offer adhesion promoters and extensive guidelines for accommodating a wide variety of adhesion problems.

4.4.4 Application of the Resist

There are three general categories that summarize the various methods for applying photoresist: spinning, spraying and dipping. These categories are discussed in detail below. Several less relevant methods are also worth mentioning. A common process used for coating large quantities of printed circuit boards is roller coating. The specimens are fed into one end of a machine and pass under an applicator roller. This applicator roller is in turn fed with a precisely controlled quantity of resist, which it deposits onto the specimens as they pass. Extremely uniform coatings are obtained in this way, but it is not a practical method for

small quantities of specimens. Two more primitive methods are brush coating, and flow coating. Flow coating is where a small volume of resist is poured on to the specimen, which is tilted in various directions to allow the resist to cover the surface. The excess resist is then drained off. Both the brush and flow coating methods are not recommended because they produce non-uniform and inconsistent coatings.

Common to all coating methods, is the problem of edge effects. As the solvent evaporates, the surface tension of the photoresist equilibrates, resulting in a uniform film thickness over the surface of the specimen. However, near the edges of a specimen, a meniscus will form, resulting in an edge-bead. Likewise, other discontinuities (including dust particles) will tend to disturb the film thickness variation. The affected areas are generally quite localized, extending no more than approximately 0.1 mm. In cases where the edges are far removed from the areas of interest, there will be no noticeable effect.

In general, the liquid resist must be applied to the substrate to form a uniform, continuous coating. This is essential to achieve consistent photographic properties. Thin coatings generally give higher resolution while thick coatings provide more chemical resistance during an etching process.

When making photoresist diffraction gratings, the best strategy is to make the coating as thin as possible. The final thickness of the grating is not critical (as it is when one performs an etching process). The quality of the final grating will be determined by its efficiency, (which will be a function of the thickness of the resist), but it is not possible to predict what thickness will give the best results *a priori*.

Spin-coating: Spinning methods provide the most reliability, in both small and large quantity production. This is the preferred method to use, provided the specimens to be coated are not too large and have a suitable aspect ratio.

The specimen must be fastened to a turntable, which is to be rotated at 3000–6000 rpm. Resist is poured from the middle of the specimen out towards the edge. The faster the rotation, the thinner the final coating. Coatings made this way can be as thin as 0.30 μm, with a tolerance of $\pm$0.10 μm. The only variables that must be controlled are spin rate and resist viscosity. These factors are relatively easy to maintain, which is why this method has such high repeatability.

Industrial uses of photoresist, such as micro-electronic circuit fabrication, rely primarily on the spinning technique. The only equipment needed for this method is a rotating wheel, and a shield to catch the excess resist. However, this method can not be used if the specimens to be coated are too large to spin properly (or safely).

Apart from safety concerns, it is not recommended to coat areas larger than approximately 20 cm^2. Areas larger than this tend to exhibit non-uniform coating profiles. Also, if the specimen has an awkward aspect ratio (such as the case with the edges of delamination specimens), a uniform coating will again, be difficult to achieve.

Spray-coating: This method is particularly suited to awkward specimen sizes. Coatings made in this way can be as thin as those made by spin-coating, but the variability is much larger.

The resist normally requires thinning with a proper solvent before spray-coating. Acetone will suffice, but it is advisable to use the manufacturer's recommended solvent. Correct dilution must be found by trial and error, and will depend on several factors: desired coating thickness, specimen surface porosity, and the type of spray equipment to be used.

There is a range of conventional spray equipment that may be used for spray-coating photoresist. A simple and economic alternative is to use a standard artist's air-brush, but more sophisticated equipment could lead to a more automated process.

The variables to be considered when spray-coating are:

- air pressure
- nozzle size
- distance from nozzle to specimen
- dilution of the resist
- speed of traverse of air-brush over specimen

It is necessary to experiment with all of the above factors until the optimum resist thickness is achieved.

Despite the number of factors influencing coating thickness, very efficient gratings can be consistently produced using this method. By holding all but the 'speed of traverse of the air-brush' constant, it is not so difficult to develop an effective technique for coating specimens.

Dip-coating: The dip-coating method is useful in both small and large quantity productions. Consistent results can be obtained by dip-coating, but it is difficult to yield thin gratings. The specimens must be completely immersed in a tank of diluted photoresist. They are then withdrawn at a uniform controlled rate. The higher the speed of withdrawal, the *thinner* the coating of resist.

The withdrawal rate may be controlled hydraulically, pneumatically or electrically. Once the 'correct' withdrawal rate has been determined, this method will provide reasonably consistent coatings.

4.4.5 Drying the Resist

A proper drying procedure must be followed when working with photoresist. This is necessary to prevent degradation of the photoresist's exposure characteristics and insure proper adhesion of the coating to the substrate.

All photoresist is diluted with some amount of *solvent.* The addition of solvent to photoresist facilitates coating and proper drying. Several coating methods also require a further dilution of the resist prior to application. Prior to exposure, the solvent must be driven from the coating.

Inadequate removal of the solvent has two effects on the quality of the coating:

- the presence of even small quantities of solvent can reduce the photographic speed of the coating considerably

- the solvent is *soluble* in the developer: thus, its presence can lead to numerous development problems

Each of these problems considered alone will devastate the quality of the final grating. If the speed of the resist is reduced, excessive exposure time will be necessary. Solvent remaining in the coating during development will lead to severe adhesion problems and virtually no contrast.

More importantly, when excess solvent remains in the coating, it is impossible to determine correct exposure times and developer concentrations. No photoresist process can be made reliable unless this problem is eliminated.

The remedy is to provide a well controlled post-baking procedure. The coating should be allowed to partially dry in air at room temperature. Once the film is 'dry to the touch', further drying can proceed at elevated temperatures. It is important to gradually dry the resist, as a surface skin may develop and prevent the underlying solvent from evaporating.

During the post-bake, it is necessary to provide a well ventilated chamber for the specimens. Poor ventilation will retard the evaporation process. Care must also be taken to insure that the resist is not 'burned' during this stage. Photoresist will chemically break down above approximately 120°C.

4.4.6 Exposure and Development

As discussed earlier, exposure of the resist must be performed with a sufficient amount of suitable radiation. Due to long exposure times, an exceptionally stable interferometer is necessary to eliminate vibrations and fringe drift.

The exposure characteristics of a particular resist can be found by systematically varying exposure time. Due to the nature of photoresist, it is more realistic to conduct a three parameter investigation of the following variables:

- exposure time
- developer concentration and method
- development time

A complete study of the above factors was performed for this work, and the results are presented in Section 4.5.3.

Glass microscope slides can be used as generic specimens while the process is being developed. They are cheap, easy to clean and prepare, and allow simple estimation of grating thickness. But, there are several problems unique to microscope slides that limit their usefulness for determining exposure characteristics.

One of the problems with microscope slides, (or any transparent specimen) is multiple reflections. Multiple reflections cause the resist to be exposed in

unpredictable ways; dependent on coating thickness, slide thickness, angle of illumination, etc. To reduce these effects, one can paint the back of the specimen with an anti-reflective coating (black paint), or index match the specimen to a non-reflecting surface. Another problem with microscope slides is adhesion of the resist, but this can be overcome by providing a sufficient cleaning and drying procedure.

Once exposed, the specimen can be developed immediately. The procedure is extremely simple, and only requires a supply of sodium hydroxide and de-ionized water. The most important aspect of the process is providing agitation.

If the specimens are not agitated sufficiently during their development (or if they are grossly over-exposed), the diffraction surface will appear dull and misty. This is known as scumming. The reflectivity of the grating will be reduced, and the diffraction spots will not be sharp. A developing tank providing vigorous agitation, or a steady flow of developer over the specimens will reduce these problems. However, attempting to brush the surface of the developing resist will destroy the grating.

The specimens should be quenched in water at the end of the development stage. Several minutes in flowing water should be sufficient. It is generally best to gently blow the specimen surfaces dry with compressed air to reduce drying patterns on the diffraction grating.

At this stage, the specimens can be aluminized and used in a standard interferometer. The aluminization is not always necessary, but it does significantly enhance the signal-to- noise ratio of the grating.

4.5 Experimental Details

In this Section, more detailed information regarding the photoresist diffraction grating process is presented. The goal of this work was to define a standard procedure which would enable the fabrication of photoresist diffraction gratings on the edges of carbon-fibre/epoxy, as discussed in Section 4.1.

The procedure for making photoresist gratings is composed of three steps:

- Polishing and surface preparation
- Spray coating the specimens

- Exposure and development

The most difficult part of the process was the first item above. Providing an optically smooth surface on the edges of carbon-fibre/epoxy specimens was not a trivial task. The remaining two steps were fairly straightforward, although a bit of trial and error was necessary to arrive at the optimal processing conditions. Once determined, the entire process allowed for a wide degree of latitude.

4.5.1 Grinding and Polishing

Polishing carbon-fibre/epoxy specimens was difficult because there are two very different materials present: the epoxy, which is relatively soft and ductile, and the carbon-fibres, which are hard and brittle.

The presence of two materials during a grinding or polishing process obviously leads to a situation where the softer material is removed more readily than the harder one. The end result is that the softer material is preferentially removed, producing what is known as 'surface relief'.

The only solution to the problem of surface relief, was to avoid it by using grinding and polishing materials that have a 'low-knapp'. (The 'knapp' of a cloth refers to the height of its surface structure.) Soft polishing cloths generally have a very high knapp, and thus were avoided when polishing specimens that have inhomogeneous hardness properties. Figure 4.6 schematically illustrates the situation in both a high-knapp cloth, and a low-knapp micro-abrasive grinding material.

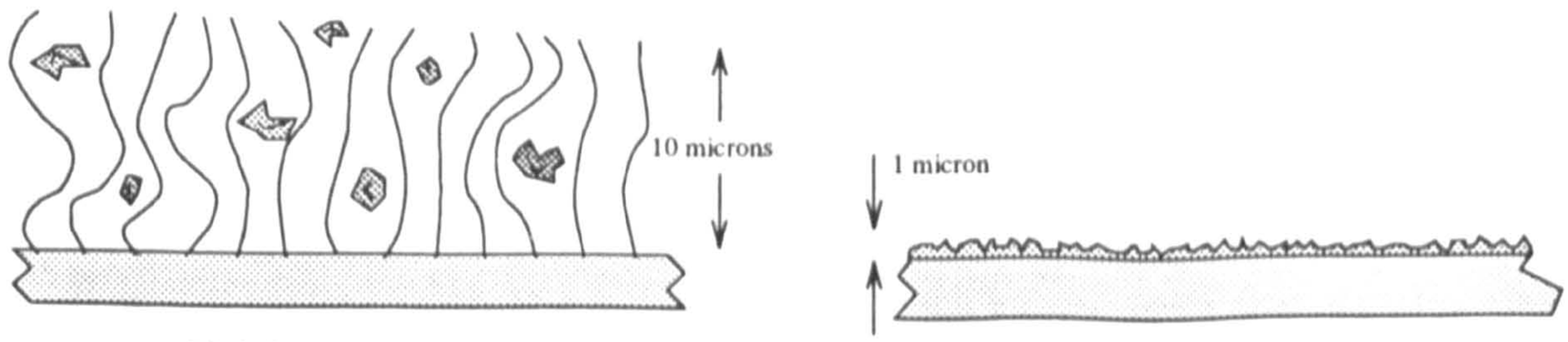


Figure 4.6: The details of the polishing process.

Materials such as silicon carbide grinding papers, and the micro-abrasives used in this work, have an intrinsically low-knapp structure, which greatly reduces relief polishing. There are two advantages offered by high-knapp cloth. First, the

knapp provides a place to 'hide' loose grains, keeping them from inadvertently scratching the surface of the specimen. Second, very fine cloth can be manufactured relatively inexpensively. Despite these advantages, high-knapp cloth should not be used when polishing carbon-fibre/epoxy specimens.

It is useful to regard the exercise of grinding or polishing as the controlled introduction of a consistent surface texture. The general idea is that one begins with a relatively coarse grinding paper, grinding the specimen until there remains no deviation in the surface of the specimen greater than that which is introduced by the paper itself. Then, after moving to the next grade of paper, the process is repeated. This is a common-sense procedure, but since the quality of the final gratings depends so strongly on this part of the process, it is necessary to consider it in such detail.

Grinding should begin with the highest grade of paper that is capable of reducing the texture of the surface to some uniform level. The presence of one large scratch in an important region would thus require starting with a relatively coarse paper. There are several disadvantages of starting on coarse paper, however. It obviously requires more work, since the papers must be used in the proper order. Furthermore, regardless of what paper one begins with, preferential gouging will always occur in the 0° layers. The severity of the gouging will increase with the amount of grinding.

Much effort can be wasted by not paying close attention to what is happening to the surface of the specimen as it is being ground and polished. Only the specimen surface itself can reveal when no further advantage can be obtained from a particular paper. Worse, prematurely moving to the next paper results in scratches remaining on the specimen surface that are nearly impossible to remove with the higher grade abrasives—a 1 μm micro-abrasive paper is powerless in the face of a 15 μm scratch.

It was therefore necessary to 'watch' the surface and verify that no abnormal scratches remain before moving to the next paper. An optical microscope was available for this purpose with a magnification of $\times 100$.

To facilitate grinding and polishing, a special jig was made to hold the specimens. It was found that working with five specimens at a time was optimal. A photograph of the holder is shown in Figure 4.8. It was crucial for the speci-

mens to be level prior to grinding; otherwise the process would have had to begin at a coarse grade of silicon carbide paper, in order to reduce the surface height variations of the specimens. Figure 4.7 illustrates the problem.

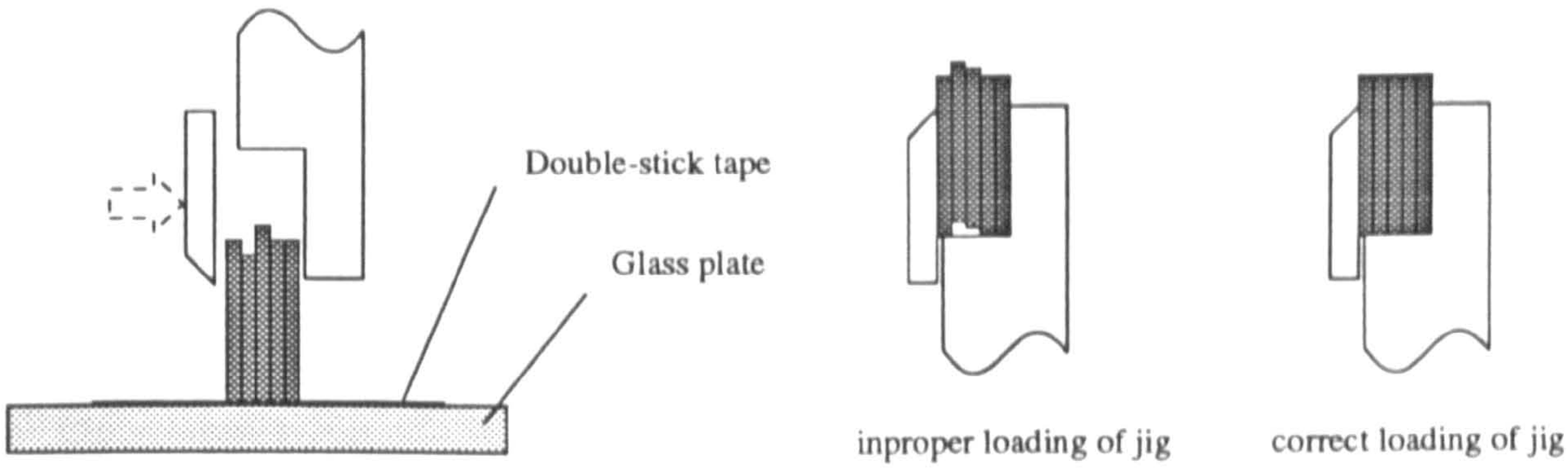


Figure 4.7: Correct loading of the grinding jig.

A procedure was developed for loading the specimens into the holder so that they would be level across the top surface. This procedure involved using double-stick tape and a smooth glass plate. First, the specimens were mounted to the glass plate using the double-stick tape, with the face to be polished facing down. The holder was then brought into place, and securely clamped around the specimens. Using this method, grinding began at the 1200-grit silicon carbide grinding paper.

This arrangement provided a wide, stable surface for the grinding and polishing procedure. Two 'blanks' were always used to flank the three central specimens, since it was not feasible to properly polish the specimens on the end. The specimens remained in this rig throughout the grinding and polishing process. This was necessary to preserve the level surface needed to polish the specimens uniformly. Unfortunately, this made it very difficult to thoroughly remove the loose grains that happen to become lodged between the specimens. It is obviously not desirable to carry coarse grains onto the finer grinding papers.

The manual grinding process was preceded by an automated surface grinding. The stone used in the surface grinder was approximately 600 grit, but due to the automation, it was possible to achieve a surface finish equivalent to that obtained with a 1200 grit silicon carbide paper. The use of the surface grinder not only saved much effort, but it also improved the quality of the final polished surfaces.

After grinding, the specimens were polished first using the micro-abrasives sheets, and then, a combination of the micro-abrasive sheets, conventional metallurgical polishing pads and diamond paste. The final polish was performed on an 8 inch metallurgical polishing wheel with 0.1 μm diamond paste.

After the surfaces had been properly polished, it was necessary to clean them thoroughly. A first rinse was given with tap water, using a small amount of soap, and a squirrel's hair paint brush to gently scrub the polished surface. A second rinse was given with an ultrasonic bath using distilled water. The specimens were dried with a towel except for the polished surface. Compressed air was used to blow the water off the polished surface, and the specimens were put aside to dry.

4.5.2 Spraying

Photoresist is applied to the specimens by spray coating. This procedure was surprisingly repeatable.

The spraying was done with a standard artists' airbrush. A photograph of the airbrush is shown in Figure 4.8. This particular model has an additional feature that allows the user to control the ratio of the liquid to air. This proved to be an invaluable asset, when applying the resist. A 3 μm needle/jet combination was used with the airbrush.

The spraying procedure was very simple, and it took only a little practice before optimal thickness coatings could be produced with a good degree of repeatability. For the specimens in this work, resist was thinned to 70%, and an air pressure of 18–20 psi was used. The resist was applied in one uniform, smooth stroke of the airbrush, with the nozzle at about 10 cm from the specimen surface.

After spraying, the specimens needed to dry. Prior to exposure, the solvent must be driven completely from the coating. This was a very important step in the process, since exposure characteristics are dependent on the relative concentration of resist and solvent. Furthermore, the solvent is soluble in the developer, and if it is present in the dried resist film, it will encourage uneven deterioration of the coating during development.

Specimens were first dried in a horizontal position at room temperature. This step allowed the resist to dry through the thickness, preventing a surface film from developing and trapping moisture below the surface. Once the specimen

was 'dry to the touch', further drying proceeded at approximately 60°, for at least 5 hours. The actual drying time was adjusted on a case by case basis.

4.5.3 Exposure and Development

The exposure was performed on an optical table, using a standard interferometric arrangement. A schematic of the interferometer is is shown in 4.9.

Before exposing the resist, the interferometer was aligned, and a null field generated by placing a suitable diffraction grating in the exposure plane. The fringes were verified to be stable and null. The contrast between the two beams was adjusted, using a variable density filter in one of the paths.

The laser used has 14 mW total power. A rough calculation estimates that the exposure energy at the exposure plate is approximately 0.06 mW/cm^2. This indicates that an exposure time of 20–30 minutes would be necessary. However, these calculations are approximate, due to the number of assumptions involved—in both estimating the power available, and the exposure characteristics of the resist. On average, each field of a crossed grating would require approximately 15-25 minutes.

Gratings were usually made on several specimens at once, and several 'dummy' specimens would be included. These dummy specimens allowed for establishing the ideal processing parameters, which varied slightly from batch to batch. Once the ideal parameters were determined, the batch of gratings could be processed identically, with a high degree of confidence. The parameters that were determined 'on the fly' were the exposure time and the development time and degree of development agitation.

After exposing the photoresist, the next step was to develop the specimens. This was a simple procedure that involved no more than immersing the specimens in a solution of sodium hydroxide. It was found that by performing the development in an ultrasonic bath, a greater degree of uniformity could be obtained. There was a strong relationship between exposure time, and the development procedure. The table below illustrates the relationship:

The concentration of the sodium hydroxide solution was another parameter that had to be optimized. Best results occurred when the developer was relatively strong (approaching 5%), the development time was short, and a strong degree

Exposure time	Development time and methods
25 minutes	1 minute 30s
20 minutes	25–40s (ultrasonic development)
15 minutes	30–50s (ultrasonic development)

Table 4.1: Typical exposure and developing times.

of agitation provided. Temperature was found to have a small effect on the development (warmer temperatures naturally increasing the development action), but no consistent increase in quality could be attributed to the temperature variable.

The typical parameters about which the process gravitated are given below:

- Developer concentration: 2%
- Development time: 40s
- ultrasonic bath on for 10s, off for 10s, on for 10s ...

The exact procedure for each batch of specimens was determined by closely observing the development of the dummy specimens.

The specimen were quenched in water at the end of development stage. Two minutes in flowing water was the standard. A final rinse was performed in the ultrasonic bath using distilled water. These rinses were very important to drive the developer solution off, otherwise the aluminium coating would be corroded by the sodium hydroxide.

The specimens were gently blown dry using a can of compressed air, to avoid drying patterns. The final step in the process was to vacuum deposit a thin layer of aluminium to enhance the reflectivity of the grating.

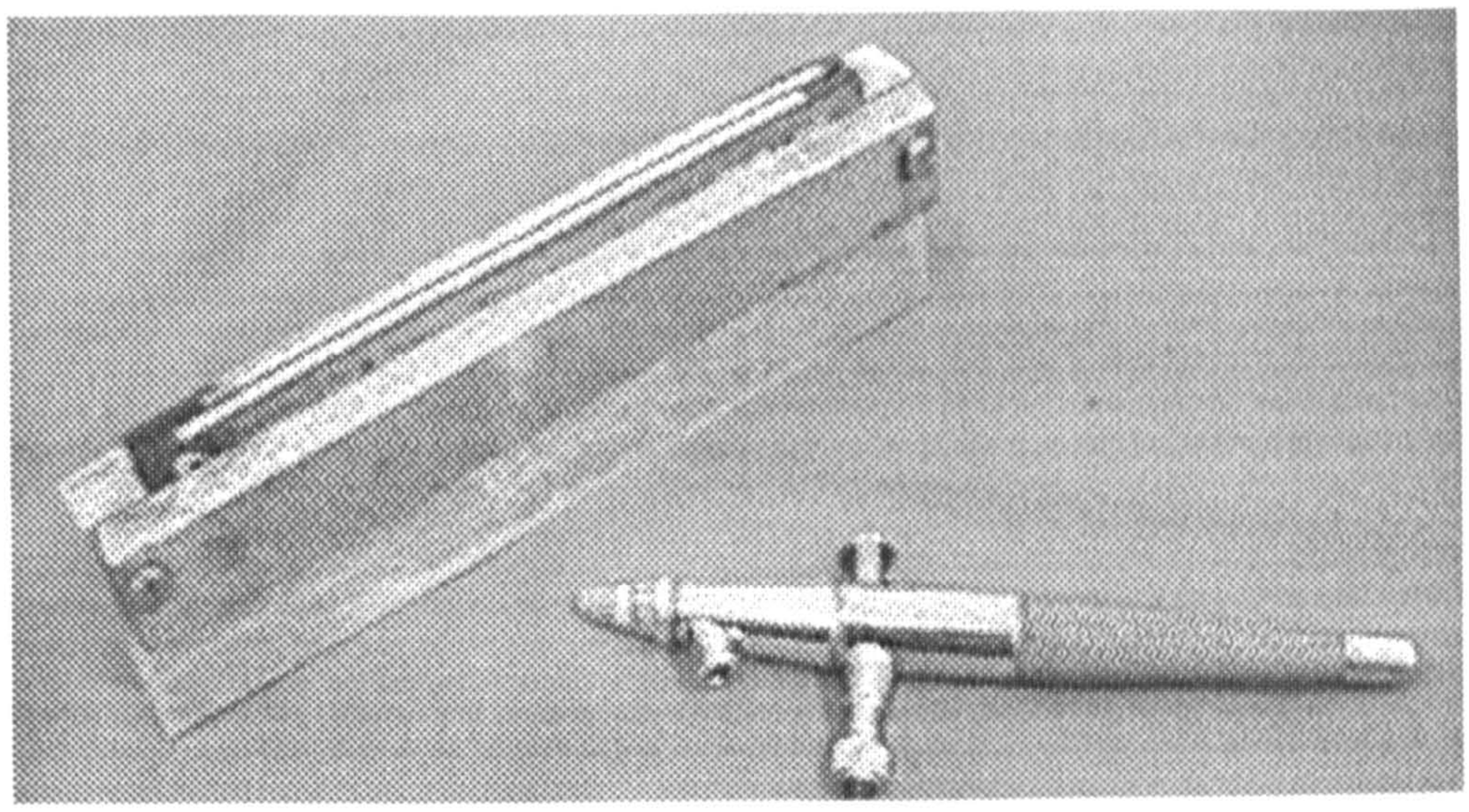


Figure 4.8: The airbrush and holder.

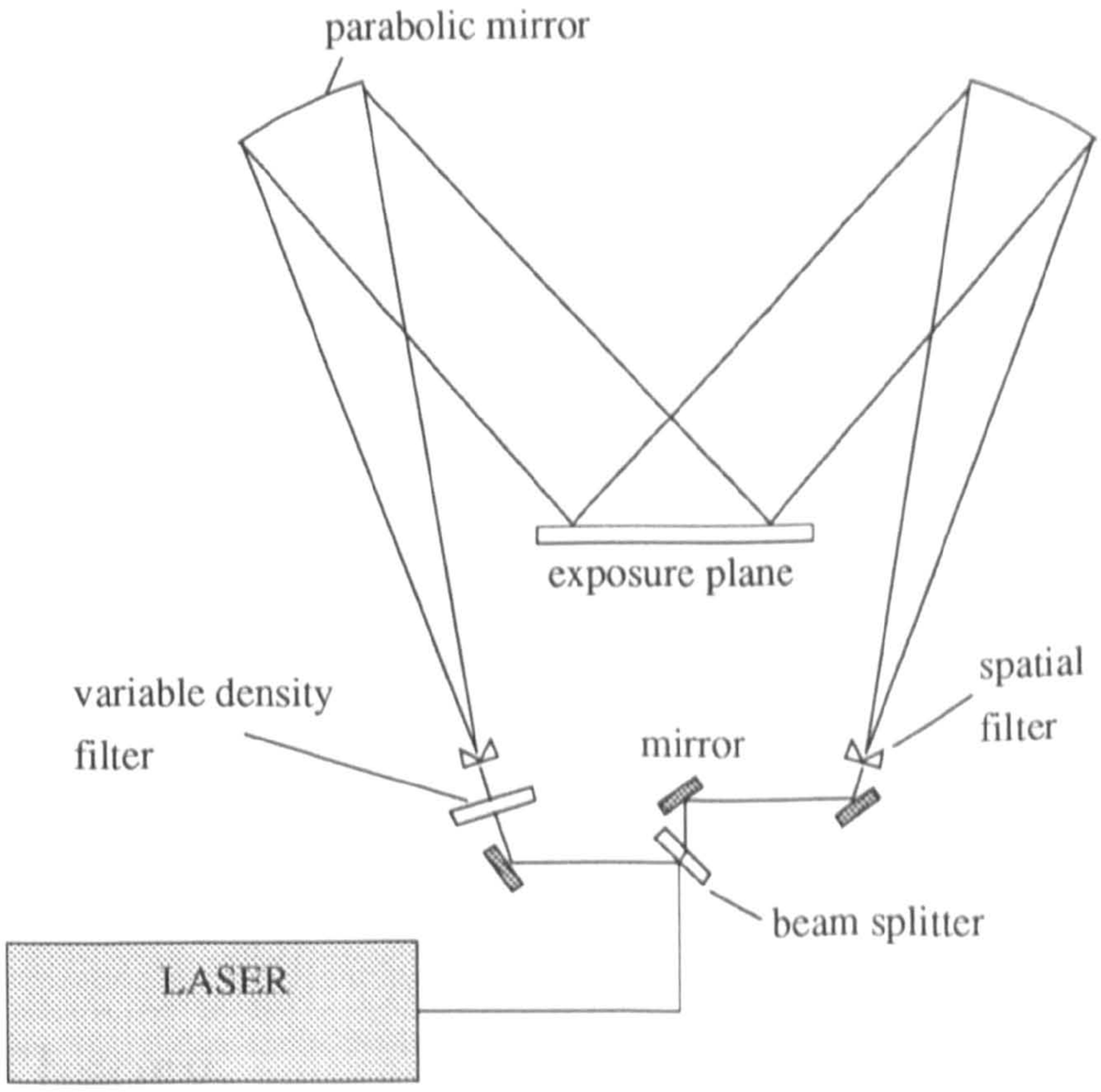


Figure 4.9: A schematic of the exposure interferometer.

Chapter 5

Phase Shifting Moiré Interferometry

5.1 Moiré Interferometry

Moiré interferometry is a high-sensitivity method of full-field displacement measurement. The technique directly provides contour maps of the two orthogonal in-plane displacement components, u and v. Strain information can be obtained from the displacement fields via differentiation, according to the usual approximations

$$\varepsilon_x = \frac{\partial u}{\partial x}, \quad \varepsilon_y = \frac{\partial v}{\partial y} \quad \text{and} \quad \gamma_{xy} = \frac{\partial u}{\partial y} + \frac{\partial v}{\partial x}.$$

Moiré interferometry offers an experimental counterpart to the powerful computational methods of solid mechanics, where displacements are also the primary output [7]. An important aspect of this similarity is that moiré interferometry is capable of providing a means to validate (or invalidate) computational schemes for describing complex mechanical behaviour.

5.2 Theory

Discussions of moiré interferometry generally begin with an illustration of the formation of geometric moiré fringes, as shown in Figure 5.1. The visual 'fringe' patterns produced by the superposition of the regular grids patterns in Figure 5.1 are manifestations of the moiré effect, or more precisely, the *geometric* moiré

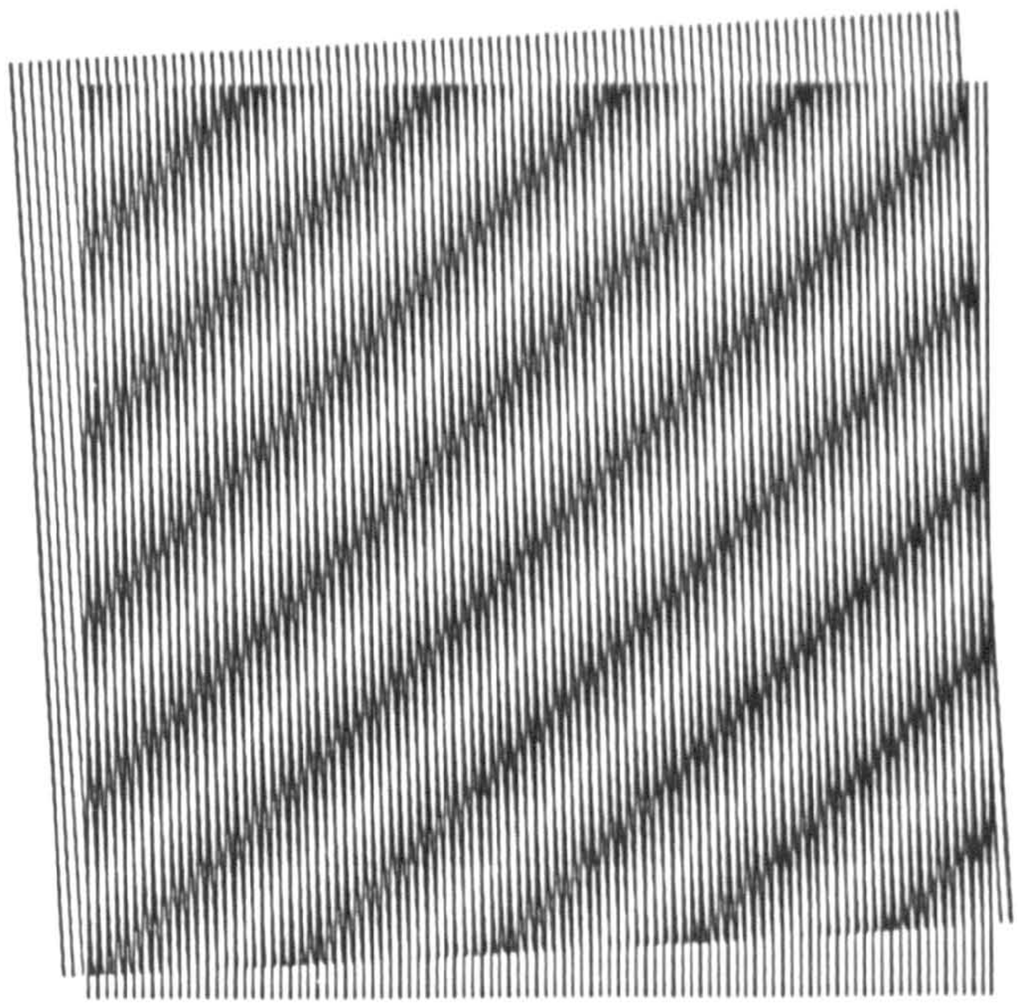


Figure 5.1: The geometric moiré effect.

effect.

It has long been recognized that the interference produced by the two sets of lines could be used to determine quantities relating to the relative extension and rotation between the pairs of lines [47]. The sensitivity of geometric moiré method is naturally increased as the line or grid pattern is made more dense. This provided the impetus for the development of moiré *interferometry*.

Instead of extending the geometric analogy to explain the phenomenon of moiré interferometry, an approach based more on the physics of the process will be adopted. The moiré interferometry fringe phenomenon is best understood by first considering two important manifestations of the wave nature of light—interference and diffraction.

5.2.1 Interference

The optical interference that occurs in moiré interferometry is called *two-beam* interference. Two-beam interference results from the combination of two independent beams of light under certain conditions—these conditions are discussed below. The effect can be constructive, meaning that the energies of the beams add together, producing a 'bright' spot, or destructive, meaning that the energy in the two beams subtracts from one another creating a 'dark' spot.

Interference is a vector phenomenon and depends on both the relative amplitudes of the two beams *and* their individual phases. This implies that the intensity distribution resulting from the interference of two beams will be determined as the vector sum of the individual components. There are several conditions that must be satisfied before two beams of light will interfere. These are summarized below.

- The two beams must have the same wavelength (if not, the fringe patterns will not be stationary)
- The beams must have a similar polarization
- The beams must be sufficiently *temporally* and *spatially* coherent
- If the intensities of the two beams are widely different, then the effect of the interference will be weak

If a conventional laser source is used, the first item is automatically satisfied. Light from most lasers is also usually *linearly polarized.* However, the polarization state of a beam can change when it is reflected off a surface such as a mirror. Care must be taken during the design of an interferometer to insure that all reflections do not inadvertently rotate the polarization of one of the beams.

Before discussing the concept of coherence, it is necessary to introduce the idea of an optical wave field. Assuming we have monochromatic and linearly polarized light, an optical wave field can be described by a real-valued scalar function of position and time [48]. This scalar function is either associated with the electric field or magnetic field vector and denoted by

$$f(\vec{r};t) = a(\vec{r})\cos[2\pi\nu_0 t - \phi(\vec{r})], \tag{5.1}$$

where $\vec{r}$ is a general position vector, t denotes time, and ν_0 is the temporal frequency of the wave. The phase, $\phi(\vec{r})$, and amplitude, $a(\vec{r})$, of the wave are, in general, both functions of position. The surfaces in space defined by the equation $\phi(\vec{r}) =$ constant are commonly called *wavefronts.*

The light emitted from a laser is said to be *collimated,* which means that the wavefronts are flat and normal to the direction of propagation (within the so-called diffraction limits). Wavefronts of this type are commonly called *plane*

wavefronts. Because the output from a laser is normally only 1–2 mm in diameter, it is common practice for a laser beam to be expanded and recollimated, so that a larger area can be illuminated. Figure 5.2 illustrates the idea of optical wavefronts by schematically demonstrating the process of expanding and collimating a laser beam.

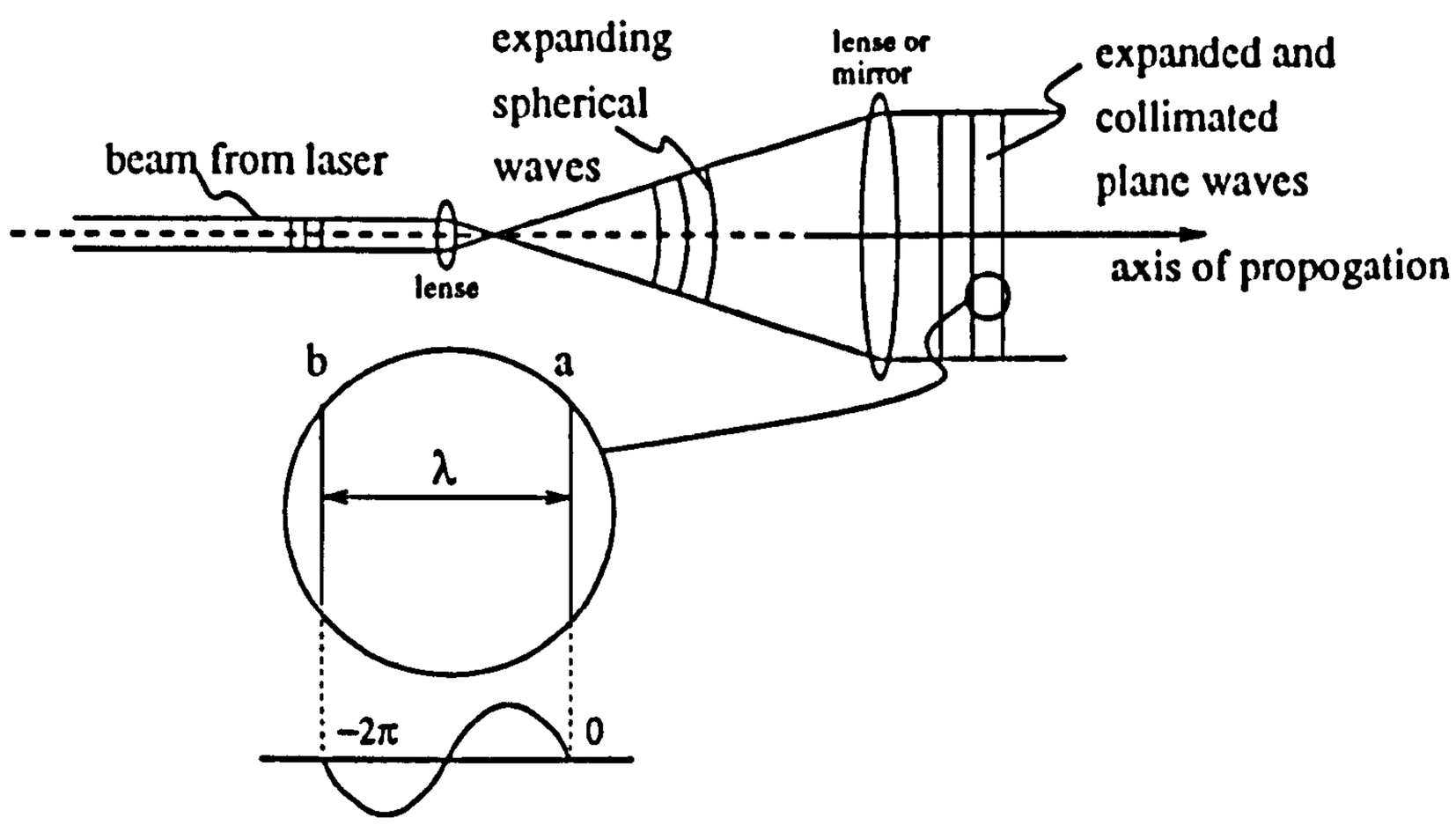


Figure 5.2: Expansion and recollimation of a laser beam, illustrating the idea of plane wavefronts.

Having introduced the idea of an optical wave field, it is now possible to discuss coherence. A distinction is generally made between *spatial coherence* and *temporal coherence* of an optical wave field. Consider the field at two separate observation points in an optical wave field, $f(\vec{r}_1;t)$ and $f(\vec{r}_2;t)$. If the phase difference between $f(\vec{r}_1;t)$ and $f(\vec{r}_2;t)$ is independent of time for all points $\vec{r}_1$ and $\vec{r}_2$, then the wave field is called spatially coherent. The wave field is said to be temporally coherent if the phase relationship between $f(\vec{r};t_1)$ and $f(\vec{r};t_2)$ depends only on the time difference $(t_2 - t_1)$ and not on t_1 or t_2 separately. Spatial coherence is a characteristic of a wave field transverse to the direction of propagation, while temporal coherence is a characteristic along the direction of propagation [48].

In practice, the condition of temporal coherence demands that the two rays must have travelled roughly the same distance from the source. The limiting distance that two beams from one laser source can differ by, and still interfere, is called the *coherence length* of the laser. It is therefore necessary to design an

interferometer in accordance with the coherence length of the laser used.

Temporal and spatial coherence are related to the optical phase of the beams. But optical interference is a vector phenomenon related to both phase and amplitude. So if the beams do not have similar amplitude, there will be an additional background intensity superimposed on the interference pattern due to the stronger beam. A fairly wide latitude is generally tolerated in two-beam interference, although the contrast of the interference pattern will suffer when the beams are not close in intensity.

When any of the conditions for interference discussed above are not perfectly satisfied, the result will be a reduction in contrast of the interference pattern. It is important to be aware of all of these factors, and make an attempt to optimize each during an experiment.

5.2.2 Diffraction

The phenomena of diffraction can be explained in a number of ways, for example, by classical wave theory or by a corpuscular quantum approach [49]. With respect to moiré interferometry, however, only the principle results are of importance.

Figure 5.3 illustrates what happens to a beam of light as it strikes a linear diffraction grating. The incoming beam is split into various orders, and the angles at which these new beams travel is governed by the following equation

$$m\lambda = d \sin \theta_m, \tag{5.2}$$

where λ is the wavelength of the incident light, d is the spacing between furrows on the grating, and the value of m denotes the various orders. The 0th order corresponds to the simple reflection of the incoming beam.

The amount of light split into each order depends on the detailed nature of the diffraction grating rulings. Generally, the intensity of a diffracted beam is strongest for ± 1 orders, and decreases quite rapidly for the higher orders.

In moiré interferometry, a crossed[1] diffraction grating is bonded to the specimen surface. The angle between the grating outward normal and the illuminating

[1] A *crossed* diffraction grating is composed of two sets of rulings that are orthogonal to one another. This allows for the interrogation of two independent displacement components.

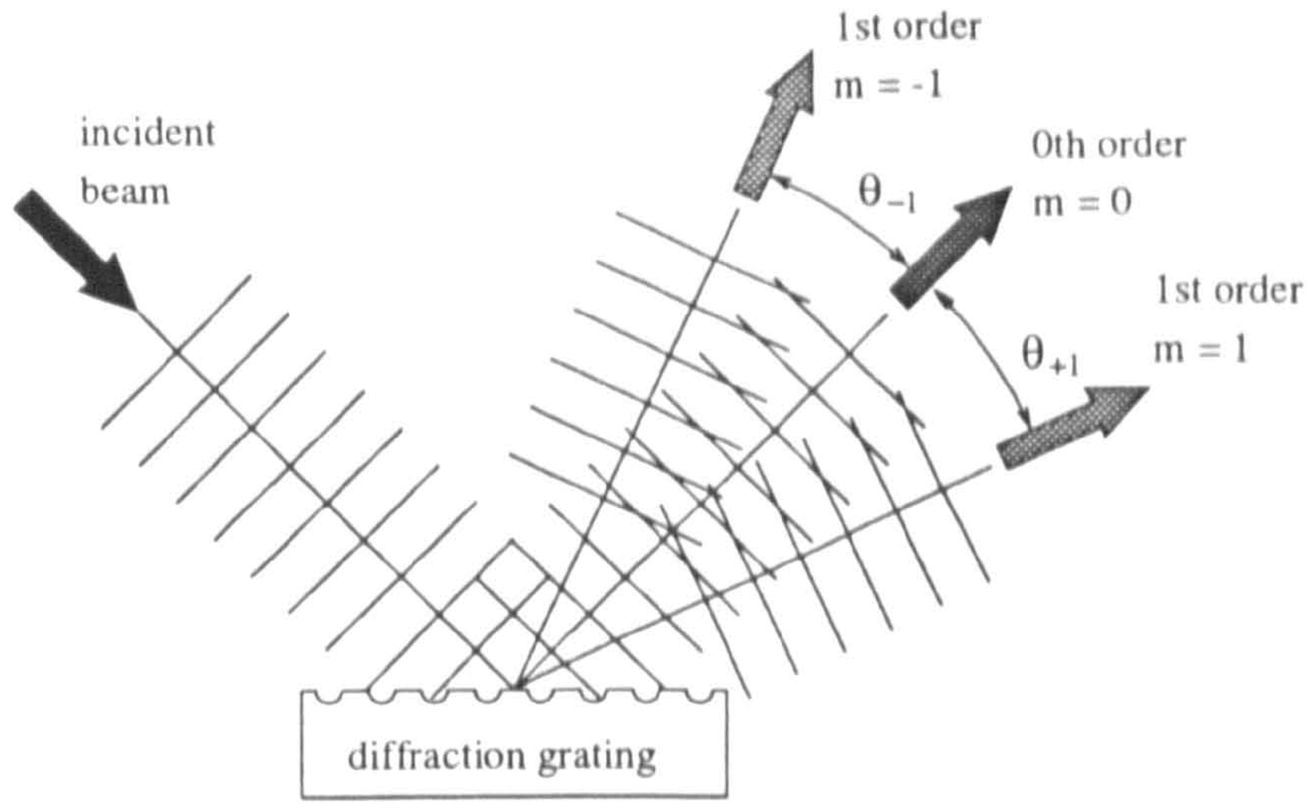


Figure 5.3: Diffraction from a grating.

laser beam is chosen so that the first order diffracted beam is directed away from the specimen surface, as illustrated in the first part of Figure 5.4. Since the diffraction grating is assumed to be undeformed, the incoming plane wavefronts are diffracted into identical plane wavefronts. These wavefronts will travel in a nominal direction according to Equation 5.2.

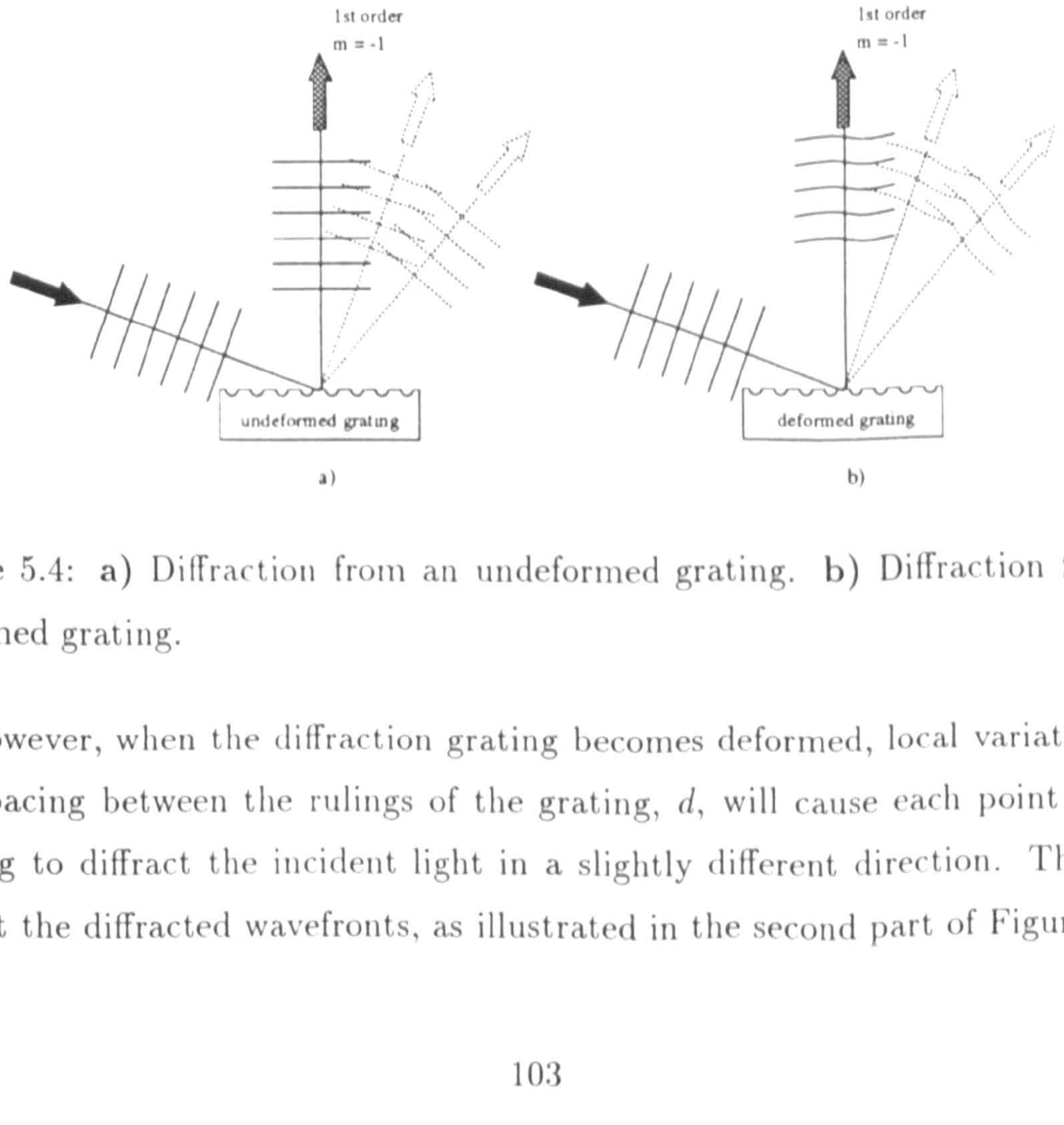


Figure 5.4: **a)** Diffraction from an undeformed grating. **b)** Diffraction from a deformed grating.

However, when the diffraction grating becomes deformed, local variations in the spacing between the rulings of the grating, d, will cause each point of the grating to diffract the incident light in a slightly different direction. This will distort the diffracted wavefronts, as illustrated in the second part of Figure 5.4.

5.2.3 Fringe Formation and Interpretation

Having discussed the phenomena of interference and diffraction, it is now time to relate them to moiré interferometry. The most simple moiré interferometry arrangement is schematically illustrated in Figure 5.5. Beams **A** and **B** are directed onto the specimen surface so that the -1 diffracted order from **A** and the $+1$ diffracted order of **B**, both depart the specimen travelling in the same nominal direction. These two sets of wavefronts are assumed to satisfy the necessary conditions for interference as discussed above, so when they are recombined by the action of the imaging lens, an interference pattern will result.

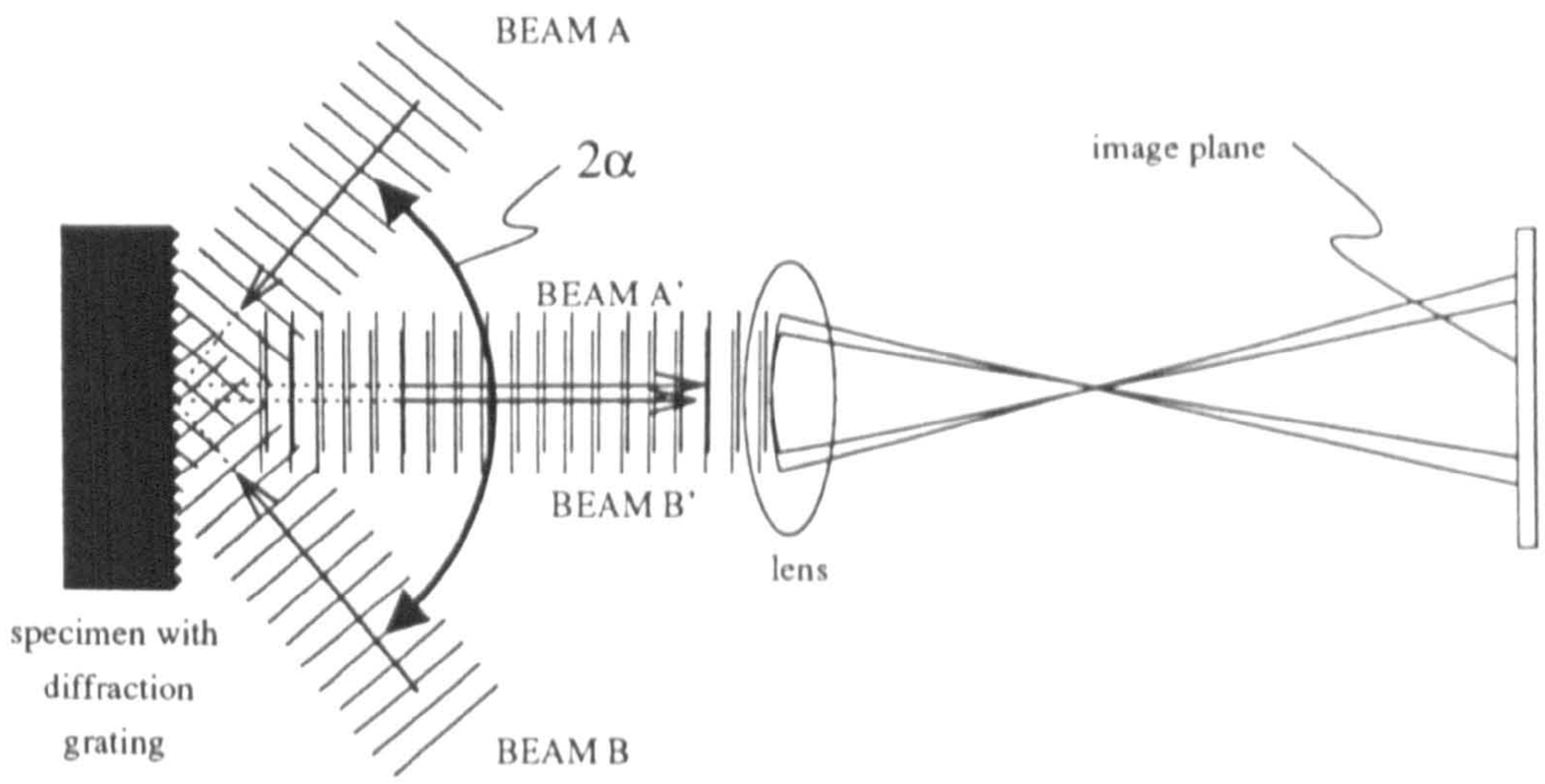
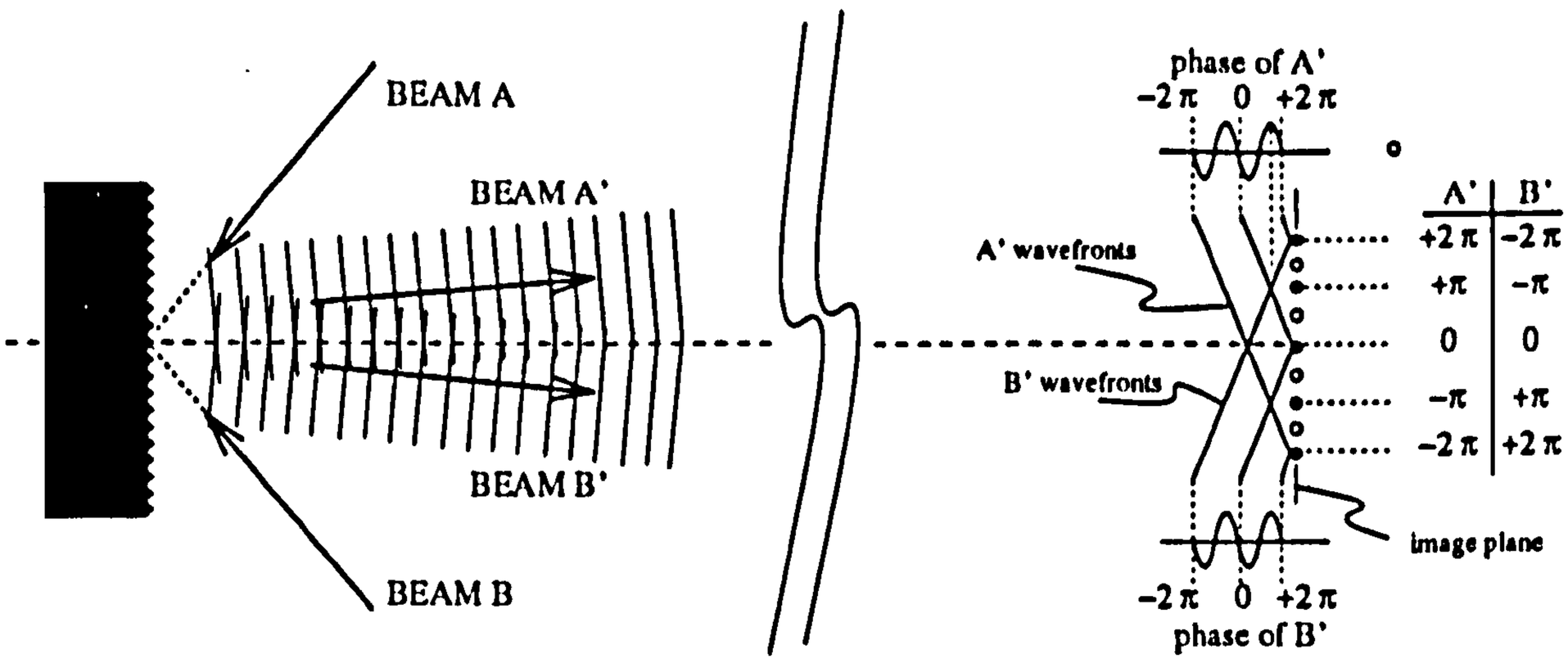


Figure 5.5: A simple configuration for producing a moiré fringe pattern.

The illustration in Figure 5.5 represents what is called the 'null-state' of the interferometer. When the diffraction grating is undeformed, the two diffracted wavefronts $\mathbf{A}'$ and $\mathbf{B}'$ will be perfect plane waves, each having a constant phase across their surface. At each point in the image plane, the interference between the two plane wavefronts will produce the same intensity.

As the specimen is mechanically loaded, the diffraction grating will also deform, which will effectively change the spacing, d, between the rulings of the grating. As d is changed, the angles at which $\mathbf{A}'$ and $\mathbf{B}'$ are diffracted will change for each point on the surface of the grating. For example, if the spacing d is made smaller by a uniform amount over the entire grating, then $\mathbf{A}'$ and $\mathbf{B}'$ will be diffracted as shown in the first part Figure 5.6. When the diffracted wavefronts of Figure 5.6 are recombined in the image plane, they will interfere, and a set



a) b)

Figure 5.6: a) Diffraction from a deformed grating where the spacing, d, between rulings of the grating have been made smaller. b) A graphical representation of the arrival of the diffracted beams in the image plane.

of equally spaced fringes will be formed, because the phase of each of the beams will have been changed. The phases of the two beams at the image plane are graphically illustrated in the second part of Figure 5.6. Fringe minima (where the image would be black) would be observed at the points marked by darkened circles, and fringe maxima would be observed at the open circles. At points between the fringe extrema, the intensity will have a sinusoidal variation. A total of 4 continuous fringes would be observed in this example.

The number of fringes produced for a given deformation of the specimen grating will be determined by the nominal frequency of the grating. The displacement at any point in the field of view in the image plane, $u(x, y)$ can be related to the number of the fringes at that point, $N(x, y)$, by

$$u(x, y) = \frac{N(x, y)}{2f} \tag{5.3}$$

where f is the frequency of the grating which is $= 1/d$. Thus, the number of fringes for a given deformation will increase as the frequency of the grating is increased.

The fringes formed in the image plane can be recorded in a number of ways. One typical method is to use photographic film. Photographic emulsion provides excellent resolution, even for large numbers of fringes, but to make use of the

information contained in the fringe pattern will ultimately require either scanning or manual digitization.

The traditional approach to fringe processing involves extracting the locations of fringe centres and using Equation 5.3 to produce a continuous, interpolated displacement map. Even with the most advanced digital image processing techniques, however, precise determination of fringe centres is difficult and subject to many sources of error. Furthermore, these methods generally require the manual ordering of fringes, since the convexity or concavity of the wavefront can not be automatically deduced from a single fringe pattern.

An alternative approach to photographic film and manual digitization, is to directly record the fringe patterns using conventional video technology. The main advantage of this method is that independent intensity measurements are obtained for each point in the field-of-view. In this way, the gray-scale information between fringe extrema can also be utilized, eliminating the need to interpolate the data.

5.3 Some Experience with Fourier Fringe Analysis

The basic Fourier pattern analysis algorithm proposed by Takeda *et al.* [50], was based on a one-dimensional Fourier transform. The method was refined by employing a two-dimensional Fourier transform by Bone *et al.* [51], greatly improving the ability to isolate the desired signal in the frequency domain.

The Fourier method has the advantage of requiring only one interferogram for the analysis. This makes it particularly attractive for dynamic applications, or instances where vibrations make interferometric measurements difficult. It is also possible to recover both in-plane displacements simultaneously by superimposing the two interference patterns with different carrier frequencies, as was done by Huntley and Field [52].

With the Fourier method, a spatial carrier is first established by tilting one of the interfering beams. The resulting intensity distribution, $f(x, y)$, of a two-

dimensional fringe pattern can be represented by the following equation [50]

$$f(x,y) = a(x,y) + b(x,y)\cos[2\pi(f_x x + f_y y) + \phi(x,y)], \tag{5.4}$$

where $a(x,y)$ is the background intensity, $b(x,y)$ is the modulation amplitude of the fringes, f_x and f_y are the linear components of the imposed tilt, and $\phi(x,y)$ is the phase information of interest. By performing a two-dimensional Discrete Fourier Transform (DFT), usually in the form of an FFT, the spatial frequency domain representation of the fringe pattern becomes

$$F(\zeta,\eta) = A(\zeta,\eta) + C(\zeta,\eta) + C^*(\zeta,\eta), \tag{5.5}$$

where $A(\zeta,\eta)$ is the transform of $a(x,y)$, and $C(\zeta,\eta)$ and $C^*(\zeta,\eta)$ are the positive and negative frequency spectra of the modulated carrier fringes. It is essential for the three terms in the above equation to be completely isolated from one another in the spatial frequency domain. This is satisfied provided the signal of interest is sufficiently band limited, and enough carrier is introduced to separate the positive and negative spectra.

Once the function $F(\zeta,\eta)$ is available, a power spectrum, $|F(\zeta,\eta)|^2$, can be formed by simply taking complex magnitudes of $F(\zeta,\eta)$. Using the power spectrum as an indicator of where the energy of the signal lies, a suitable filter can be constructed that selectively attenuates the undesired portion of the spectrum. A simple rectangular bandpass filter, with a value of unity in the region containing the signal, and zero elsewhere, will usually yield satisfactory results. A more precise method of designing the filter is presented below.

5.3.1 Practical Difficulties

The strength of the Fourier method lies in the ability, afforded by the DFT, to identify the signal and separate it from the noise. However, difficulties are encountered due to certain implicit assumptions involved with frequency domain analysis using the DFT.

The discrete nature of the DFT assumes that the function $f(x,y)$ can be identically decomposed into a sum of complex exponentials having frequencies that are an integer multiple of the fundamental frequency. The fundamental

frequency is implicitly determined by the sampling of the input function $f(x, y)$. In the case of video data acquisition, the individual pixels perform the sampling, so the fundamental frequency can be determined from the number of pixels per scan line. By the Sampling Theorem [48], if the detector array is M×N pixels, then the transform will provide a decomposition of the input function into frequencies ranging from 0–$M/2$ and 0–$N/2$ fringes per screen. Frequencies higher than these values will be aliased into this range, and frequencies that are not exact integers of the fundamental will be rounded to the nearest integer frequency.

A more serious problem with using the DFT results from discontinuities in $f(x, y)$. The DFT assumes that the original function is cyclically defined—that is, it should have the same value, and continuous derivatives across the left and right, and top and bottom edges. Thus, the edges of a fringe pattern are an inherent source of discontinuities. If left unchecked, the process of transforming the function will spuriously distribute large amounts of energy over a wide range of frequencies. This not only obscures the signal power, but also makes isolation of pure signal impossible.

The common solution to this problem is to first remove the arbitrary dc component of the fringe pattern, and then taper the data at the edges using a numerical window, such as the Hamming or Blackman Window. A further improvement can be made by first artificially creating continuity in the function and the first few of its derivatives by replacing the data around the edges with values obtained from cubic spline interpolates. Still, errors on the order of 10% can be expected near the edges. Additional problems arise because the input function is reduced to a very low intensity by the numerical window. Thus, data near the edges becomes acutely sensitive to noise, and more importantly, to leakage from nearby regions of the image.

In addition to the inherent edges of a fringe patterns, discontinuities can also occur in many practical applications of interferometry, such as with interferograms of shock waves in supersonic flow [53], or the analysis of crack-tip displacement fields. Even if the discontinuities do not involve large regions of 'dead' zones, where there are no fringes, the analysis is made difficult by the typical large variation in fringe density between different parts of the observation field. There is little that one can do to prevent leakage from occurring in such cases.

Limited success has been achieved in such cases by forcing continuity across 'dead' zones by replacing the data using interpolating splines, similar to what can be done around the edges. Another method as effective, and easier to implement is to simply replace the 'dead' zone with appropriately scaled random data. When transformed, the random data is spread uniformly over the frequency domain, so that corruption of the data near the 'dead' zone will be minimized. Obviously, the best way to handle such problems is simply to process the data on a region-by-region basis, in such a way as to completely avoid the discontinuities. This is difficult to automate however, and it will still not yield reliable results near the discontinuities.

Another difficulty with the Fourier analysis method occurs when the signal of interest is no longer confined to a narrow band of frequencies. In such cases, it may not be possible to provide sufficient carrier to entirely separate the positive and negative spectra. Even if the signal can be properly isolated, a more carefully designed filter may still be required to suppress the noise.

5.3.2 Optimal Filtering

A simple approach to filtering in frequency space that can be used for large bandwidth signals is to construct an optimal Weiner filter [54]. Since the noise encountered in typical fringe patterns is generally random in nature (white noise), the power spectrum will clearly indicate which portions of the frequency domain the signal occupies, so that a simple thresholding will suffice to isolate the signal. Assuming that the fringe pattern noise is additive, the function $f(x,y)$ can be described as

$$f(x,y) = s(x,y) + n(x,y),$$

where $s(x,y)$ is the pure signal to be measured, and $n(x,y)$ the noise. The power spectrum of $f(x,y)$, can then be expressed as

$$|F(\zeta,\eta)|^2 = |S(\zeta,\eta)|^2 + |N(\zeta,\eta)|^2. \tag{5.6}$$

The terms involving cross products can be neglected because it is assumed that the signal and the noise are un-correlated. (This is precisely the definition of noise!)

The ideal filter $\Gamma(\zeta,\eta)$, is one which, when applied to $F(x,y)$, would yield the spectrum of the signal of interest

$$|S(\zeta,\eta)|^2 = |\Gamma(\zeta,\eta)|^2 |F(\zeta,\eta)|^2. \tag{5.7}$$

Simple algebraic manipulation of the following two equations yields the optimal filter as

$$|\Gamma(\zeta,\eta)|^2 = \frac{|S(\zeta,\eta)|^2}{|S(\zeta,\eta)|^2 + |N(\zeta,\eta)|^2}. \tag{5.8}$$

The optimal filter is easily constructed by first smoothing and then thresholding a normalized gray-scale version of the (log magnitude) power spectrum $|F(\zeta,\eta)|^2$, so that the peaks representing the signal $|S(\zeta,\eta)|^2$ become isolated. Once $|S(\zeta,\eta)|^2$ has been determined, $|N(\zeta,\eta)|^2$ can be deduced from Eq. 5.6. The optimal filter can then be constructed and extended to negative frequencies as necessary.

5.3.3 Demonstration of Fourier Fringe Analysis

Figure 5.7 is an example of a fringe pattern obtained by taking the difference of two acquired images, with a phase shift of 180° between them. A carrier pattern of approximately 17 fringes/screen was introduced, to facilitate isolation of the signal in frequency space. The signal modulating the carrier pattern represents horizontal displacements on the surface of a deformed composite laminate specimen, under a bending load. The field of view is a 2.67×2 millimeter rectangle, and the fringe contour interval is 0.25 μm. The specimen is composed of layers of different materials, and the relative displacements between each layer is of particular interest.

After removing the arbitrary dc component and applying a Hamming window, a two-dimensional FFT was performed, resulting in a frequency space representation of the fringe pattern. An optimal filter was constructed from the power spectrum using the method discussed above, and a wrapped, filtered fringe pattern was obtained, and is shown in Figure 5.8. The wrapped and filtered pattern was unwrapped using the phase unwrapping method presented below to yield a continuous phase map (displacement field). A topological plot of this result is presented in Figure 5.9, where each contour represents a change of 0.15 μm.

This simple example illustrates how the Fourier method can be used to im-

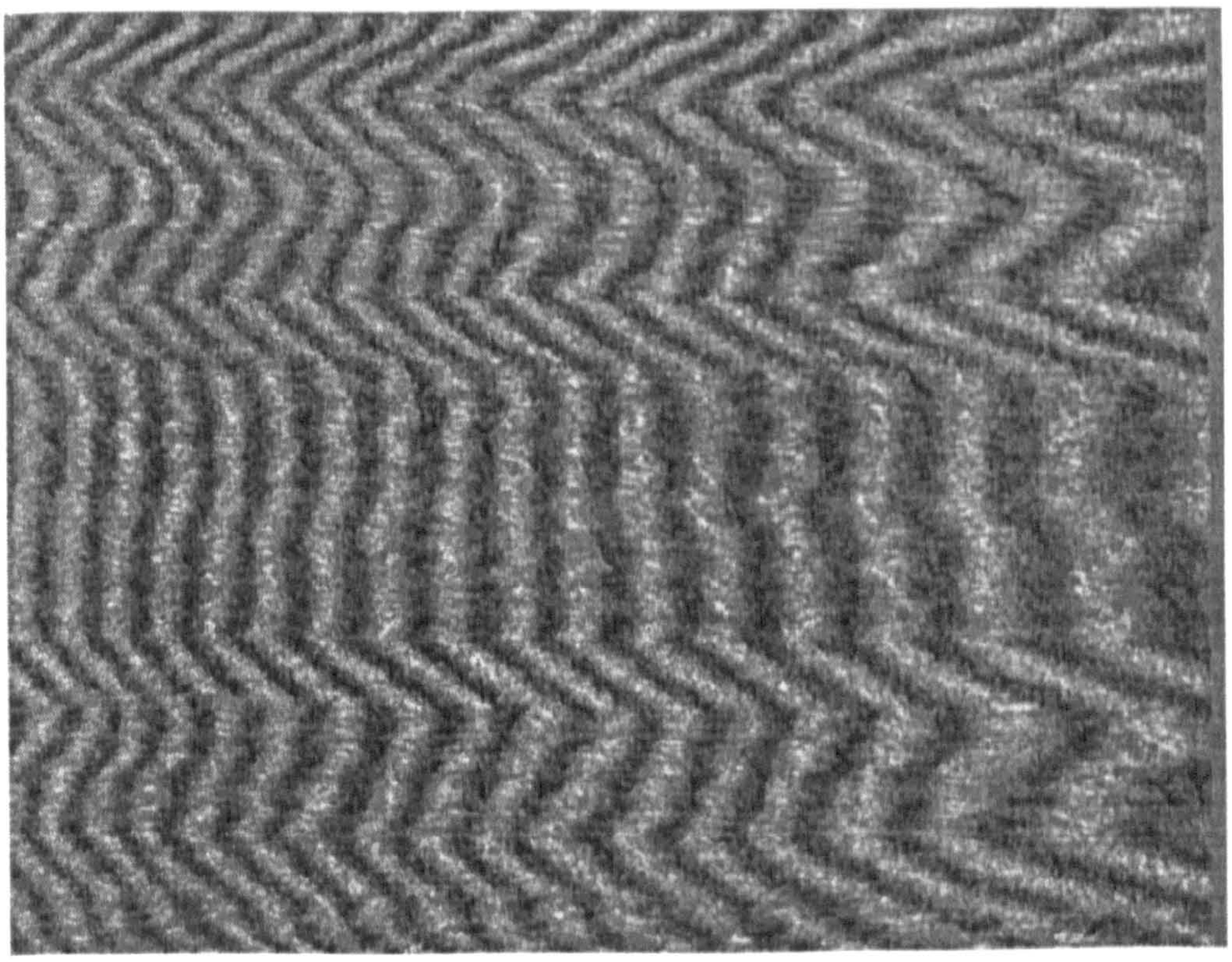


Figure 5.7: Raw fringe pattern of a composite laminate under a three-point bending load.

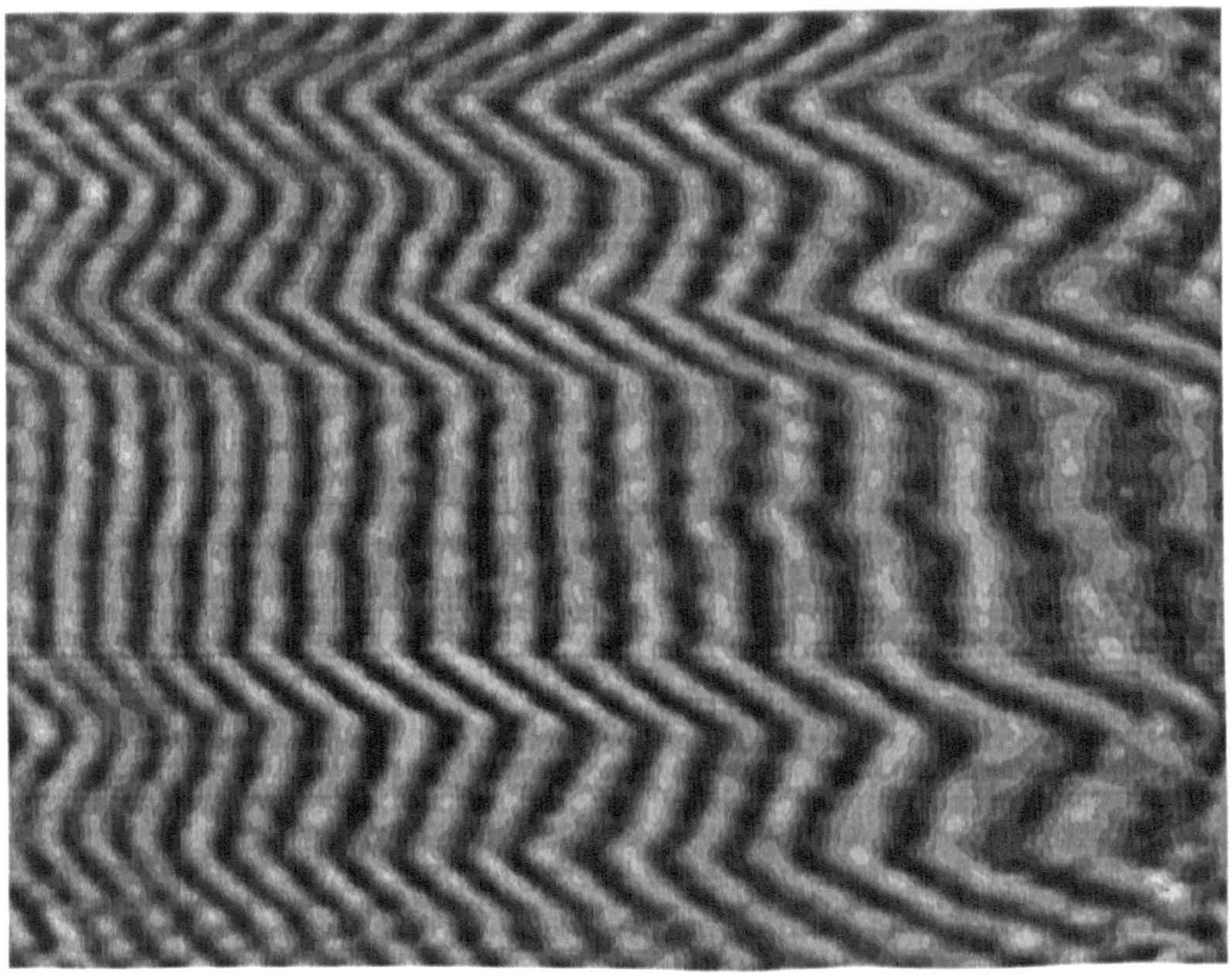


Figure 5.8: Fourier filtered fringe pattern.

prove the signal-to-noise ratio of an image, and produce a continuous phase map. However, the improved fringe contrast comes at the expense of errors introduced particularly at the corners and edges of the image. This effect can be seen clearly in Figures 5.8 and 5.9.

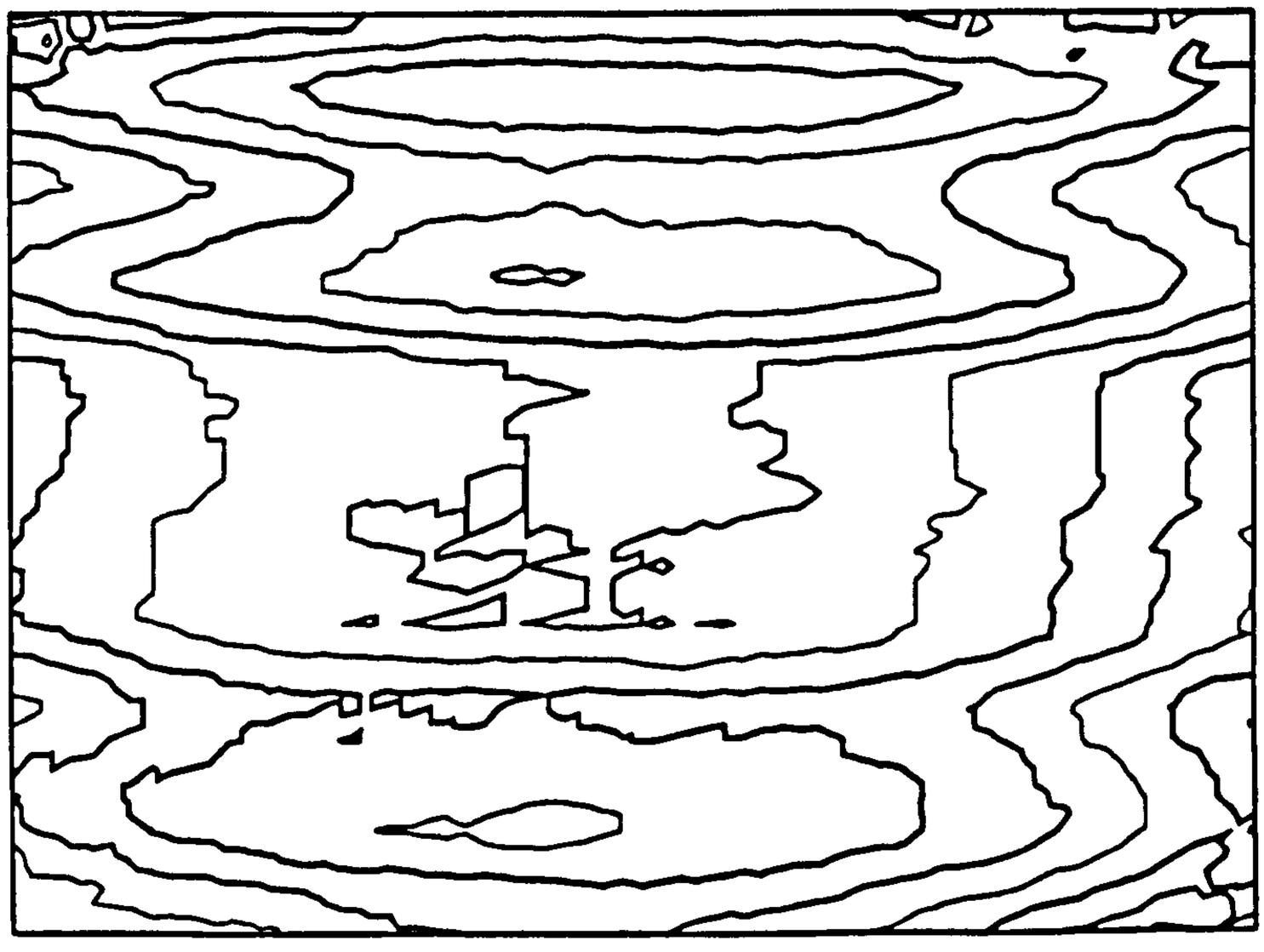


Figure 5.9: Topological plot of the phase extracted from the Fourier filtered wrapped fringe pattern.

In the next section, an alternative method for processing interferometric fringe patterns is presented, based on the technique of phase shifting interferometry. Both the Fourier and phase shift methods are based on virtually an identical analysis—the main difference is that the phase shift algorithms perform computations in the spatial domain, while the computations for the Fourier method are performed in the spatial *frequency* domain [55].

5.4 Phase Shifting Interferometry

Phase shifting interferometry (PSI) is becoming an increasingly popular technique for measuring optical path differences in a variety of applications. The technique has been successfully used with holographic, multiple-wavelength and speckle interferometry to produce surface contours and deformation measurements [56].

The method of PSI dates back to at least the work of Carré [57], and has received much attention more recently, due to the increased availability of inexpensive piezo-ceramics and fast video frame buffer technology. An extensive review of phase shifting interferometry can be found in [55]

Traditional analysis of interferograms, such as the identification and ordering of fringe centres, results in a precision of not better than $\lambda/10$, which is well below that contained in most interferometric data. Furthermore, the full potential of most experimental data often becomes lost due to the difficulty in providing a systematic data reduction process. The PSI approach makes automated processing of the data straightforward, and provides the principle motivation for extending the technique into the realm of moiré interferometry.

The data acquisition system and the phase shifting moiré interferometer developed for this work are described in Section 5.5. First, the basic theory and several important practical aspects of phase shifting interferometry are discussed in the following sections.

5.4.1 Basic Theory

The biggest single change in all types of optical instrumentation recently has been the integration of computers into the the measurement system. A prime example of this situation has been the development of phase shifting interferometry. PSI is not a specific optical hardware configuration, but a complete data collection and analysis method.

Although it was originally developed as a method to assist in the assessment of optical components, the techniques of PSI are applicable to almost any type of interferometer. The PSI technique is based upon recording several intensity distributions arising from the interference of two beams while known, uniform phase shifts are introduced across one of them between frames. The resulting

intensity distribution can then be processed in a relatively simple manner to recover the original phase relationship between the two interfering beams.

There are numerous ways to induce the necessary phase shifts and record the resulting intensity distributions. One can either ramp the phase continuously, or provide discrete phase steps. Discrete or continuous phase modulation can be achieved by moving a mirror or a grating, tilting a glass plate, rotating a half-wave plate or analyzer, heating or deforming an optical fibre, or using an acousto-optic or electro-optic modulator.

If the interferogram is viewed during the phase shifting operation, the fringe pattern will appear to move across the field of view. Because the induced phase shift is constant for each point in the image, the fringes will march in unison.

A more useful way to visualize the action of phase shifting is to consider the variation in intensity at a single point during the operation. Consider a single pixel of a conventional video camera detector that digitizes brightness in the range of 0–255, where 0 and 255 represent black and white, respectively. If the brightness of this pixel is originally black, then as the phase is gradually ramped from 0 to 2π, the detected intensity for this pixel will vary sinusoidally—the brightness will change from black to white and back to black.

The relationship between the brightness of the pixel and the induced phase shift is quite simple. As will be discussed below, PSI utilizes this relationship to extract the underlying phase distribution of a fringe pattern, from several recorded interferograms separated by discrete phase shifts.

A unique aspect of PSI is that the phase of the fringe pattern under study is determined *independently* at each point—without the need to consider data from other points in the image. This is because all of the necessary information is encoded in the point-by-point variations of intensity with respect to the induced phase shifts. As a consequence, the noise rejection is achieved without degrading the phase data close to discontinuities in the fringe pattern—an important consideration in experimental mechanics applications.

5.4.2 Phase Extraction Algorithms

There are numerous methods available for extracting the phase function from a set of shifted intensity distributions. The method described below is the most general,

and can be used when the individual phase shifts are not evenly spaced. A more thorough discussion of the various extraction methods can be found in [55, 56].

The acquired intensity data for a two-beam interference fringe pattern can be written

$$I_{ij} = A_j + B_j \cos[\phi_j + \delta_i], \tag{5.9}$$

This is essentially a restatement of Equation 5.4, where ϕ_j is the unknown phase (including tilt terms), A_j is the background intensity variation, B_j is the modulation, and δ_i is the induced phase shift for each of the N acquired frames. The subscript j is used to identify the individual detector (pixel) locations, and replaces the (x, y) range used in Equation 5.4.

The solution for the wavefront phase, ϕ_j, is obtained by first writing Equation 5.9 in a manner amenable to a least-squares solution [58],

$$I_{ij} = a_j + b_j \cos(\delta_i) + c_j \sin(\delta_i) \tag{5.10}$$

where $a_j = A_j$, $b_j = B_j \cos(\phi_j)$, and $c_j = -B_j \sin(\phi_j)$. For each detector location, j, Equation 5.10 represents a system of i independent equations in the three unknowns a, b, and c. Provided at least three frames of information are available, and the δ_i are known, the solution to the above problem is written

$$\left\{ \begin{array}{c} a_j \\ b_j \\ c_j \end{array} \right\} = [\mathrm{A}]^{-1} \mathrm{b}_j,$$

with

$$\mathrm{A} = \left[\begin{array}{lll} N & \sum \cos(\delta_i) & \sum \sin(\delta_i) \\ \sum \cos(\delta_i) & \sum \cos^2(\delta_i) & \sum \cos(\delta_i) \sin(\delta_i) \\ \sum \sin(\delta_i) & \sum \sin(\delta_i) \cos(\delta_i) & \sum \sin^2(\delta_i) \end{array} \right], \quad \mathrm{b}_j = \left\{ \begin{array}{l} \sum I_{ij} \\ \sum I_{ij} \cos(\delta_i) \\ \sum I_{ij} \sin(\delta_i) \end{array} \right\}. \tag{5.11}$$

and the limits of the summations are from 1 to N. The matrix [A] will depend only on the known phase shifts, and can be assembled and inverted only once.

The wrapped phase is recovered as

$$\phi_j = \tan^{-1}\left(\frac{-c_j}{b_j}\right). \tag{5.12}$$

The data reduction process is much simplified when only four phase shifts of 90° are used. Then, $\delta_i = i(\pi/2)$, $i = 0 \ldots 3$ and

$$\phi_j = \tan^{-1}\left[\frac{I_{3j} - I_{1j}}{I_{0j} - I_{2j}}\right]. \tag{5.13}$$

Equations 5.12 and 5.13 will produce a least-squares estimate of the phase, ϕ_j, assuming the reference phase shifts are known accurately. In practice, the actual phase shifts are susceptible to a number of systematic error sources including hysteresis and linear drift of the phase shift device. These types of systematic errors have been studied extensively [59, 60, 61]. Their effect can be minimized by carefully determining the response of the phase shift device, and using phase extraction algorithms that are insensitive to quadratic reference phase shift error [57, 62]. For well calibrated systems, a precision greater than $\lambda/100$ can be achieved [56].

Random phase error sources, such as caused by mechanical vibrations or laser source instabilities, are naturally more difficult to analyze, and generally must be dealt with on a case-by-case basis. Schwider *et al.* [60] discussed a method for reducing errors arising from small reference phase deviations based on averaging independent runs separated by a $\pi/2$ phase shift. Schwider [63] also demonstrated that at least an order of magnitude improvement could be attained by least-squares fitting a function of the form $\Delta\phi = a + b\cos(2\phi) + c\sin(2\phi)$ to the measured phase data, and then subtracting it from the initial phase results. This approach makes use of the well documented form of the $\cos(2\phi)$ error in the wrapped phase measurements resulting from deviations in the reference phase shifts [55, 56].

Of course, a most effective method for minimizing the influence of random errors, is to perform several measurements, and average the results. In some cases, however, it is not always possible to acquire more than one data set—such as when studying mechanical strains that change over time.

5.4.3 Practical Aspects of Data Acquisition

Unlike conventional optical testing, moiré interferometry employed for material deformation analysis is generally performed under non-ideal operating conditions. The most serious problem encountered during a typical experiment is that of vibration elimination. Although great care may be taken to stabilize the optics and isolate the system from mechanical disturbances, testing conditions may make the complete elimination of vibrations of the specimen impossible.

Vibrations affecting fringe formation are generally classified as either low or high frequency, depending on whether they have a period less than or greater than that of the video frame rate. High frequency vibrations cause the fringe system to be in motion during the integration time of the detector, resulting in diminished fringe contrast. Since it is unlikely that each of the acquired frames will be affected in the same way, an intensity discrepancy will be introduced between frames and a slight error can be expected when extracting ϕ_j [64, 65]. Another problem of high frequency vibrations is that of 'phase jitter', which is understood as an artifact of interlaced video sampling.

Low frequency vibrations occurring during the data acquisition process primarily result in a discrepancy between the nominal and actual reference phase shifts. When the amplitude of the disturbance is low, there is little degradation of the image but an error is introduced in the shifted phase. Instead of processing the data using the nominal phase shifts, it would, in principle, be preferable to estimate the actual phase shifts directly from the intensity measurements, and use these values to extract the phase using Equation 5.12.

Such an approach would be reasonable only if the reference phase shifts could be determined with sufficient confidence to justify the additional effort required to use Equation 5.12 instead of Equation 5.13. Thus, it is first necessary to analyze how much error in measuring ϕ_j can be expected from using Equation 5.13 when the relative shifts are not the expected 90°, 180°, and 270° values. The difficulty with such an analysis, however, is that the resulting estimated error, $\Delta\phi$, in measuring the phase will depend on the actual values for *each* of the N reference phase shifts.

To overcome this complication, a statistical analysis was performed whereby the average error in measuring the phase was determined as a function of a charac-

teristic reference phase shift error, $\bar{\delta}_\epsilon$. Synthetic intensity data was generated using Equation 5.9, but the phase shift values were obtained from $\delta_i = i(\pi/2) + \sigma_i \bar{\delta}_\epsilon$, where the σ_i represent normally distributed random numbers with a variance of unity. Since the distribution of the σ_i values was chosen to be Gaussian, it is implied that $\bar{\delta}_\epsilon$ will represent the mean error in the individual phase shifts over many simulations.

For each run, the ϕ_j were extracted from the synthetic fringe data using Equation 5.13, and $\Delta\phi$ was determined as one half of the error range when ϕ_j was compared to the generating phase function. This definition of $\Delta\phi$ is more relevant to the analysis than the rms of the individual errors at each j location, since it characterizes the error for the entire data set, as opposed to the error at each point *within* the data set. For each of the six $\bar{\delta}_\epsilon$ considered, approximately

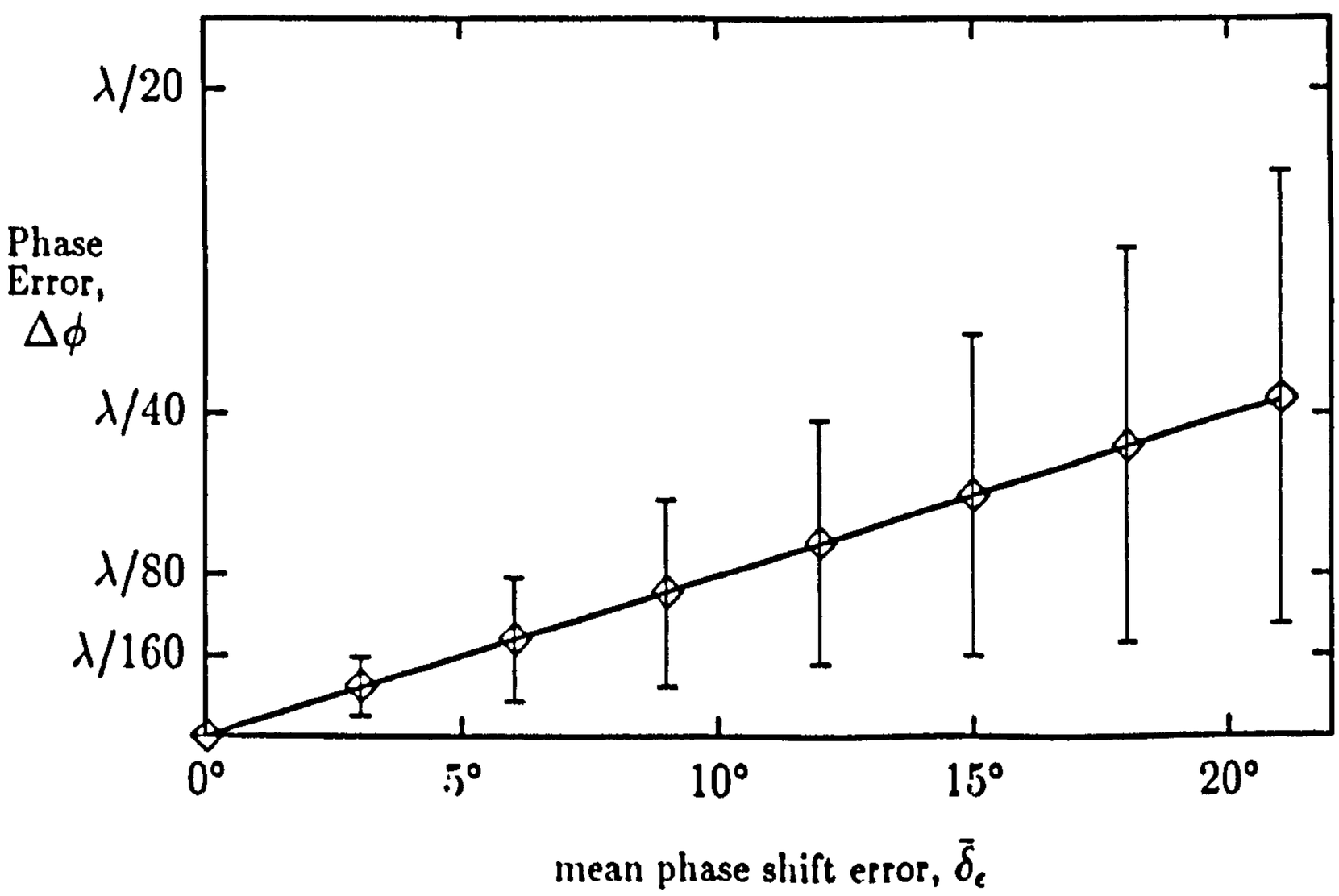


Figure 5.10: Predicted error in extracting ϕ as a function of characteristic phase shift error.

700 runs were required before the average phase error, $\overline{\Delta\phi}$, converged. The results of the statistical simulation are presented in Figure 5.10, where the bars indicate the standard deviations for the $\Delta\phi$ values. Thus, to insure a precision on the order of $\lambda/100$, the phase shifts must be known to within 5°.

5.4.4 Reference Phase Shift Determination

Assuming that the phase shifts are linear and evenly spaced, a simple expression was proposed by Schwider *et al.* [60] for calculating the nominal phase step α, when five intensity measurements are made ($N = 5$). The equation is $\cos\alpha = (I_4 - I_0)/2(I_3 - I_2)$. Using this equation, Cheng and Wyant [59] have reported that α can be determined with a precision of 0.139°. However, when low frequency vibrations become significant, the reference phase shifts will be perturbed about their nominal value, and it becomes necessary to determine each phase shift separately.

As a first approximation, conventional fringe analysis methods can be used to provide such information. For instance, slices of intensity data can be extracted from each acquired frame and processed to determine the symmetry points of each strip. An algorithm proposed by Snyder [66] for simultaneously low-pass filtering and differentiating arrays of intensity data facilitates this task. The locations of the extrema for each strip are found as zero-crossings of the filtered data. These values can be fitted to a least-squares line, where the underlying phase is recovered as the slope, and the constant phase offset as the intercept. The relative phase shifts can be determined by comparing the individual offset values.

This method can be improved by differentiating the original strip of data two times, thus obtaining twice as many data points for the regression, and by interpolating between the discrete pixel locations to determine the zero-crossings more precisely. In practice, however, this method has several shortcomings. The precision will be limited by the length of the strip taken, but longer strips require disproportionally more computation time due to the convolutions involved. Finally, since a linear phase profile is assumed, any deviation from this condition will also degrade the accuracy. Experience indicates that with typical fringe patterns this method can determine the phase shifts only to within $\pm 15^{\circ}$.

Lai and Yatagai [67] proposed a method for directly evaluating the relative phase shifts in real time using an additional set of parallel Fizeau fringes recorded along with the interference data of interest. Due to the essentially linear phase profile of these fringes, the first order peaks in the Fourier domain could be easily isolated, and the reference phase shifts determined from complex logarithms of

the discrete frequency coefficients. Using an eight step algorithm ($N = 8$) these authors reported a precision on the order of $\lambda/500$ in the measurement of optical surfaces.

Evaluating the reference phase shift for each strip in the Fourier domain provides sufficient precision and can be performed relatively fast using conventional DFT algorithms. But, like the convolution method described above, it requires that a suitable strip of data with a linear phase profile be available, which will not, in general, be feasible. Another method that utilizes additional information was presented by Lassahn *et al.* [68]. Here the authors used two additional shuttered images to provide information on the relative intensity of each beam. This information, along with their acquired phase shifted fringe patterns, was used to determine the relative phase shifts for a set of images.

A graphical technique that reveals useful qualitative information about sequentially phase shifted images was proposed by Kinnstaetter *et al.* [65]. The method involves selecting two pixels in approximate phase quadrature and displaying these intensity values in a Lissajous pattern. Deviations in the nominal phase shift due to low frequency mechanical disturbances displace the data points along the rim of the figure, while high frequency vibrations collapse the figure inward, indicating a loss of contrast. Farrell and Player [69] extended this approach to quantitatively determine the relative phase shifts by fitting an ellipse to their Lissajous figures. The authors concluded that this method could determine the individual phase shifts to within $\pm 4°$.

For this work, it was necessary to implement a procedure for determining the actual phase shifts during an experiment, so that Equation 5.12 could be used with these δ_i to extract ϕ_j. This process ought not to impose an excessive burden on the user, nor require any additional information or hardware than would be available during a conventional moiré interferometry experiment. Below, three general methods are discussed for determining the reference phase shifts for a set of fringe patterns, in the case of $N = 4$. A synthetic intensity data model that incorporates several types of fringe pattern disturbances is then described and used to evaluate each of the methods. All three methods require only a single strip of intensity data from each acquired frame to determine δ_i, and can be used with an arbitrary number of shifted fringe patterns.

Fourier Series Method

The first method that was considered was based on computing the coefficients, p_1 and q_1, of the fundamental harmonic, $p_1 \cos \omega x + q_1 \sin \omega x$, in the Fourier series expansion for each strip of intensity data. This is equivalent to least-squares fitting a sinusoid to the data with A_j and B_j modeled as constants. In general, A_0 and B_0 will be different for each strip of intensity data, since the method only considers each strip individually. The δ_i are determined for each strip as $\tan^{-1}(-q_1/p_1)$, A_0 is taken as the mean of the I_{ij} values, and $B_0 = \sqrt{p_1^2 + q_1^2}$.

The Fourier Series Method requires that the fringe period be known in advance, and assumes that ϕ_j is linear. It was found that the inclusion of higher order Fourier coefficients was detrimental to the performance of the method, while the removal of linear background trends improved the calculation.

Iterated Linear Method

The Iterated Linear Method uses the nominal phase shifts and Equation 5.12 to compute an initial estimate of ϕ_j, as in the conventional phase extraction algorithm. These ϕ_j values are then used to compute a new set of δ_i by assembling Equation 5.11 with the roles of ϕ and δ reversed—thus, a corrected version of the phase shifts are recovered as $\delta_i = \tan^{-1}(-c_j/b_j)$.

This method was originally used by Okada *et al.* [70], who suggested that about three iterations of the above scheme are sufficient. However, no significant improvement can be expected after just the first iteration, since this approach does not completely couple the interaction of δ and ϕ in Equation 5.9. The Iterated Linear Method has the advantage of being robust with respect to variations in the underlying ϕ_j values, because it computes these on a pixel-by-pixel basis.

Nonlinear Least-Squares

The most general approach to the problem of determining the reference phase shifts is to use Equation 5.9 and the given intensity data I_{ij}, to simultaneously solve for A_j, B_j, ϕ_j, and δ_i using the the technique of nonlinear least-squares, as was done by Lassahn *et al.* [68]. This method requires an initial estimate of the unknown parameters and proceeds to refine these values through successive iterations. The advantage of this method over the Iterated Linear Method is that

the relevant variables (particularly B_j, ϕ_j and δ_i) are allowed to mutually interact with one another during the iterated solution process.

Three implementations of this method were considered, differing from each other in how the A_j and B_j are modeled. In the first case, call it NLS0, A_j and B_j were assumed to be constants, A_0 and B_0. In NLS1, A_j and B_j were modeled as linear functions of position, i.e. $A_j = A_0 + A_1 x_j$ and $B_j = B_0 + B_1 x_j$. Similarly, in NLS2, A_j and B_j were represented as quadratic functions. The phase distribution was modeled as $\phi_j = \phi_1 x_j + \phi_2 x_j^2$ in all cases. The A_j, B_j, ϕ_j and δ_i values from the Fourier Series Method were used as the initial estimates for the iteration procedure.

It was found that all three NLS methods were quite sensitive to the initial estimates for ϕ_j and δ_i. In fact, the solution procedure diverged when a δ was more than 15° in error. However, all three implementations showed considerable tolerance to large deviations in the A_j and B_j estimates.

The Numerical Simulations

The purpose of the numerical simulation was to compare the three methods discussed above on a quantitative basis. Five types of fringe pattern 'disturbances' were included in the generation of synthetic data by rewriting Equation 5.9 as

$$I_{ij} = A_j + B_j \cos[\phi_j + \delta_i] + \sigma_\eta \eta + \sigma_\mu f(\Phi)\mu, \tag{5.14}$$

where $\sigma_\eta \eta$ is an additive noise term with σ_η being a random number[2] between ± 1, and η a nominal scaling factor. The term $\sigma_\mu f(\Phi)\mu$ represents a coherent noise effect, defined in the same way, except for the additional functional dependence on the instantaneous phase, $f(\Phi = \phi_j + \delta_i)$. The phase dependence of this noise term can be explained by considering small statistical variations in the complex amplitudes of the reference and object waves due to, for example, dust diffraction in the interferometer. In the case of constructive interference at a fringe maximum, a considerable speckle fluctuation will be present in the resulting intensity distribution, while this effect is suppressed at a fringe minimum [71].

[2]The random numbers (σ_ν and σ_μ) used in Equation 5.14 are *not* uniform Gaussian deviates. They were intentionally chosen to be *randomly* distributed between ± 1.

Thus, $f(\Phi)$ can be approximated by any smooth function that satisfies $f(\pi n) = 0$ and $f(2\pi n) = 1$, such as $f(\Phi) = \cos^2[(\phi_j + \delta_i)/2]$.

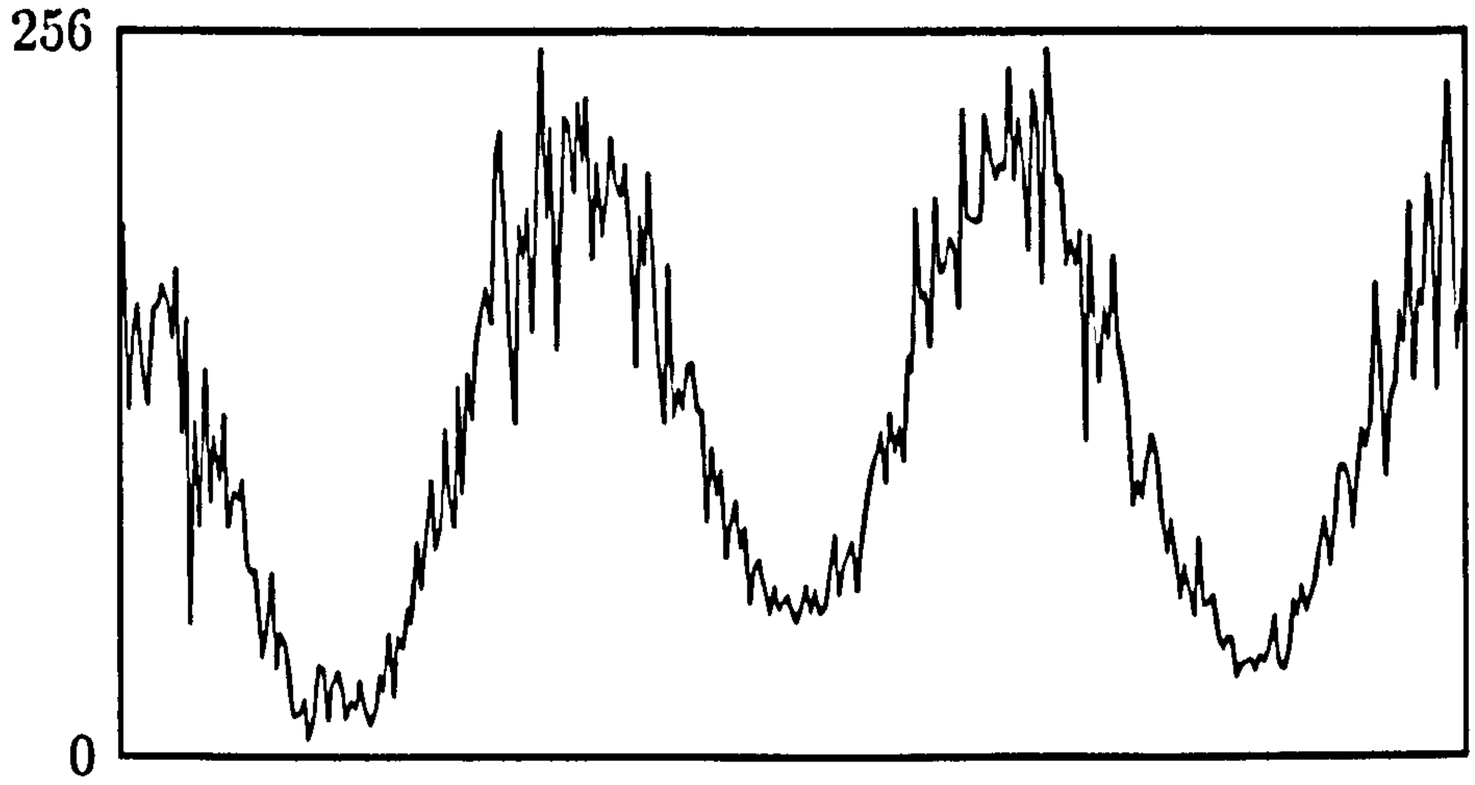


Figure 5.11: A typical example of the synthetic fringes used in the analysis.

The background, or dc variation was modeled using $A_j = \sigma_{A_0} A_0 + \sigma_{A_1} A_1 x_j + \sigma_{A_2} A_2 x_j^2$, where again, σ_{A_0}, σ_{A_1} and σ_{A_2} are random numbers between ± 1, and A_0, A_1 and A_2 are nominal magnitudes. Likewise, the modulation variation was modeled as $B_j = \sigma_{B_0} B_0 + \sigma_{B_1} B_1 x_j + \sigma_{B_2} B_2 x_j^2$ and the underlying phase as $\phi_j = \sigma_{\phi_1} \phi_1 x_j + \sigma_{\phi_2} \phi_2 x_j^2$.

The nominal disturbance scaling factors A_0, A_1, A_2, B_0, B_1, B_2, ϕ_1, ϕ_2, η and μ were selected one-by-one on a qualitative basis by visually inspecting the synthetic fringes and choosing values that would be representative of a 'worst case' scenario. The intent was not to rigorously characterize fringe pattern disturbance, but merely to provide a framework within which different methods for determining the reference phase shifts could be evaluated. A typical strip of synthetic data that includes all five types of error sources is shown in Figure 5.11.

In the first part of the analysis, the effects of each type of disturbance acting alone were evaluated by setting all of the nonrelevant scaling factors in Equation 5.14 to zero. The δ_i values were chosen arbitrarily as 5°, 88°, 191°and 265°. Synthetic data with a fringe period of 32 pixels was produced using Equation 5.14, and the three methods were used in turn to estimate the reference phase shifts. The fringe period and length of the strip was found to have little influence on the results for all of the methods, and a length of 90 pixels was arbitrarily chosen for the analysis.

Estimation Method	Quantity Varied					
	Background Variation (A_j varied)	Modulation Variation (B_j varied)	Additive Noise (η non-zero)	Coherent Noise (μ non-zero)	Phase Variation (ϕ_j varied)	All Error Types Included
Fourier Series	2.2±0.6	0.77±0.4	2.4±1.9	2.2±1.8	2.1±1.7	9.5±5.7
Iterative Linear	3.2±2.9	1.5±0.8	2.5±2.0	2.6±2.1	0.0±0.01	10.7±5.5
Nonlinear 0	2.0±1.4	0.77±0.4	2.4±1.8	2.2±1.7	2.2±1.7	9.5 ±2.2
Nonlinear 1	1.8±1.2	0.75±0.4	2.9±1.7	2.4±2.0	2.1±1.6	4.6±2.0
Nonlinear 2	1.1±0.8	0.76±0.4	3.1±1.9	2.7±2.0	2.2±1.6	5.1±2.5

Table 5.1: Mean phase shift errors (in degrees) for several types of fringe pattern disturbance.

Since the actual phase shifts were known, the mean phase shift error, $\bar{\delta}_\epsilon$, was determined for each method at the end of each run. This process of generating synthetic data, computing the 'apparent experimental' phase shifts using each method and determining the mean shift error was repeated with a new set of random numbers until the $\bar{\delta}_\epsilon$ converged. The resulting $\bar{\delta}_\epsilon$ values and their standard deviations for each type of fringe pattern disturbance are presented in the first five columns of Table 5.1.

The results confirm that the Iterated Linear Method is indeed insensitive to phase variation, while the other methods all exhibit approximately the same error. The NLS methods show an increasing ability to accommodate variations in A_j, and all methods seem to be relatively insensitive to variations in B_j. Additive and coherent noise had about the same effect on all of the methods, but was slightly more pronounced for the higher order NLS methods.

Finally, all five types of errors were allowed to operate concurrently to obtain an estimate of how accurately the methods could determine the reference phase shifts in the general case. The last column of Table 5.1 summarizes these results. The NLS0 Method offers little improvement over the initial estimates provided by the Fourier Series Method, but a significant refinement is available in the NLS1 Method. It is interesting to note that the NLS2 Method does worse than the NLS1 Method with respect to almost all of the different types of disturbances. This is probably because the additional degrees of freedom in the NLS2 model

introduce ambiguous relationships between the parameters of the model, thus allowing the iteration procedure to prematurely settle at a low value of χ^2 for the nonlinear least-squares fit.

5.4.5 Demonstration of Reference Phase Determination

The ability to measure the phase shifts between frames eliminates the need to rigorously calibrate the phase shifter. Nevertheless, a higher degree of noise rejection is possible when the shifts are evenly spaced, so an initial calibration procedure was performed. This was done by using a master grating mounted directly to the optical table so that mechanical disturbances were negligible. By using the above methods for determining the phase shifts, an iterative approach similar to that of Schwider et al. [60] was used to determine the nominal voltage ramps for each pair of mirrors in the interferometer.

As discussed earlier, the experiments necessarily involve mounting specimens on a load stage designed to allow for several degrees of motion. After all of the usual precautions were taken to avoid mechanical instability, considerable fringe movement was still present. Typically, low frequency vibrations cause the fringes to fluctuate about some nominal position such that the actual reference phase, δ, differs from the nominal value by approximately $\pm\pi/8$. High frequency vibrations are also present, and their effect ranges from simply decreasing the fringe contrast and producing fringe jitter to completely obliterating the image.

To overcome these problems during the experiments, a 'manual' approach for acquiring the shifted fringe patterns was adopted. For each frame, the phase shifter is moved into the required position, and the user then judges when to shutter the image. If the acquisition is deemed unsuccessful for any reason, the image was rejected, and another image grabbed. Otherwise, the image was stored in the frame buffer, and the process repeated for the next frame. In this way, high frequency disturbances were eliminated at the user's discretion, and poorly spaced phase shifts were avoided.

Once the required number of frames had been acquired, a suitable reference strip was selected in a low-noise region of the fringe pattern to be used for determining the phase shifts between each image. The reference strip spanned at least one period, but did not need not be exact, since the actual fringe spacing was

automatically determined by using the conventional fringe processing methods discussed above. The δ_i between the frames were then determined using either the Fourier Series, Iterated Linear or Nonlinear Method.

For the purpose of validating this approach, a simple solid mechanics example with a well known theoretical solution was chosen to evaluate the three methods for determining the δ_i. The problem is illustrated in Figure 5.4.5, and involved a homogeneous orthotropic beam loaded in a conventional three-point bend arrangement. Before loading, a small region of the specimen was chosen and the horizontal displacement fringe pattern was nulled. After the load was applied, the specimen was repositioned so that the selected region was centered, and rotated to impose symmetry.

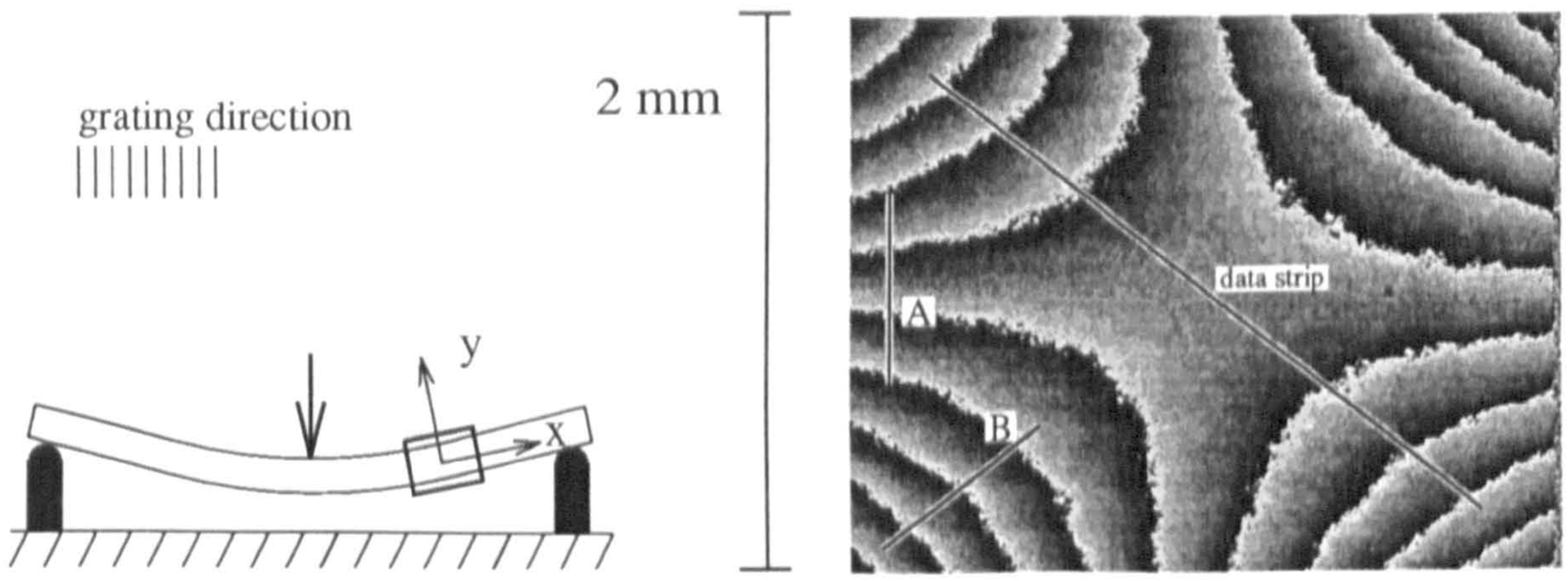


Figure 5.12: Solid mechanics three-point bend example.

A particularly simple theoretical expression [22] is available for the horizontal displacement component along a diagonal strip, as indicated in Figure 5.4.5. For such a strip, a first order approximation to the displacements are given by $u_j = a + b(x_j - c)^2$, where a and c are arbitrary constants, and b is related to the flexural rigidity of the beam, EI, and the applied moment, M, as $b = M/EI$.

Two reference strips were chosen for the determination of the phase shifts, denoted 'A' and 'B' in Figure 5.4.5. Along 'A', the displacements, and hence ϕ_j, are predicted to vary linearly with position, while along 'B' they will have a quadratic profile. Four shifted fringe patterns were recorded using the manual method described above, and the δ_i were calculated by the Fourier Series Method, the Iterated Linear Method, and all three Nonlinear Methods, using data from 'A' and then 'B'. This procedure was repeated for several different applied loads,

and each time, ten wrapped fringe patterns were determined using Equation 5.12 with the δ_i from each of the five methods using reference data along both 'A' and 'B'.

Diagonal strips of data were extracted from these images, unwrapped, and fitted to the theoretical model so that the rms error could be determined for each of the cases. A typical strip of data and the theoretical model are illustrated in Figure 5.13. Not only was it surprising to find no significant differences between the various methods, but the results indicated that with $N = 4$, an accuracy of not better than $\lambda/24$ could be achieved. In all cases, the actual errors appeared to be distributed randomly, and no detectable $\cos(2\phi)$ trend was observed. In addition, the maximum discrepancy between the δ_i determined during the experiments was only 11°, so it was concluded that the limit of $\lambda/24$ was not dominated by errors in estimating the phase shifts.

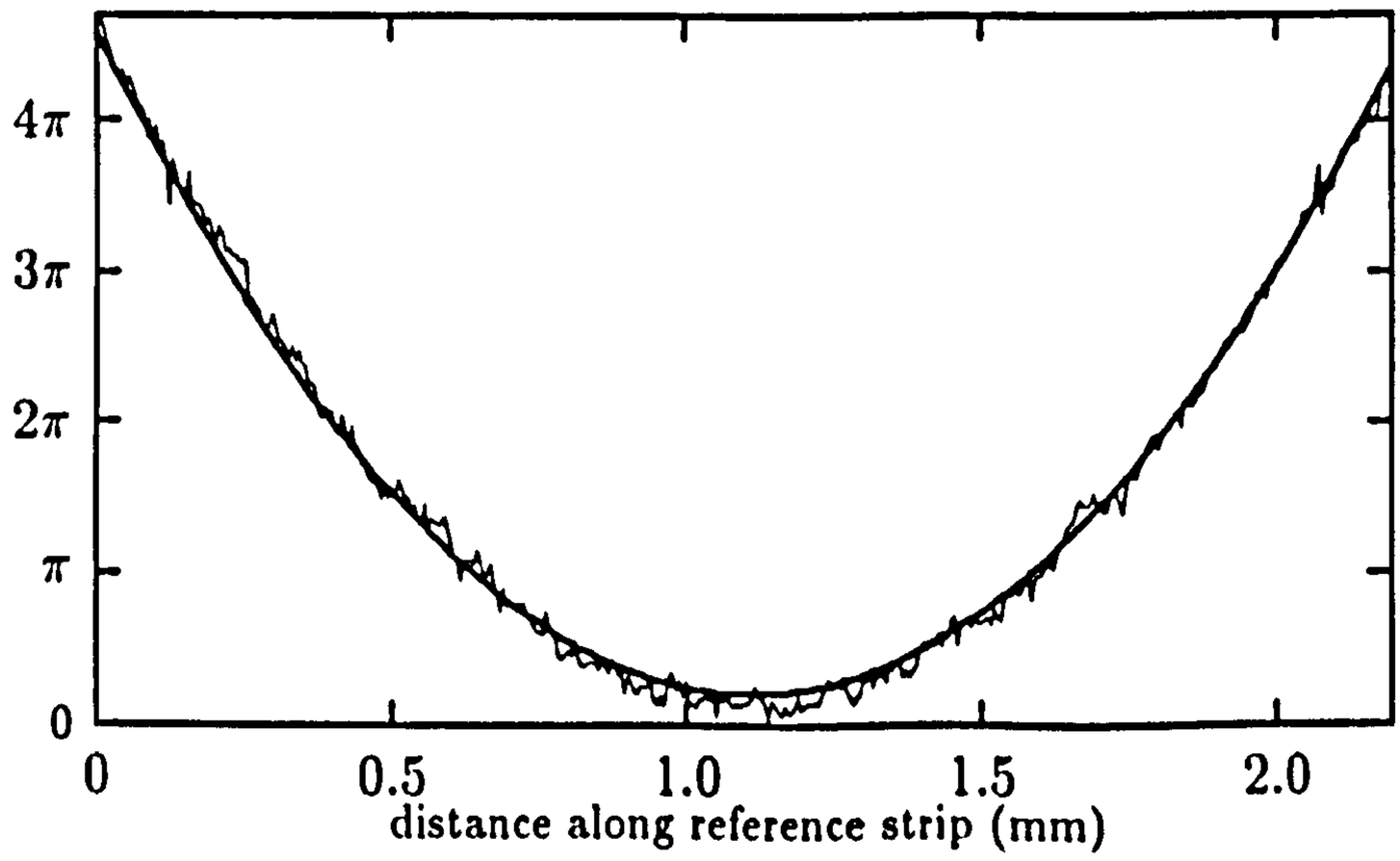


Figure 5.13: Comparison of experimental and theoretical displacements for the experiment described.

This poor result was attributed to the relatively noisy diffraction grating used in the experiment. Because an interest to preserve the integrity of surface strain gradients motivates this work, photoresist gratings that are holographically exposed directly on the specimen surface, were regularly used. Thus, gratings are fabricated with a thickness of no more than a few microns, at the expense of

N	Photoresist Grating	Epoxy Grating
4	$\lambda/24$	$\lambda/35$
8	$\lambda/45$	$\lambda/58$
16	$\lambda/50$	$\lambda/64$

Table 5.2: Experimental rms error for the solid mechanics three-point bend example.

having to tolerate high noise levels due to the inability to polish the specimen surfaces adequately. Conventionally, diffraction gratings that are imposed on the specimen surface using an epoxy or silicone cast, provide much lower noise levels.

Some additional experiments were performed to verify these findings. If the random grating noise was truly the limiting error source, then taking more images should help to improve the precision of the measurement. The above experiment was repeated, but this time more images were used ($N = 8$ and $N = 16$), and as expected, a significant improvement was attained. Next, a new specimen was prepared using the conventional epoxy casting method, and the same experiments performed. Table 5.2 summarizes these results, and it can be seen that the data obtained from the epoxy gratings does indeed show an improvement over the data from the photoresist gratings.

5.5 The Optical System

5.5.1 The Conventional Three Mirror Moiré Interferometer

The commonly used three mirror system was adopted for this work for several reasons. This simple interferometer design required only one collimated six inch beam for illumination, and provided for both the u- and v-field interferograms.

In the standard arrangement, collimated laser light simultaneously illuminates the specimen and the three mirrors. This is illustrated in Figure 5.14. The u-field interferogram is obtained by interfering the diffracted beams A and B, while the v-field interferogram is a result of the interference between beams C and D.

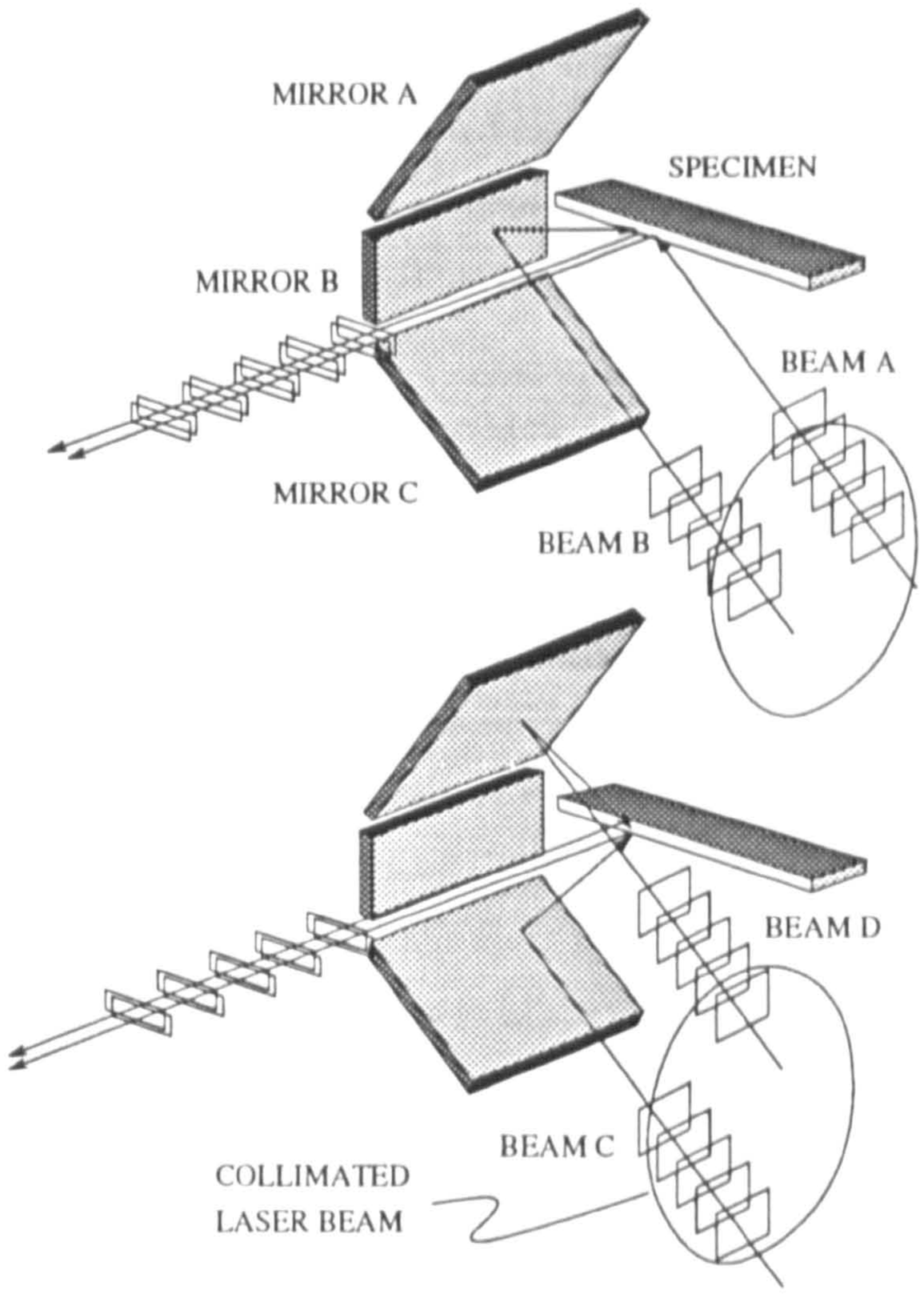


Figure 5.14: The standard three mirror arrangement

Construction and assembly of the interferometer was relatively simple Three laboratory grade ($\lambda/4$) flat mirrors were first mounted on standard two-axis mirror mounts (J.A. Knoll Co., model # 52–54). The mounted mirrors were then positioned on an aluminium base, which was machined roughly in the form of an angle bracket. This entire assembly was then mounted on an optical post, and held to the table using a magnetic base.

A field switching beam block was placed in the path of the collimated laser beam so that either the central mirror or the upper and lower mirrors could be illuminated. In this way, the separate u- and v-field interferograms were viewed.

It was necessary to insure that the field switching beam block correctly masked the beam, since any stray light leaking through to the camera would produce spurious interference patterns.

The initial alignment of the system began by calculating the necessary angle of

illumination using the equation $\lambda = d\sin\theta$, where d was the inverse of the grating frequency, and θ was the diffraction angle measured from the outward normal of the specimen. The necessary optical hardware for producing an expanded and collimated laser beam would then be aligned on the optical table to provide the illumination at the required angle. The four diffracted beams would then be 'tweeked' into alignment.

The three mirror interferometer provided excellent coverage at values of θ around 60°, permitting a field of view of up to $2\,\text{cm}^2$. This was the case when the HeCd laser was used ($\lambda = 442.5\,\text{nm}$) with 2000 lines/mm gratings. When the grating frequency was changed to 1200 lines/mm, and the HeCd laser used, θ was reduced to 30°, and the field of view was only $1\,\text{cm}^2$.

A more serious problem was encountered when using the three mirror system at 30°, however, that was an artifact of the polarization of the beams. In order for two beams to interfere with one another, it is necessary that they have the same polarization. In the case of the u-field, both beams would travel the same optical paths, except for the one additional reflection experienced by beam B. Since the light emitted from the laser was linearly polarized, with an axis exactly normal to the optical table, this additional reflection had no detrimental effect.

In the case of the v-field, however, the upper and lower mirrors would rotate the polarization of the incoming beams in opposite senses, so that when they were made to intersect, their polarizations were not aligned. Because the amount of polarization rotation was a function of θ (due to the geometry of the mirrors), this problem was almost undetectable at large values of θ, becoming more severe as θ was reduced.

Thus, in order to utilize 1200 lines/mm gratings, the HeCd laser had to be replaced with a HeNe laser ($\lambda = 632\,\text{nm}$), so that θ was increased to 50°.

5.5.2 Modifications for Phase Shifting

Central to the development of the phase shifting moiré interferometer used in this work were the modifications rendered to the conventional three mirror system discussed above.

To provide phase shifting capability, mirrors B and C were mounted on a high-quality translation stage driven by a piezo-electric (PZT) ceramic. See Figure 5.15.

The PZT ceramic responds to voltages supplied by a personal computer during data acquisition. Thus, mirrors B and C could be moved with a sub-micron resolution, and the path length of beams B and C were varied accordingly. Path length differences at this order of magnitude result in phase changes between the two beams that will be constant across the illuminated field.

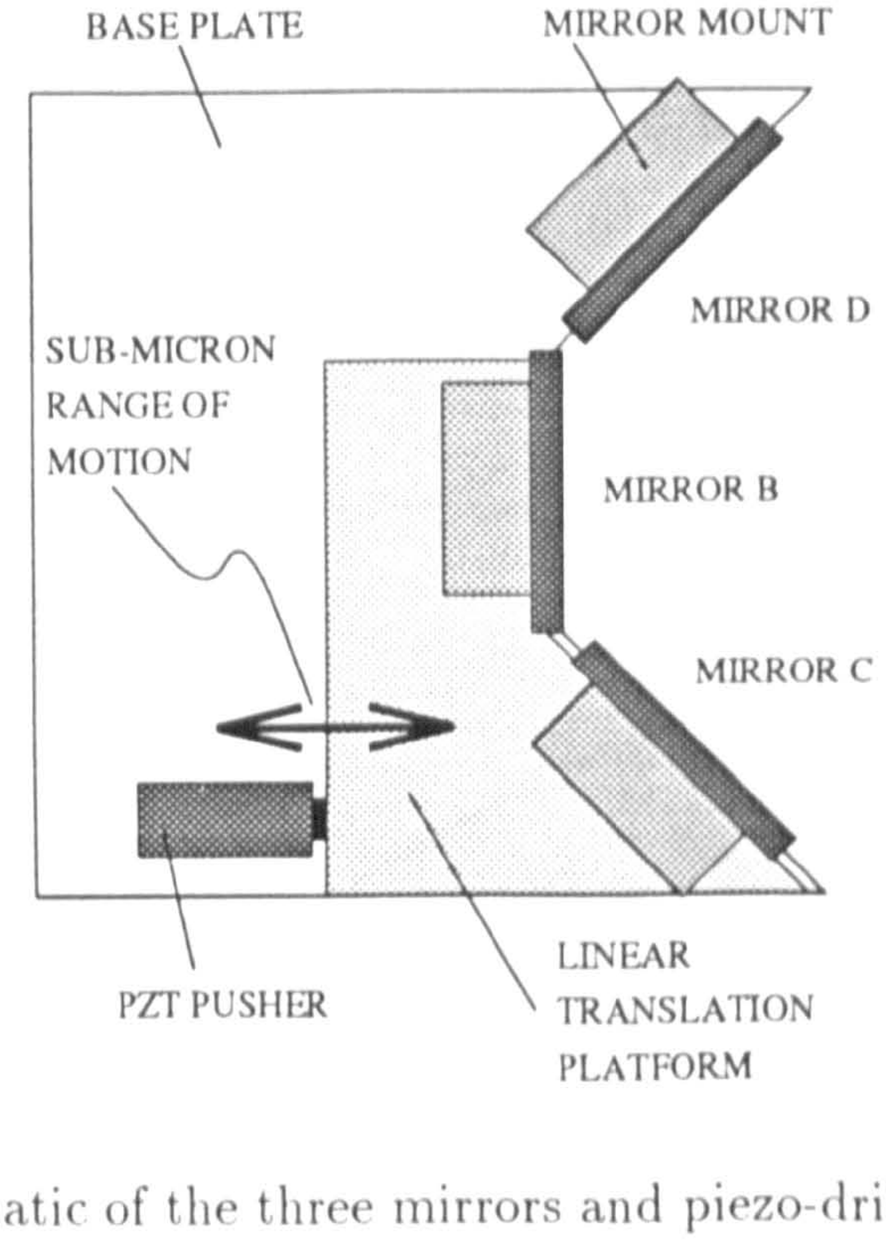


Figure 5.15: Schematic of the three mirrors and piezo-driven translation stage

The phase of the beams was deliberately shifted during the data acquisition process by ramping the voltage to the PZT ceramic in a controlled manner. This was achieved by converting a predetermined array of digital voltage values into analog values using a D/A converter installed in the laboratory personal computer. Because the PZT device was designed to operate in the 0–150 V range, an arbitrary voltage offset of 60 V was placed in series with the D/A output. This insured that the PZT was being operated in its most linear regime.

The translation of the stage imposed a path length difference into each pair of interfering beams according to

$$\delta = \frac{4\pi d}{h\lambda} \cos \alpha, \tag{5.15}$$

where d is the linear translation of the stage, α is the illumination angle, λ is the wavelength of the illuminating light and $h = 1$ for the horizontal field and $\sqrt{2}$ for

the vertical field. During the course of the experiments, two lasers with different wavelengths were used. Also, diffraction gratings with different frequencies were utilized. Since the illumination angle, α, is a function of grating frequency, and Equation 5.15 is dependent on both α and λ, different arrays of digital voltages were necessary for each combination of laser and grating frequency.

For each combination, a rigorous calibration routine was performed to ensure linearity, and simple periodic checks ensured that the system was behaving appropriately. A typical set of voltage values for $\lambda = 442.5\,\text{nm}$ and a 2000 lines/mm grating are shown in Figure 5.16.

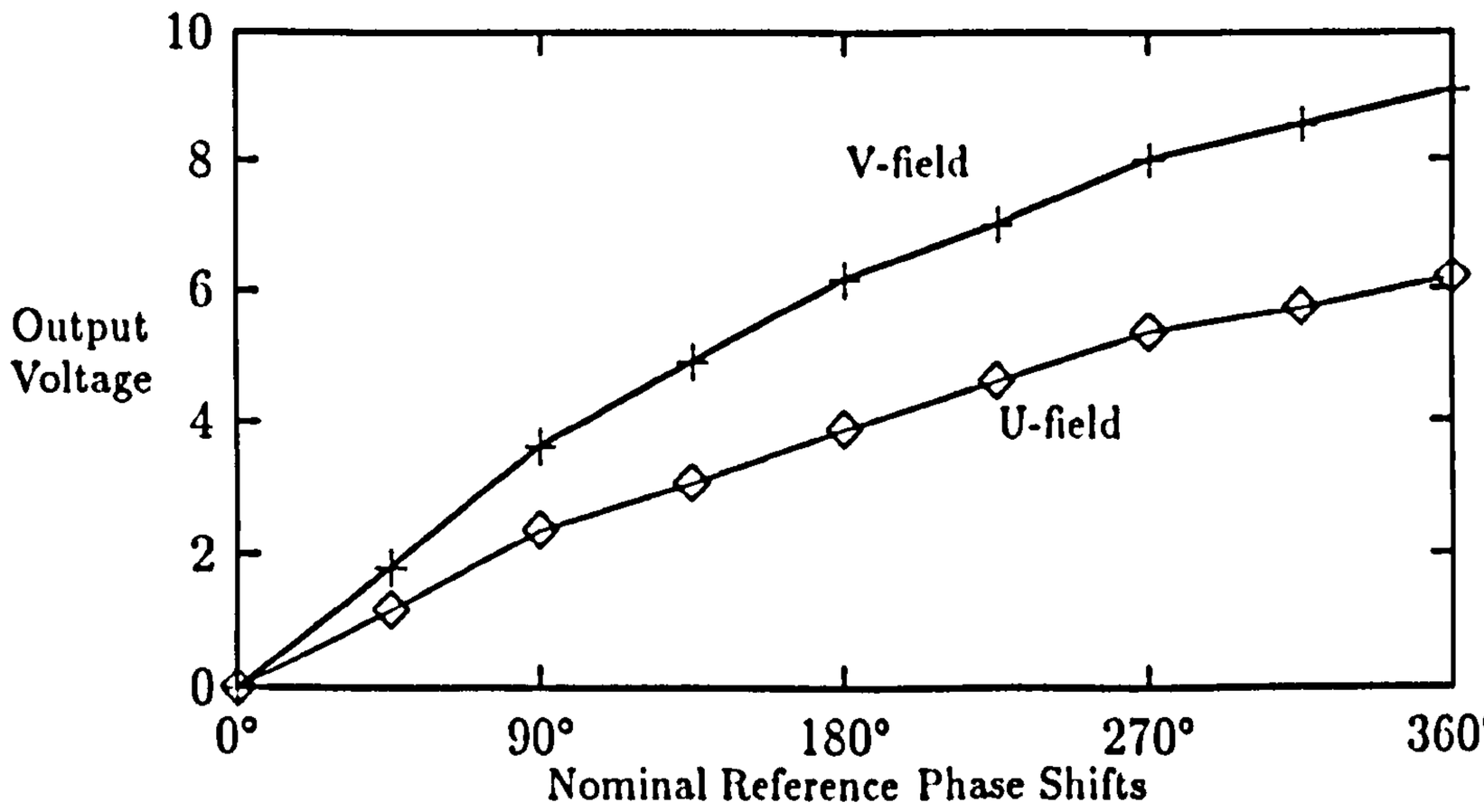


Figure 5.16: Typical voltage values for the phase shifter.

5.5.3 Imaging and Recording of the Fringes

The fringe patterns were imaged directly onto the exposed CCD array of a conventional video camera using a standard field lens. This arrangement provided a wide range of possible magnifications and was easy to align.

The field lens was made by Zeiss Jena and had a focal length of 16.5 cm and an f-stop of 2.7. The geometry of the imaging system is shown in Figure 5.17, and Table 5.3 summarizes the important dimensions for the range of magnifications used in this work. The optimal imaging ratio for the lens was 1:1, but very little degradation of the image was observed even at magnifications of 1:15. Beyond

this limit, aberrations became noticeable.

Also shown in the last column of Table 5.3 are the nominal dimensions for a single pixel, in real data units. Thus, at an optical magnification of 4 (imaging ratio of 1:4), 1 pixel represents a 6.3×4.7 μm rectangular region of the specimen surface. The difference between horizontal and vertical dimensions reflects the 4:3 aspect ratio of the video detector array.

One particular advantage of this imaging arrangement was the relatively long working distance between the lens and the camera. This facilitated the initial alignment of the three mirror system, since the diffracted plane waves from the surface of the specimen would come to a focus in the back focal plane of the lens.

Figure 5.18 schematically illustrates this general imaging property of lens. This is useful for the initial alignment of a moiré system, since obtaining the null pattern basically involves determining the appropriate mirror angle such that the interfering plane waves are coincident with one another.

The video camera used for this work was a Sony model TLC-450, with an external control/power unit. This camera allowed the setting of γ to unity, which is crucial for the deployment of PSI, since the basic equations depend on a linear relationship between the incident intensity and the recorded values.

An external gain control proved to be another essential feature, allowing for the precise adjustment of the image characteristics so as to permit full utilization of the available detector range.

The camera and lens were both mounted on a movable optical rail. This simplified the alignment procedure by insuring that both elements were located along the optic axis. Additional mechanical positioning devices were incorporated into the mounting of the lens and the camera, so that each could be located as necessary.

It was necessary to take great care during the alignment procedure to insure that the plane waves travelling from the specimen surface were passing through the centre of the lens. Both the u-field and the v-field beams were made to come to the same focus in the back focal plane of the lens, and this spot was verified to be on the optical axis, as defined by the optical rail. Incorrect alignment of the system resulted in a poor focus, particularly at the higher magnifications.

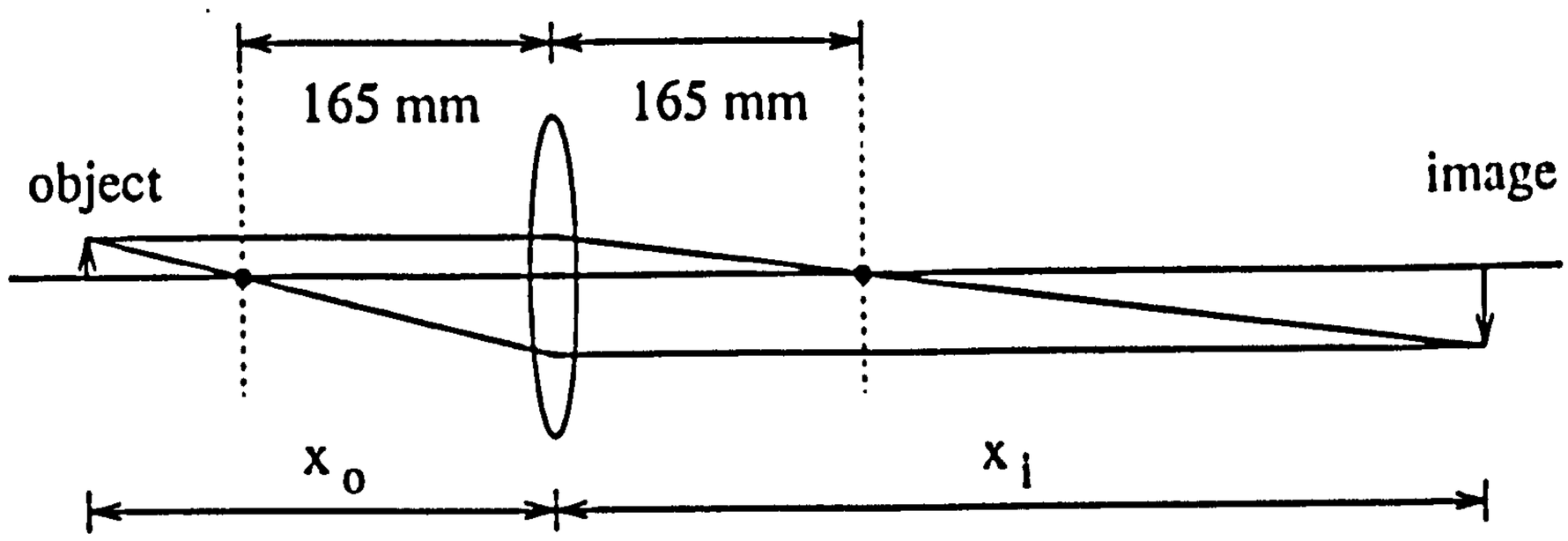


Figure 5.17: Geometry of the imaging system.

ratio	x_o (mm)	x_i (mm)	pixel (μm)
1:2	248	495	12.4×9.3
1:4	206	824	6.3×4.7
1:5	198	990	4.9×3.7
1:6	192	1155	4.1×3.1
1:8	185	1480	3.1×2.4
1:10	182	1815	2.5×1.9
1:15	176	2640	1.6×1.2

Table 5.3: Critical dimensions for various imaging ratios.

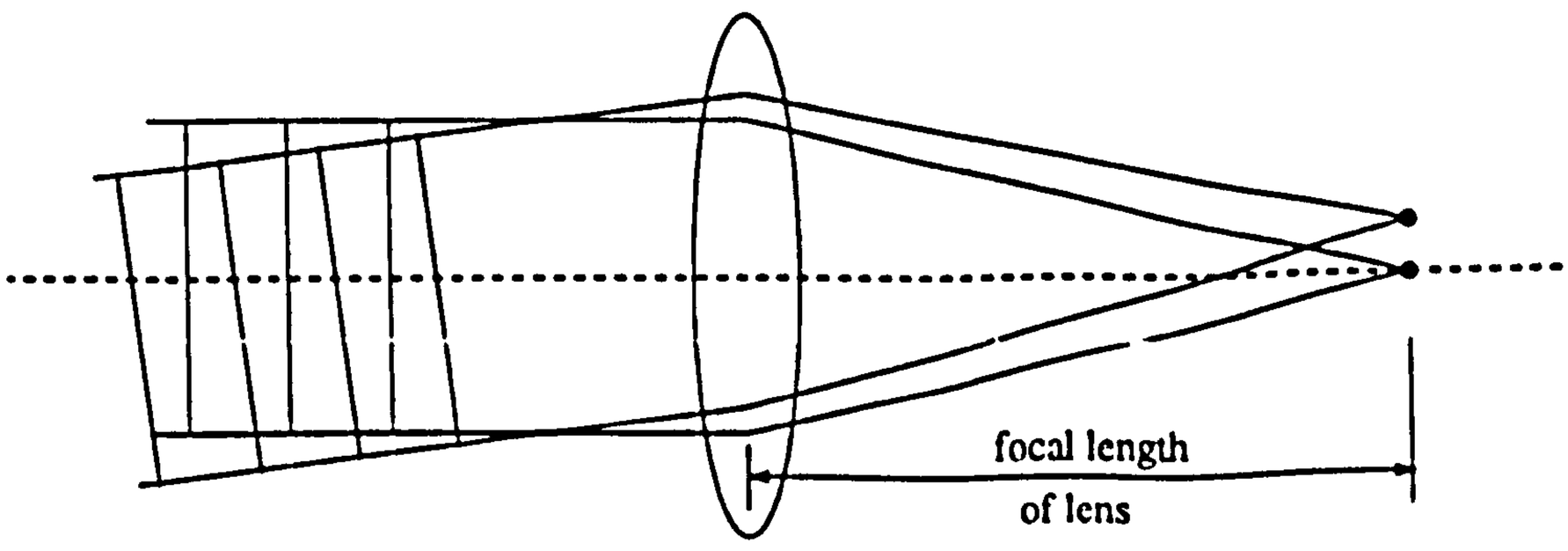


Figure 5.18: How a lens maps incoming plane waves to points in its back focal plane.

5.5.4 Software Control and Data Acquisition

A centralized, menu driven data acquisition system was responsible for both the capture of images and the shifting of phase, as well as recording load displacement data. The system developed for this work, called XMELT, also performed the operations necessary to produce a wrapped fringe pattern from the separate phase shifted interferograms. In this section, the details of the image capture and processing are described.

A Data Translation DT2861 frame grabber with 4 on-board frame buffers was used for this work. The frame grabber digitized a field of 512×512 pixels to 256 grey levels. The routines necessary to acquire and manipulate images were provided by the manufacturer as software libraries that were called directly from the XMELT programme.

The frame grabber was hardware connected to a Data Translation DT7020 array processor, which was used to process the images and produce the wrapped fringe pattern. The array processor allowed very fast operations to be performed on large vectors of data. Because the PSI data processing is performed on a pixel-by-pixel basis, an entire image could be transferred to the array processor as a vector of length 262144 (512×512). The array processor had enough local memory to store up to three full images at one time.

Implementing the PSI phase extraction procedures on the array processor was a relatively simple task. To accomplish this, individual array processor modules were written to perform basic operations such as addition, subtraction, multiplication, etc. A higher level function would then transfer the images to the array processor in the correct order, and evoke the necessary array processor modules to perform the computations. Finally, the $\tan^{-1}$ was computed on the array processor using the array processor's `atan2(y,x)` routine, which uses the signs of `x` and `y` to return the $\tan^{-1}$ value between $-\pi$ and π, instead of in the range of $-\pi/2$ and $\pi/2$.

The wrapped fringe patterns were available within 4 seconds, when Equation 5.13 was used, and within 10 seconds with Equation 5.12. Also during the data acquisition process, it was necessary to compute what became known as the

modulation mask, which for the $N = 4$ case (Equation 5.13) was defined as

$$B_j = \sqrt{(I_{3j} - I_{1j})^2 + (I_{0j} - I_{2j})^2}$$

and for the general least squares algorithm (Equation 5.12)

$$B_j = \sqrt{c_j^2 + b_j^2}.$$

These images were arbitrarily scaled between 0 and 255, and displayed to the user after the wrapped images were computed. Due to the nature of the PSI phase extraction equations, the modulation mask proved to be an invaluable guide in determining whether the freshly wrapped fringe pattern was corrupted by spurious phase shifts. By using the wrapped fringe pattern and the modulation mask, it was possible to reject the wrapped image if necessary, and perform the data acquisition again.

A typical data acquisition session proceeded as follows:

1. the video camera gain control was adjusted by monitoring a real time display of the histogram of the image intensities

2. the phase shifter was returned to its nominal position, and the user was presented with a real time display of the fringe pattern

3. the first frame was captured at a signal from the user (hitting a key on the keyboard)

4. if the image was not disturbed by vibrations (as deemed by the user), then it would be stored in the frame buffer, and the user would proceed to step 5, else, the user would return to step 3

5. the PZT voltage was ramped to produce the necessary phase shift

6. the user again would initiate an acquisition, again with the option of rejecting corrupted images

7. after the second frame was captured, the phase would be shifted and this process repeated until all of the necessary frames had been acquired

8. the user would browse through the captured images, and decide to process them or acquire a new set (back to step 2)

9. if it was decided to process the images, a wrapped image and a modulation mask would be computed on the array processor

10. the user could then make a decision as to the quality of the wrapped fringe pattern, and elect to save the wrapped image, or return to step 2

It was imperative to allow the user the freedom to abort either a part of or the entire acquisition process, since there was a subjective element in judging the quality of the fringe patterns. Uncontrollable external disturbances (vibrations) continually plagued the acquisition process, and the only way to defeat this problem was to develop a system that provided flexibility and a means to asses the quality of the data at various steps.

At the end of an experiment, all of the wrapped images were transferred to a Silicon Graphics workstation for the remaining unwrapping procedure and post-processing operations.

5.6 Phase Unwrapping

The PSI extraction procedure produces a noise rejected fringe pattern from a set of phase shifted images. However, due to the nature of the $\tan^{-1}()$ calculation, the resulting phase values are wrapped between $-\pi$ and π. It is a relatively simple matter to correct the wrapped fringe patterns and this process is known as phase unwrapping.

The basic idea behind phase unwrapping, or phase integration, is that an integer number of cycles must be added to or subtracted from each wrapped phase value. This is needed to restore continuity to the phase values that are implicitly 'wrapped' by the $\tan^{-1}()$ calculations of Equation 5.12 or 5.13.

As a simple example, consider the following one-dimensional sequence of wrapped phase measurements where the units are in cycles, and the phase, ϕ, can only assume values in the range $0 \leq \phi < 1$

0.5, 0.6, 0.7, 0.8, 0.9, 0.0, 0.1, 0.2, ...

It is clear that one cycle should be added to each of the last three values to restore continuity to the entire sequence. Thus, the unwrapped phases would be

$$0.5,\ 0.6,\ 0.7,\ 0.8,\ 0.9,\ 1.0,\ 1.1,\ 1.2,\ \ldots$$

The above example illustrates that the unwrapping process represents a trivial operation, that lends itself easily to automation. Occasionally, however, ambiguities may creep into the wrapped phase measurements at the local pixel level and introduce errors that will be propagated down the line. An additional procedure should be performed to detect and isolate these ambiguities, thereby preventing them from affecting the unwrapping algorithm.

Local ambiguities are introduced when the signal[3] of interest becomes partially or completely obscured. This can arise due to a damaged diffraction grating, (a scratch, finger print, etc.), or it may occur in high-fringe gradient regions, where the spacing between fringes become too small. The latter situation represents a physical limitations of resolving fringe patterns using video technology, and is discussed in more detail in the next section.

5.6.1 The Sampling Theorem

There are many versions of the sampling theorem, but basically it may be stated as follows:

> *"Any band-limited function can be specified exactly by its sampled values, taken at some regular intervals, provided that these intervals do not exceed some critical sampling interval."* [48]

In the context of PSI, the sampling theorem manifests itself during the digitization of the individual phase shifted fringe patterns. If the distance between fringes approaches the dimension of the individual pixels, aliasing will inevitably occur, and ambiguities will be introduced. The intensity values that are recorded at these points will *not* yield reliable wrapped phase values, since the sampling theorem has been violated.

---

[3]In this discussion, the term signal will refer to the point-by-point intensity variation of the fringe pattern as detected by the imaging chip in the video camera.

It is well known [55], that the level of noise rejection afforded by PSI algorithms is dependent on the local fringe density within an interferogram. This can be explained by considering the relative dimensions of the fringes and pixel sizes, without recourse to the sampling theorem.

Because pixels have a finite dimension, some degree of averaging will accompanying every measurement of fringe pattern intensity. When the spacing between fringes is much larger than the pixel size, the errors introduced by this averaging will be small, since the intensity variation of the fringe over the finite pixel dimensions will be limited. However, the magnitude of these errors will gradually increase as the fringe spacing becomes smaller, and the resulting sinusoidal intensity variation between fringe peaks changes more rapidly.

5.6.2 Detecting Ambiguities in the Fringe Pattern

A direct consequence of the sampling theorem is that a fundamental assumption must be made regarding the wrapped phase measurements. Between any two sampled points, the true phase must not change by as much as one-half cycle. If the true phase changes by more than this amount, the signal will be *undersampled*, and the resulting wrapped phase value will be unreliable.

Other situations can arise that can complicate the phase unwrapping procedure, such as when a true discontinuity exists in the true phase values. This will occur when the phase values represent the displacements on opposite sides of a crack. The algorithm should respect this discontinuity, and unwrap the phase values accordingly.

High fringe gradient regions are confined to a small portion of a fringe pattern, and in principle, they can be removed from the image by an appropriate 'masking' operation applied by the user. But, these high fringe gradient regions are generally of most interest in experimental mechanics investigations, so masking of the data would have to be done carefully, so as to preserve as much information as possible.

Since unwrapping proceeds on a point-by-point basis, any errors that are introduced during the unwrapping process will inevitably propagate from that point on. The result is that the local error will become global, affecting the entire unwrapped fringe pattern.

Fortunately, *all* types of ambiguities can be detected *automatically*, and ap-

propriate measures can be taken without the need for the user to interact with the data directly. To understand how this is done, it is first necessary to clarify what constitutes an *ambiguity* in a set of phase measurements. A simple example from Goldstein *et al.* [72] will help.

Consider the following two-dimensional set of 'noisy' phase measurements:

0.0	0.1	0.2	0.3
0.0	0.0	0.3	0.4
0.9	0.8	0.6	0.5
0.8	0.8	0.7	0.6

For the case of two-dimensional data, it is necessary to unwrap in both directions—i.e. along the rows and then down the columns. Using the rule that no two phase measurements can differ by more than one-half cycle, and scanning along the rows of the data, it can be seen that no corrections are necessary.

Scanning down the columns, however, results in the following complete solution:

0.0	0.1	0.2	0.3
0.0	0.0	0.3	0.4
-0.1	-0.2	-0.4	-0.5
-0.2	-0.2	-0.3	-0.4

Note that there are two values that differ from their vertical neighbours by more than one-half cycle, indicated by the horizontal line. This is a contradiction that violates the 'differ by one-half cycle' assumption.

If we return to the original data and scan down the columns first and then across the rows, we obtain:

0.0	0.1	0.2	0.3
0.0	0.0	0.3	0.4
-0.1	-0.2	0.6	0.5
-0.2	-0.2	0.7	0.6

Not only are the unwrapped phases different, but the location of the contradiction is changed. This inconsistency is an inherent property of the data set, and can not be eliminated by any different method of scanning. Different solutions will be obtained for different paths taken during the unwrapping. The problem then, is

to determine the cause of the inconsistency and to develop a procedure to avoid it.

Goldstein *et al.* suggested that the source of such an inconsistency could be detected by evaluating the sum of the phase differences clockwise around each set of four adjacent phase measurements. The result can be only one of three values, +1, −1 or 0. A value of 0 indicates that the path travelled began and ended at the same phase value. A result of ±1 indicates that an erroneous phase step was encountered along that particular path. This test of the data is analogous to demanding that one begin and end a hill-walk at the same elevation.

Consider the following set of four points chosen from the center of the original data:

$$\begin{array}{ccc} 0.0 & \rightarrow & 0.3 \\ \uparrow & & \downarrow \\ 0.8 & \leftarrow & 0.6 \end{array}$$

Beginning at the top-left hand corner and travelling clockwise we find that no adjustment is needed from 0.0→0.3, 0.3→0.6 or 0.6→0.8. But the last step, 0.8→0.0, requires that we add one cycle to respect the 'differ by one-half cycle' rule. Thus the net phase around the four points is +1 cycle.

If the same test is applied to each set of four-point circuits possible in the original data, then the following is obtained:

$$\begin{array}{cccc} 0.0 & 0.1 & 0.2 & 0.3 \\ & o & o & o \\ 0.0 & 0.0 & 0.3 & 0.4 \\ & o & +1 & o \\ 0.9 & 0.8 & 0.6 & 0.5 \\ & o & o & o \\ 0.8 & 0.8 & 0.7 & 0.6 \end{array}$$

The small numbers represent the net resultant cycles, and will be called 'residues'. It should be emphasized that a residue is *not* associated with a single point, rather, it is derived from the topological relationship between (at least) four pixels.

It has been shown by Goldstein *et al.*, and independently by Huntley [73],

that any phase unwrapping path that encloses a single residue will produce an inconsistency in the unwrapped phase. However, if a path encloses an equal number of plus and minus residues (the sum of all of the residues is zero), no inconsistency results. Thus the phase measurements can be unwrapped in a consistent manner if the residues are identified and suitable 'branch-cuts' are made between the residues to prevent any unwrapping path from crossing these cuts.

Huntley [73] presented essentially the same residue detection and isolation algorithm as that of Goldstein *et al.*, but has implemented the algorithms differently. Bone [53] also discusses the method of using residues, but implemented a fundamentally different masking scheme based on the second differences calculated from the locally unwrapped phase measurements.

5.6.3 Isolating Ambiguities in the Fringe Pattern

It is a relatively simple matter to determine a table of residues for any particular unwrapped fringe pattern. This can be done by scanning the fringe pattern and computing the residues for each four-pixel circuit, while storing the result in a suitable data structure. Once the table of residues has been determined, it remains to introduce branch-cuts between the residues.

The only purpose of the branch-cut is to keep the unwrapping algorithm from attempting to unwrap along a path that would enclose a finite number of unpaired residues. Therefore, it is sufficient to 'mask' a path between two residues, so that these points will be excluded from the unwrapping procedure. This is illustrated in Figure 5.19 for two types of branch-cuts. Residue near a boundary of the image, can simply be connected to it.

The implementation of the complete branch-cut routine is based on a recursive algorithm, and proceeds as follows:

1. The residue 'state' is queried for each point in the image by scanning along each consecutive row. When a ± 1 residue is found, the scanning is temporarily suspended.

2. A box of size three is placed around the new residue location and a search is initiated for the nearest neighbouring residue.

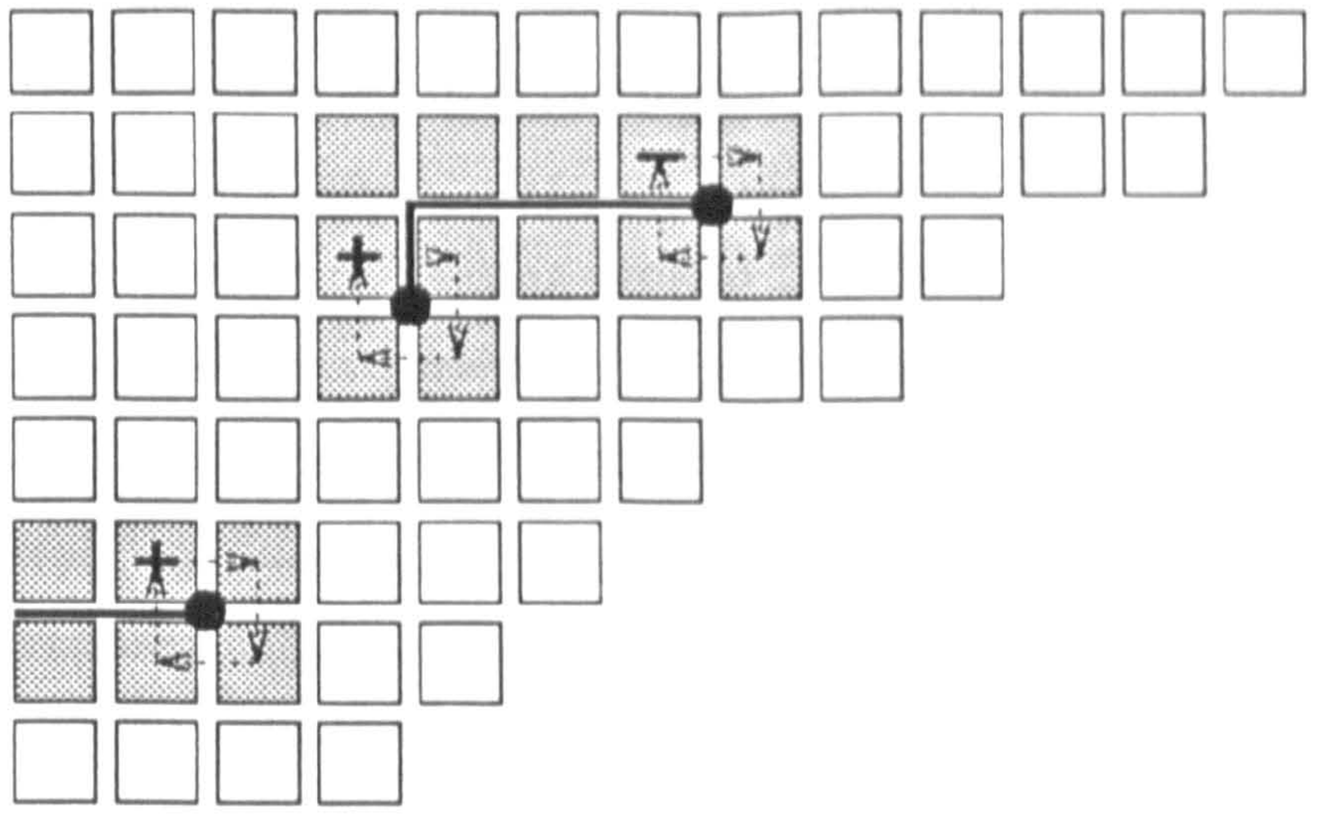


Figure 5.19: A schematic illustration of the branch-cut scheme adopted for this work. The shaded pixels would have had their mask value set so that the flood-fill unwrapping algorithm would ignore them.

- If a residue of the same sign is found, a branch cut is placed between the two residue locations, and a *new* search is initiated. This means that an equal number of oppositely residues must be found before the current branch-cut is complete.
- If no residue is found, the box is extended by one pixel on each side, and the search is extended to this new area (back to 2, but there is no new invocation of the recursive function).
- If a residue of the opposite sign is found, or if the sum of all the residues connected by the current branch-cut becomes zero, a branch-cut is placed between the two locations, and the row-by-row scanning is resumed (back to 1).
- If a boundary of the image is encountered, then the branch-cut will be extended to it and the row-by-row scanning will be resumed (back to 1).

5.6.4 Flood-Fill Unwrapping

After all of the necessary branch-cuts have been constructed, the phase measurements can be unwrapped along any path, producing the same result. The unwrapping could proceed row-by-row first, and then column-by-column, or vice

versa. However, because the algorithm must respect the branch-cuts, an approach based on a row-by-row or column-by-column will be difficult to implement.

The most straightforward procedure would be one that chose its own path through the data, while navigating around branch-cuts when they are encountered. This is precisely the case with a typical flood-fill algorithm. The flood-fill algorithm is recursive, meaning that it performs the unwrapping by repeated calls to the same software function.

The unwrapping begins when an arbitrary 'seed' point is chosen. Each of the four pixels directly surrounding this pixel will then be checked in turn, to see if they belong to the set of pixels that are to be unwrapped. (A pixel will belong to this set if it has not had its 'mask' value set during the branch-cut operation.) If any of the surrounding pixels belong to this set, then they will be unwrapped relative to the 'seed' pixel. But, before they are actually unwrapped, each of *their* four neighbours will be checked in turn, to see if they belong to the set of pixels to be unwrapped. But, before they are actually unwrapped, each of *their* four...

In this way, each of the pixels that have not been excluded by the branch-cuts will be unwrapped relative to the 'seed' pixel. The primary advantage of the flood-fill approach to phase unwrapping is that is simple to programme, and easy to understand.

5.6.5 Automated versus Automatic Fringe Processing

The idea of a completely automatic fringe processing system is attractive, although it is probably unrealistic. Each application of interferometric fringe analysis has its own unique set of objectives and peculiarities, and these bear directly on how the fringe processing should proceed.

The intimate melding of video technology and computer control that forms the basis of the PSI approach does make the task more manageable, but still, occasions arise where a human being must intervene. As discussed previously, no amount of software sophistication can replace the unique human capacity to judge the quality of fringe patterns during the acquisition.

Even after the data has been acquired, there are several preventative measures that a human can render to the data that will make the subsequent processing proceed much more efficiently. Furthermore, these steps can anticipate the eventual

post-processing of the experimental displacement data, such as differentiation to obtain strains or the calculation of conservative contour integrals.

As an example, consider the extraction of shear strains from conventional moiré displacement data. Both the u- and v-field must be available so that $\partial u/\partial y$ and $\partial v/\partial x$ can be evaluated. It is imperative to insure that data from each unwrapped fringe pattern is taken relative to a common coordinate system. The most convenient time to establish this coordinate system is *prior* to the unwrapping, but after the actual experiment. Another practical matter that must be dealt with at some time in the data reduction process is the scaling of the images from arbitrary pixel units into the actual dimensions of the specimen.

The approach adopted for this work integrates these necessary bookkeeping tasks with some preventative measures designed to improve the efficiency and quality of the unwrapping. A menu driven image editing programme was written so that several operations could be performed on the images prior to unwrapping. The following options were supported:

- The crack tip could be located, with an option to 'mask' the region behind the crack-tip where there was a finite gap between the upper and lower crack surfaces. Location of the crack tip would provided an absolute reference point that would be used when extracting displacement and strain data from the unwrapped image.

- The scale of the image could be set by defining the distance in real data units from the top of the image to the bottom. This would fix the scale for both dimensions, since the aspect ratio of 4/3 was known.

- A border could be applied to the image, if the entire field-of-view was not of interest.

- Regions of the fringe pattern containing significant noise could be 'masked' out by selecting them with the cursor.

- Several preprocessing filters could be applied to selected portions of the wrapped fringe patterns to reduce the noise level of the data

Only two of the above operations were absolutely necessary—locating the crack tip and scaling the image. Of the remaining, only the masking of the crack was

performed on a regular basis. Only in exceptional cases did the wrapped fringe patterns require additional editing.

Chapter 6

Analysis of Displacement Fields

In this chapter, some basic principles of solid mechanics are presented for the purpose of establishing a consistent notation for the remainder work. Since the focus of the analysis will be on interpreting displacement data from moiré interferometry, this section will begin with the coordinate system for the displacement measurements, and proceed to define strains and then stresses.

The details of extracting displacement and deriving strains from experimental fringe patterns will be discussed in Section 6.2. In Section 6.3, the definition and implementation of the J-integral will be covered.

6.1 Notation and Material Properties

The coordinate system used for the analysis of the experimentally determined displacement fields is shown in Figure 6.1. The precise location of the origin was determined independently for each fringe pattern, and was associated with the apparent crack tip location. Location of the crack tip was performed after the experiments but prior to unwrapping the fringe patterns.

The x- and y-axes of the coordinate system were, by necessity, assumed to be horizontal and vertical. There was no facility for rotating an image after it had been acquired, so it was imperative to insure that the delamination was located along a horizontal line *before* the image was acquired (by rotating the whole test rig).

The full-field displacements in the horizontal, or x, direction will be denoted $u(x, y)$, while the displacements along the vertical, or y, direction will be denoted

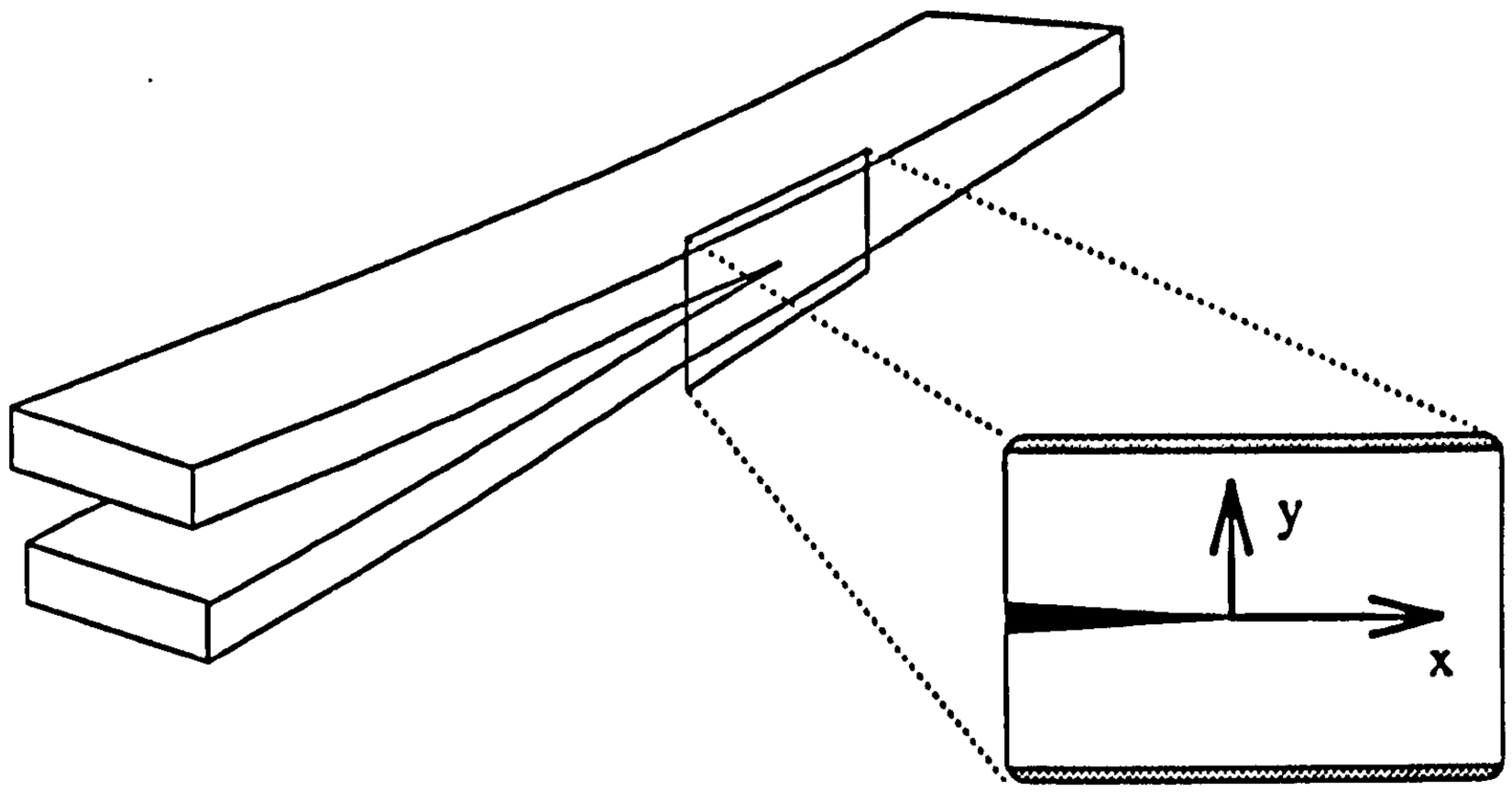


Figure 6.1: Coordinate system used for local displacement fields.

as $v(x, y)$. Each (x, y) pair is associated with a discrete pixel location.

Any particular fringe pattern represented *either* a u-field *or* a v-field, so it was necessary to acquire two fringe patterns, one for each field, during an experiment to characterize the complete state of deformation. In the remainder of the discussion, a *data set* will refer to a pair of u- and v-field fringe patterns.

Strains were derived from the displacements according to

$$\begin{aligned} \varepsilon_x &= \frac{\partial u}{\partial x} \\ \varepsilon_y &= \frac{\partial v}{\partial y} \\ \gamma_{xy} &= \frac{\partial u}{\partial y} + \frac{\partial v}{\partial x} \end{aligned} \tag{6.1}$$

The details of the implementation of the differentiation will be discussed in the next section.

Because the shear strain, γ_{xy}, depends on both $u(x, y)$ and $v(x, y)$, it was important to insure that no deformation or rigid body rotation occurred between the acquisition of the two fringe patterns composing a data set. It was also necessary to render compatible the crack tip locations between the the two separate fields. This was accomplished by computing an offset (in pixel units) between the two different crack tip locations.

Stresses were determined from the strains according to the relations for a

plane-stress orthotropic material, as follows:

$$\begin{Bmatrix} \sigma_x \\ \sigma_y \\ \tau_{xy} \end{Bmatrix} = \begin{bmatrix} Q_{11} & Q_{12} & 0 \\ Q_{12} & Q_{22} & 0 \\ 0 & 0 & Q_{66} \end{bmatrix} \begin{Bmatrix} \varepsilon_x \\ \varepsilon_y \\ \gamma_{xy} \end{Bmatrix}. \tag{6.2}$$

The Q_{ij} were determined as functions of the engineering properties and are given by:

$$\begin{aligned} Q_{11} &= \frac{E_x}{1 - \nu_{xy}\nu_{yx}} \\ Q_{12} &= \frac{\nu_{xy}E_y}{1 - \nu_{xy}\nu_{yx}} \\ Q_{22} &= \frac{E_y}{1 - \nu_{xy}\nu_{yx}} \\ Q_{66} &= G_{xy} \end{aligned}$$

The above definition of the Q_{ij} along with the material properties presented in Table 3.1 define the constitutive relation for a 0° unidirectional lamina with the fibres oriented along the x-axis. However, since 90° and 45° lamina were also encountered in the multidirectional specimens, it was necessary to interpret the engineering constants in Table 3.1 appropriately. For the 90° lamina, this meant reversing the roles of E_x, E_y, ν_{xy} and ν_{yx}. For the 45° lamina, it was required to transform the 0° material properties. A summary of the material properties used for each type of layer is shown in Table 6.1.

	0°	45°	90°
E_x (MPa)	128,615	13,834	9,672
E_y (MPa)	9,672	13,834	128,615
G_{xy} (MPa)	5443	5443	5443
ν_{xy}	0.476	0.527	0.036
ν_{yx}	0.036	0.527	0.476

Table 6.1: Material properties used for each type of lamina.

6.2 Extraction of Information from Fringe Patterns

Each fringe pattern acquired during an experiment represented a total of 262,144 individual displacement measurements (512×512 pixels). Under normal magnification, this meant that each pixel covered a 5.2×3.9 μm rectangular region of the specimen surface.

The noise rejection of the phase shift technique is unique in that there is no local averaging of the data, so the initial gauge length of the measurements was given by the dimensions of a pixel. However, a certain amount of noise was still present in the raw data, and it was necessary to perform additional smoothing operations. Specifically, there were two types of smoothing applied to the data. The first type of smoothing occurred during the initial extraction of the displacement and strain values, and was accomplished using digital filters.

The second type of filtering was performed after the displacement and strain data had been extracted from the raw data set, and is discussed in Section 6.3. The smoothing was performed in two distinct steps to preserve as much detail as possible, while at the same time, adequately conditioning the data before the differentiation process. The final balance obtained from the action of the two different types of smoothing was found to be superior to either of the two methods operating alone.

Digital filtering involves convolving the raw data with a carefully chosen 'kernel'. In the case of two-dimensional data, a typical kernel used to smooth data is:

$$\frac{1}{100}\begin{pmatrix} 1 & 2 & 4 & 2 & 1 \\ 2 & 4 & 8 & 4 & 2 \\ 4 & 8 & 16 & 8 & 4 \\ 2 & 4 & 8 & 4 & 2 \\ 1 & 2 & 4 & 2 & 1 \end{pmatrix}$$

This 5×5 kernel is the integer representation of a two-dimensional Gaussian function with its peak at the central position. Smoothing with this filter increases the effective gauge length by a factor of 5, but effectively attenuates the high frequency components of the data.

Differentiation can be performed by using a filter with anti-symmetric coefficients. A simple horizontal differentiator is given by the coefficients:

$$\begin{pmatrix} -1 & 0 & 1 \end{pmatrix}$$

This filter provides no smoothing, and will produce output that is highly susceptible to noise in the data.

It is possible to design a digital to achieve a specific frequency response [74]. In the case of differentiation, we want to respect low frequency components of the data, while attenuating the high frequency components—especially during differentiation. A better set of coefficients for differentiation based on a fitting a cubic spline to the data are:

$$\frac{1}{12}\begin{pmatrix} 1 & -8 & 0 & 8 & -1 \end{pmatrix}$$

By combining the above filter design with Gaussian smoothing in the vertical direction, a more robust horizontal differentiator is obtained:

$$\frac{1}{108}\begin{pmatrix} 1 & -8 & 0 & 8 & -1 \\ 2 & -16 & 0 & 16 & -2 \\ 4 & -32 & 0 & 32 & -4 \\ 2 & -16 & 0 & 16 & -2 \\ 1 & -8 & 0 & 8 & -1 \end{pmatrix}$$

The above coefficients were used for differentiation in the horizontal direction, while vertical differentiation was accomplished by transposing the coefficients.

The actual extraction of information from a data set usually occurred prior to the evaluation of a contour. A contour geometry would be chosen, and data along the contour would be extracted by using the appropriate digital filters. For $u(x,y)$ and $v(x,y)$ the 5×5 Gaussian filter was used, for the partials in the x-direction, $\partial u/\partial x$ and $\partial v/\partial x$, the 5×5 horizontal differentiator was used, and similarly, the 5×5 vertical differentiator was used for $\partial u/\partial y$ and $\partial v/\partial y$. As an intermediate step, these values would be stored in a data structure for each point along the contour to be evaluated.

In addition to the displacement and strain data, it was also necessary to

store information relating to the geometry of the contour. In particular, the (x, y) position and the two components of the outward unit normal vector to the contour, N_x and N_y, were retained. From this information, dy and ds could be recovered. After the strain data had been extracted from the raw data set, the stresses were obtained using the appropriate constitutive relations for each layer.

6.3 J-integral Analysis

The J-integral provided a convenient means for determining the linear elastic Energy Release Rate, G, from the stress and strain fields surrounding delaminations. Apart from the strain and stress fields obtained directly from the moiré interferometry fringe patterns, no additional information was required to compute J.

The definition of the J-integral is

$$J = \int_{\Gamma} (W n_x - T_x \frac{\partial u}{\partial x} - T_y \frac{\partial v}{\partial x})\, ds. \tag{6.3}$$

where Γ is any closed contour enclosing the crack tip, n_x is the x component of the outward unit normal to the contour and ds is the differential arc length around the contour. The strain energy density function, W, and the x and y components of the traction vector, T_x and T_y, are given by

$$\begin{aligned} W &= \frac{1}{2}(\sigma_x \epsilon_x + \sigma_y \epsilon_y + \tau_{xy} \gamma_{xy}) \\ T_x &= (\sigma_x n_x + \tau_{xy} n_y) \\ T_y &= (\tau_{xy} n_x + \sigma_y n_y). \end{aligned}$$

There were two contour geometries used during the course of the work—circular contours and rectangular contours. Circular contours were preferred over the rectangular ones because they were easier to implement and generally exhibited less susceptibility to noise. The rectangular contours were used only for the multidirectional specimens when the individual legs of the contour had to be kept within the thin 0.125 mm layer. In both cases, however, the data were obtained from the fringe patterns using the digital filters described above and stored in the same way.

The second smoothing operation was performed after the displacement and strain data had been assembled into a contour data structure. It involved fitting smoothing splines to the data, by using the position around the contour as the dependent variable. Smoothing splines differ from ordinary splines in that they are not required to go through each data point. They are controlled by a *tension factor* and a *smoothing factor*, and attempt to minimize an error functional based on the expected standard deviation of the data [75].

The purpose of this second smoothing operation was to further reduce the general noise level of the data and to force the stress data to conform to the appropriate boundary conditions. For example, in the case of a DCB specimen, both σ_y and τ_{xy} should be zero at the traction free surfaces behind the crack tip. Enforcing this constraint improved the quality of the J measurements, especially since data directly on the two crack surfaces had to be extrapolated.

In order to arrive at the appropriate coefficients for the smoothing splines (and the balance between the smoothing afforded by the splines versus the smoothing provided by the digital filters), a programme was written to compute $u(x, y)$ and $v(x, y)$ for 'ideal' data. This programme utilized the analytic solution for the displacements in the vicinity of a crack tip for a centre-cracked, infinite plate. The analytic model accounted for orthotropic material behaviour and generated output files equivalent to those obtained from unwrapped u- and v-field fringe patterns. This 'ideal' data was subjected only to a quantization error, as it was stored in the same format as a normal image, using only 8 bits per sample (0–255).

For the circular contours, the implementation of the smoothing splines was trivial, as the data to be smoothed were continuous and varied smoothly from $-\pi$ to π. Two examples of the smoothing process for circular contours are illustrated in Figures 6.2 and 6.3. The data set was generated using the analytic model described above, with $G = 100$ J/m^2, and displacements and strains were taken from a contour with a radius of 0.454 mm from the crack tip. The discrete data points extracted using the digital filters are shown as diamonds, while the smoothed functions are shown as solid lines.

An undesired consequence of smoothing the data is that peaks become rounded and sharp edges loose their definition. These effects can be seen in both Figures 6.2 and 6.3. The digital filter coefficients and the smoothing spline param-

eters were adjusted incrementally to arrive at the final balance between the two methods. Any more smoothing by either of the procedures resulted in a more drastic reduction in character of the data, while any less did not provide enough noise rejection. Ultimately, the smoothing parameters had to be chosen subjectively, but examination of the individual displacement, strain and stress plots helped to guide the process.

The path independence of the J-integral computation was verified using the 'ideal' data and circular contours. Approximately 20 different contours with radii between 0.1 mm and 0.7 mm from the crack tip were evaluated, with the results shown in Figure 6.4.

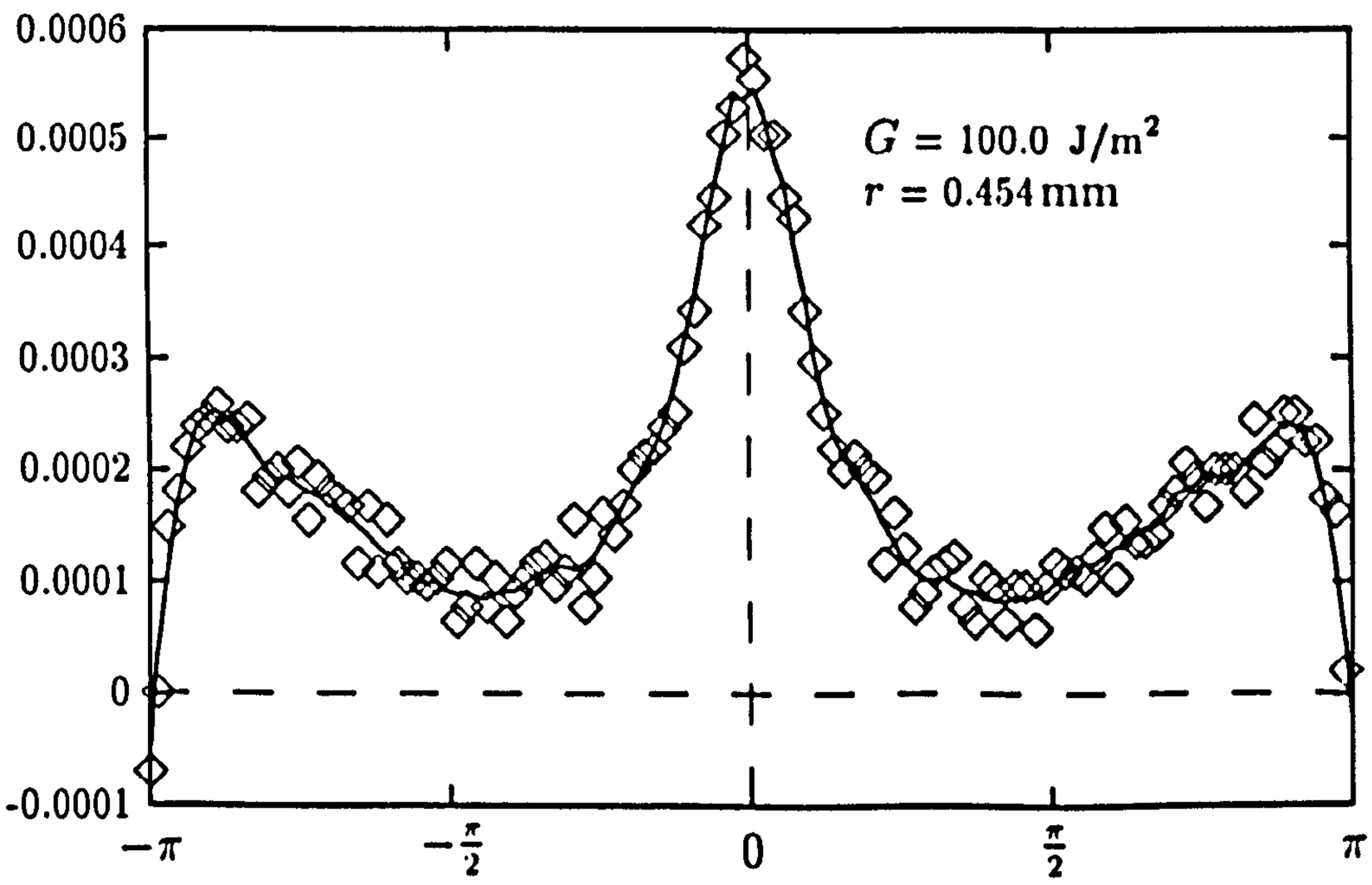


Figure 6.2: $\partial u/\partial x$ on a radial contour for 'ideal' data.

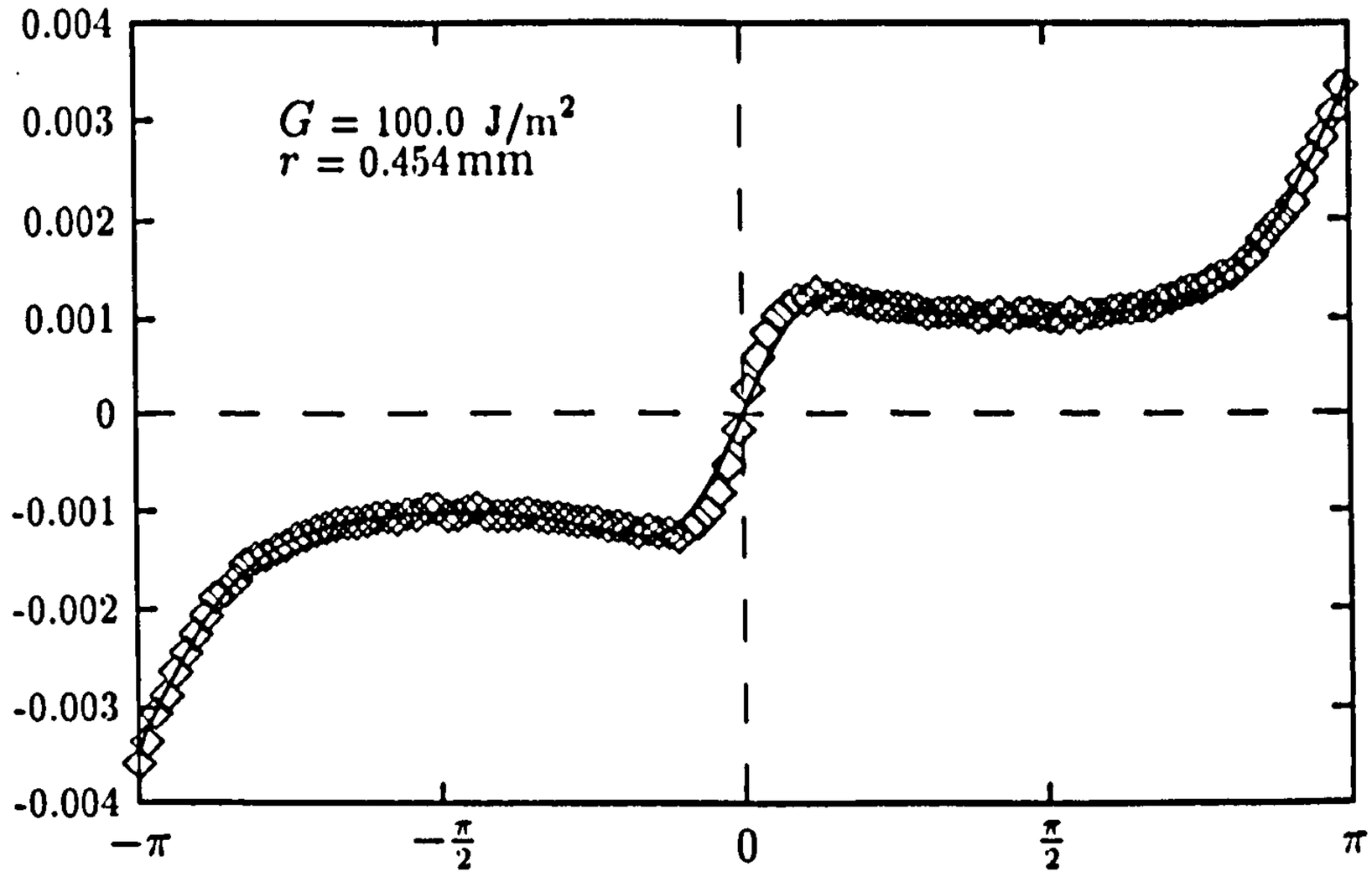


Figure 6.3: $\partial u/\partial y$ on a radial contour for 'ideal' data.

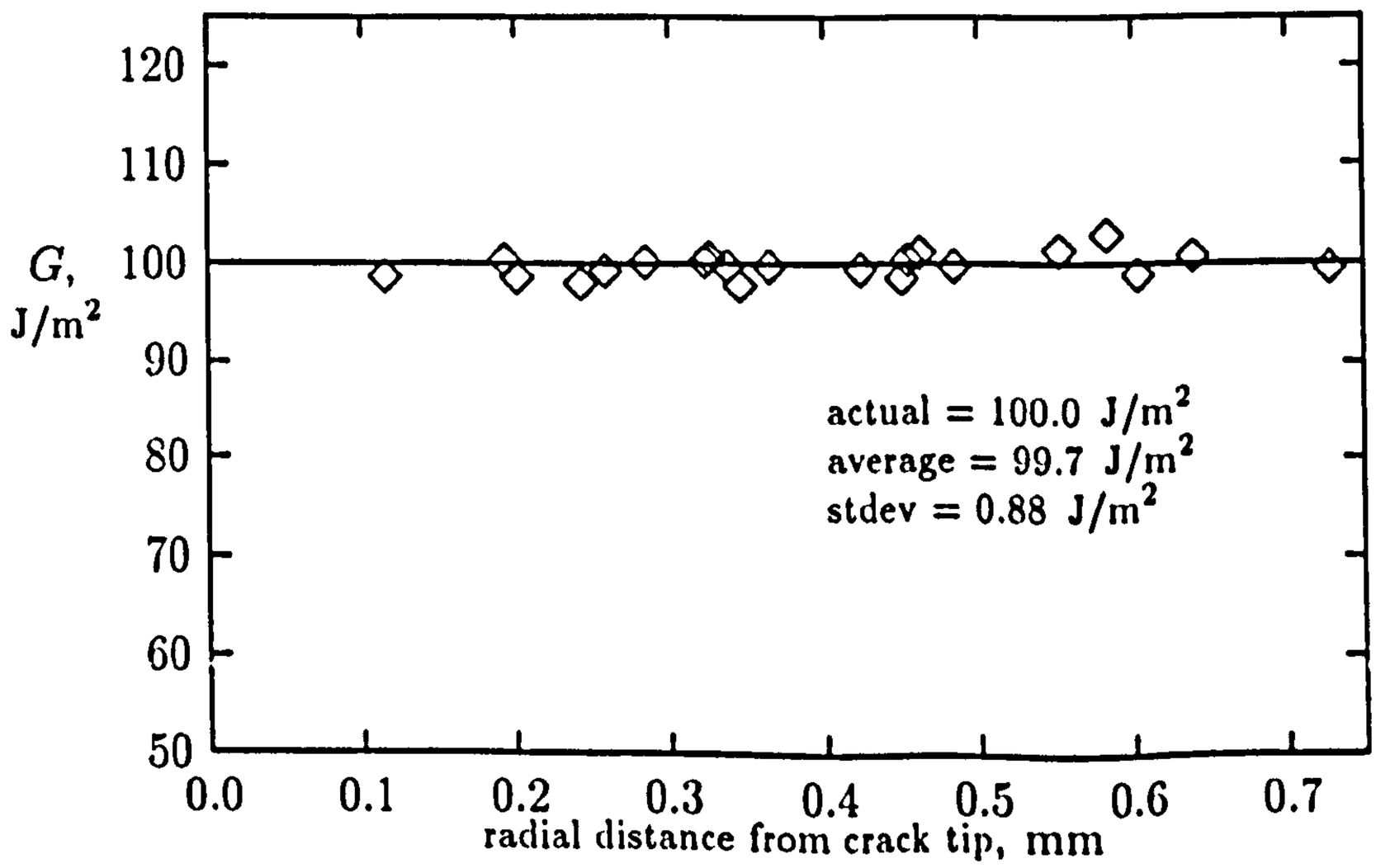


Figure 6.4: J values for contours of different radii for 'ideal' data.

Smoothing of rectangular contours was more difficult because the data were not expected to vary smoothly as a function of position around the contour. At the corner points between a horizontal and a vertical leg, the data will not have continuous derivatives, although it is expected that the same value would be approached from either direction.

It was therefore necessary to fit individual spline functions to each leg, while enforcing continuity between legs. This was accomplished by appropriately weighting the data points to be smoothed, and smoothing each leg independently. The residuals at the common points between legs would then be evaluated, and the weighting of the data points would be modified accordingly. The smoothing of the each leg would then be performed again with the new weighting factors. This procedure would be iterated until the residuals at the common points were no longer improved.

An example of the piecewise smoothing operation for a rectangular contour is shown in Figure 6.5, where the piecewise-smoothed shear strain is plotted as a position around the contour. The digitally filtered data points are not shown for clarity.

Once the various strain and stress values were smoothed by the piecewise splining process, three intermediate functions were constructed. Inspection of Equation 6.3 shows that there are three separate terms that comprise the integrand of the J-integral

$$\begin{aligned} Wn_x &= \frac{1}{2}(\sigma_x\epsilon_x + \sigma_y\epsilon_y + \tau_{xy}\gamma_{xy})n_x \\ T_x\frac{\partial u}{\partial x} &= (\sigma_x n_x + \tau_{xy}n_y)\frac{\partial u}{\partial x} \\ T_y\frac{\partial v}{\partial x} &= (\tau_{xy}n_x + \sigma_y n_y)\frac{\partial v}{\partial x}. \end{aligned}$$

These were determined as a function as a position around the contour, so that a more accurate integration using spline quadrature could be performed.

Spline quadrature [76] is simply a numerical integration method that utilizes the actual coefficients of the cubic splines. This approach provides a more accurate estimate of an integral than the trapezoidal or Simpson's rule, especially in cases where the function to be integrated changes rapidly.

Specifically, Wn_x, $T_x\partial v/\partial x$ and $T_x\partial u/\partial x$ were determined from the smooth

stress and strain data, and then each was fit using ordinary cubic splines (no additional smoothing was applied). Spline quadrature integration was used to integrate each individual term, and these were combined according to Equation 6.3 to obtain a value of J.

A useful plot to visualize which portions of the contour provided the most contribution to the J integral is shown in Figure 6.6. In this plot, the integrand of the J-integral has been evaluated at each point along the contour, using the spline-quadrature integration method described above, and plotted as a function of position.

Additional insight as to which stresses and strains were responsible to the overall value of the J-integral could be obtained by plotting Wn_x, $T_x \partial v/\partial x$ and $T_x \partial u/\partial x$ separately, in a similar manner.

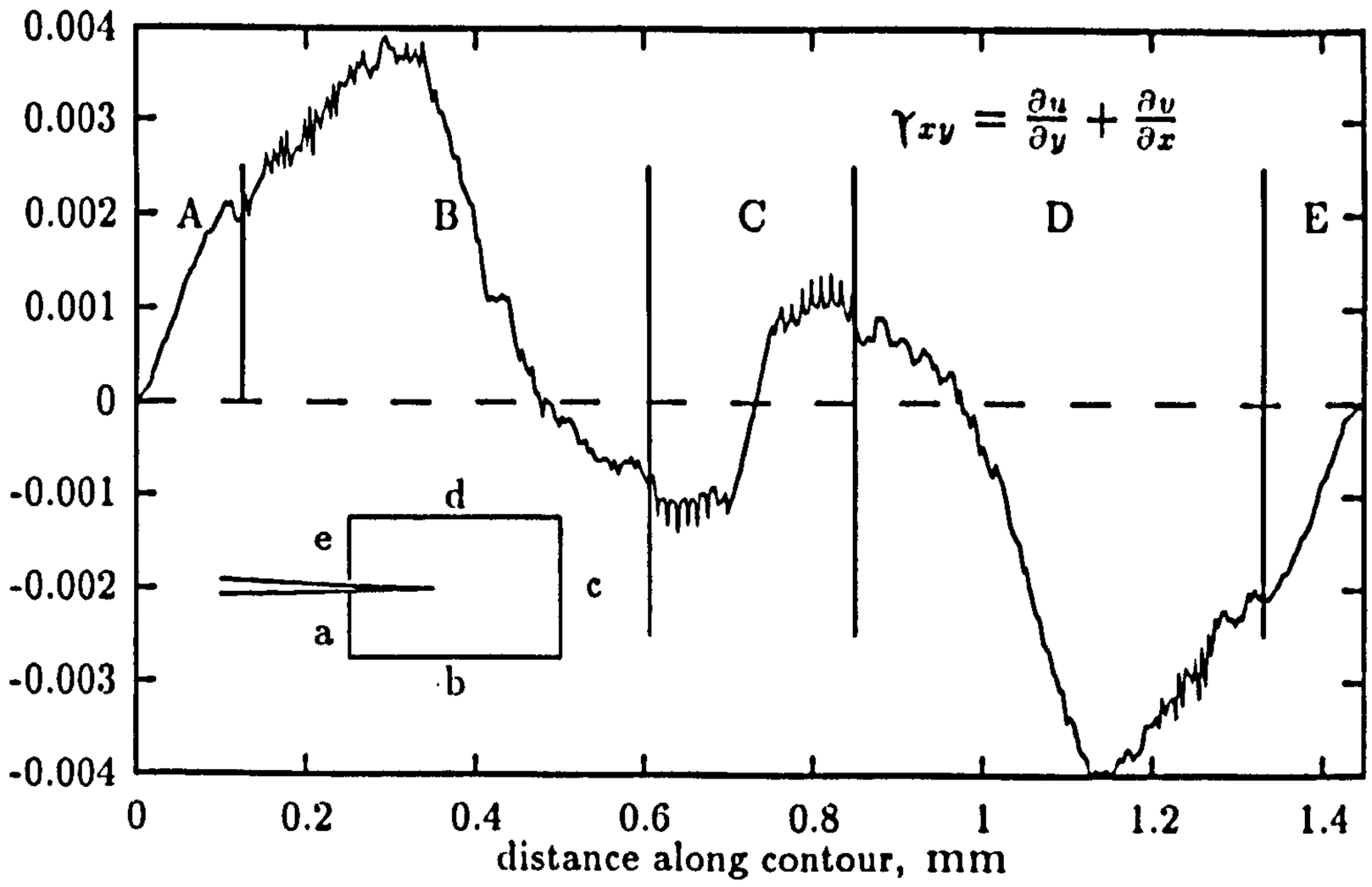


Figure 6.5: Shear strain on a rectangular contour for 'ideal' data.

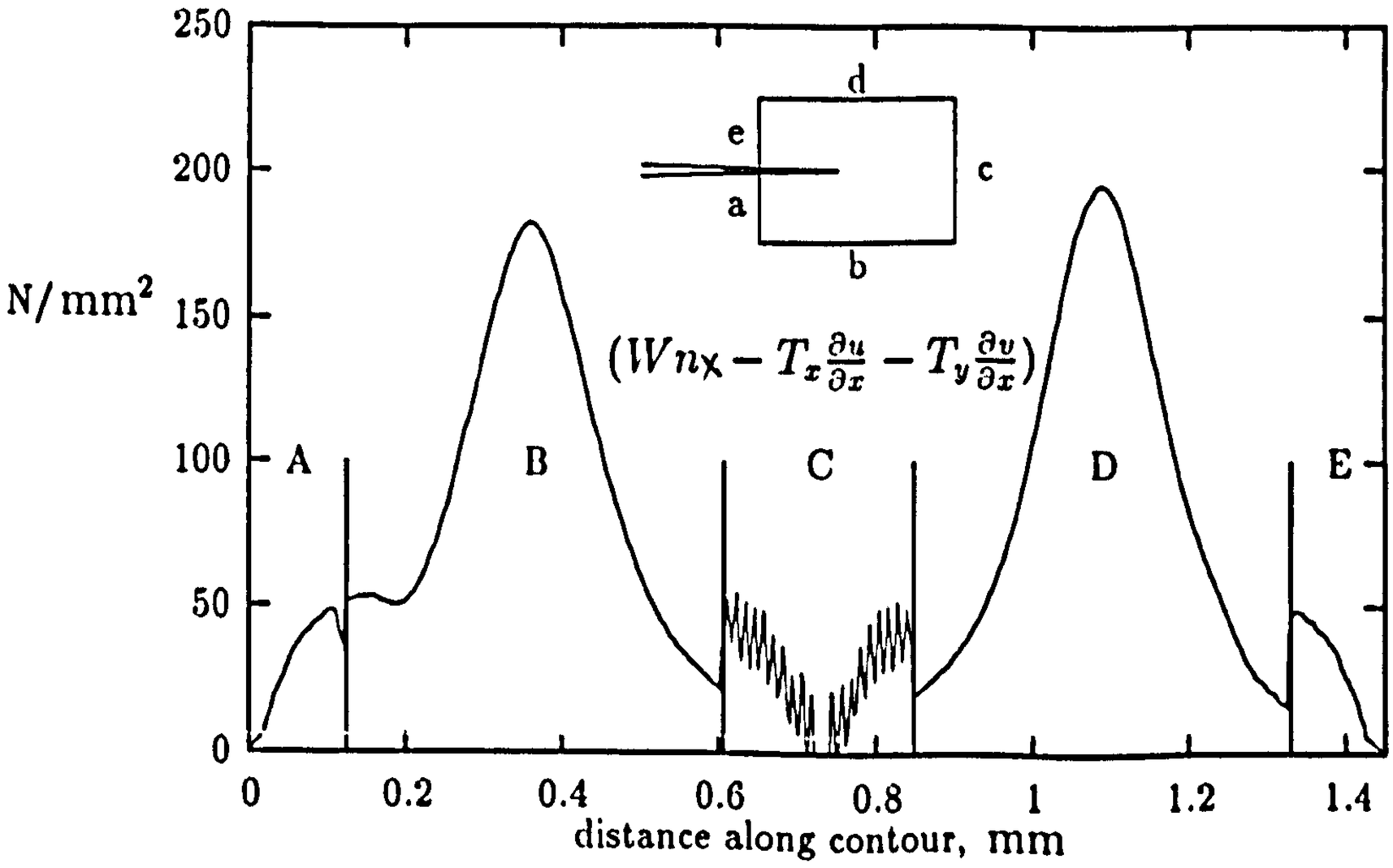


Figure 6.6: Sum of the three J-integral terms on a leg-by-leg basis for a rectangular contour using 'ideal' data.

Chapter 7

Presentation of Results

The first objective of the experimental work was to validate the moiré interferometry method for measuring the energy release rate. This was accomplished by testing unidirectional carbon-fibre/epoxy laminates, and comparing the moiré results to those obtained by conventional methods.

The final set of experiments extended the moiré technique to measure the energy release rate for multidirectional carbon-fibre/epoxy laminates. Since the conventional methods could not be used reliably for these specimens, only moiré results are presented.

7.1 Validation of the Method

Three types of experiments were performed to validate the moiré method for measuring the energy release rate. These were the double cantilever beam (DCB), end notched flexure (ENF) and mixed mode bending (MMB) experiments. The DCB and ENF experiments are well-established, and standardized test methods have been proposed by the European Group on Fracture [45].

The original plan was to perform the DCB and ENF tests in accordance with the European Group on Fracture protocols, and acquire moiré fringe patterns at the point of incipient fracture. In this way, the energy release rate calculated from the moiré fringe patterns would represent a critical value. These could then be compared directly to the global values obtained using the conventional data reduction methods.

However, the sensitivity available with the moiré technique provided an un-

precedented observation capability of the delamination process. The moiré fringes indicated that sub-critical delamination growth occurred at a much lower load level than required to propagate the macroscopic delamination. This made acquisition of fringe patterns *at* critical loads nearly impossible

Sub-critical delamination growth was found to be most prominent in the ENF specimens, and in MMB specimens with considerable G_{II} contribution relative to G_I. Even with the DCB experiments, however, it was difficult to obtain stationary fringe patterns near critical loads.

For the above reasons, an alternative method of validating the moiré approach was adopted. The approach was based on a wider interpretation of the global data available from the conventional tests. Validation was accomplished by comparing sub-critical values from the moiré fringe patterns to values determined using the conventional data reduction methods at the corresponding sub-critical loads.

The following three sections present the results of both the conventional data reduction and the moiré interferometry procedures, for the DCB, ENF and MMB validation experiments. Comparisons are made between the two methods to demonstrate the agreement.

7.1.1 Double Cantilever Beam Experiments

Numerous DCB experiments were performed during the development and validation of both the conventional and moiré procedures for measuring the energy release rate. Since the DCB specimens provided stable delamination growth, several experiments could be performed on a single specimen. Thus, a substantial database of mode I delamination behaviour was collected.

For each test, three global data reduction methods were used to generate G values. A standard data reporting form for a DCB experiment is shown in Table 7.1. The G values obtained for two typical DCB experiments are presented in Figure 7.1. The corrected beam theory results were generally slightly higher than those computed using Berry's method.

As it can be seen in Figure 7.1, no rising R-curve effect was detected for this material. Observation of the fracture surfaces did not reveal any fibre bridging, or signs of plastic deformation.

The average value and standard deviation of G according to the corrected

beam theory method were 177±9 J/m^2, and according to Berry's method 155±8 J/m^2, while the area method gave 188±20 J/m^2. These statistics were based on a consolidation of 4 complete DCB experiments. The errors in determining critical values for the DCB experiments were mostly attributed to the uncertainty in interpreting the critical loads at fracture.

In Figure 7.2, the load/deflection curves for the actual DCB validation exercise are shown. Each curve corresponds to a different crack length. Compliance values were extracted as the slope of each of the six curves and used to generate a log-log plot to recover the empirical correction factor, n, for Berry's method. These values and the least-squares best fit line are shown in Figure 7.3.

For this particular experiment, n was determined to be 3.05, which is very near the beam theory value of 3.00. For the four complete DCB experiments performed, the average value and standard deviation of n was 3.01±0.1.

For the purpose of the validation, sub-critical energy release rates were determined from this value of n, and the corresponding measured load, according to Equation 2.2. Because the sub-critical loads could be known with more confidence than the critical loads, it is argued that the error estimate for the sub-critical G values is at least smaller than that given above for the critical G values.

Typical u- and v-field wrapped fringe patterns for a DCB loading are shown in Figures 7.4 and 7.5. These are an example from epoxy gratings, with a nominal frequency of 1200 lines/mm (each fringe corresponds to 0.4 μm of displacement). Epoxy gratings were used instead of photoresist gratings since they exhibited less noise.

Location of the crack tips for v-fields was generally easier than for u-fields. This is because the v-fields were anti-symmetric with respect to the crack, while u-fields were symmetric. For this reason, it was necessary to determine the precise pixel location of the crack tip for the u-fields during an experiment, by closely examining the actual fringe pattern.

Figure 7.6 shows the $\partial u/\partial x$ (ε_x) strain along a circular contour at a radius of 0.412 mm from the crack tip. The diamonds are the values determined using the digital filters, and the solid line is the result of the spline smoothing.

Two stresses, σ_y and τ_{xy}, are plotted in Figure 7.7 along the same contour. The individual data points in Figure 7.7 were removed for clarity. It can be seen

that the surfaces behind the crack tip are indeed traction-free.

The strain and stress values for each location around the contour were combined to obtain the three individual terms comprising the integrand of the J-integral. These three terms are plotted in Figure 7.8 as a function of position around the contour.

Another useful plot is the sum of the three terms, and this is presented in Figure 7.9. As can be seen, most of the contribution to the final value of G occurs directly above and below the crack tip (at $-\pi/2$ and $\pi/2$).

It should be pointed out that the detailed contour data of Figures 7.6 through 7.9 were selected along a relatively 'noise free' contour. In general, the data were not always so well behaved. Differentiation of experimental data is not an easy task. Even relatively small fluctuations of the data due to noise will result in large deviations from the true signal.

In contrast, integration of experimental data generally produces smoother results, as noise effects are averaged out. In the case of computing the J-integral, however, the integration provides only a moderate amount of relief, since it essentially depends on the *square* of differentiated values.

It was shown in the previous chapter (specifically, Figure 6.4), that the J-integral computation produced values with a standard deviation of $0.88\,\mathrm{J/m^2}$ for 'ideal' data. In the case of real data, however, the standard deviation was naturally much larger.

The error associated with evaluating the J-integral for real moiré data was investigated using the fringe patterns presented in Figures 7.4 and 7.5. The J-integral was evaluated on 120 contours, ranging from a 0.1 mm to 0.7 mm radius. The average value and standard deviation were $120 \pm 18\,\mathrm{J/m^2}$—representing an error of approximately 15 %. Close inspection of the contour-by-contour J values did not reveal any significant trends, rather, the values were normally distributed about the mean.

In practice, the large degree of error was tolerated by performing an additional averaging during the J-integral evaluation process. Instead of just using the value of J obtained on a contour of radius r, usually ten contours were evaluated over an area defined by $r \pm \Delta r$, where Δr was typically 0.01 mm.

When this additional averaging procedure was adopted, the standard devia-

tion of the measurements was reduced to only $\pm 4\,J/m^2$. A plot of the energy release rate as a function of radial distance away from the crack tip obtained in this way is presented in Figure 7.10. The path independence is, in essence, demonstrated.

Having clarified the implementation of the moiré approach, it is now possible to present the comparison between the sub-critical G values obtained using the moiré approach to those obtained using the conventional method. A direct comparison of the conventional and moiré results is shown in Figure 7.11. This is a plot of $G^{moiré}$ versus G_I obtained using Berry's method. The slope of this plot was found to be 1.01, compared with 1.00 for perfect agreement.

a_i (mm)	a_f (mm)	δ_c (mm)	δ_f (mm)	P_c (N)	P_f (N)	C (mm/N)	δ/a (—)	Beam Theory	Berry's Method	Area Method
41.5	46.0	6.3	7.3	19.1	15.8	0.440	0.151	**173.2**	**142.6**	**176.9**
46.0	52.0	7.9	9.0	17.8	14.1	0.620	0.172	**184.0**	**151.4**	**160.9**
52.0	56.0	10.0	11.0	16.2	13.5	0.780	0.192	**186.5**	**153.5**	**220.4**
56.0	60.0	11.8	13.1	15.2	12.7	1.000	0.210	**191.7**	**157.8**	**243.7**
60.0	66.5	14.0	15.3	14.1	11.9	1.200	0.234	**196.8**	**162.0**	**146.3**
66.5	72.5	16.4	17.1	13.6	11.0	1.440	0.246	**201.0**	**165.4**	**173.9**
72.5	78.5	18.0	20.0	12.5	10.5	1.790	0.248	**185.2**	**152.4**	**200.3**
78.5	87.0	21.0	24.0	11.8	9.4	2.370	0.268	**189.4**	**155.9**	**200.6**

Table 7.1: ERR values by the three global methods for a typical DCB experiment.

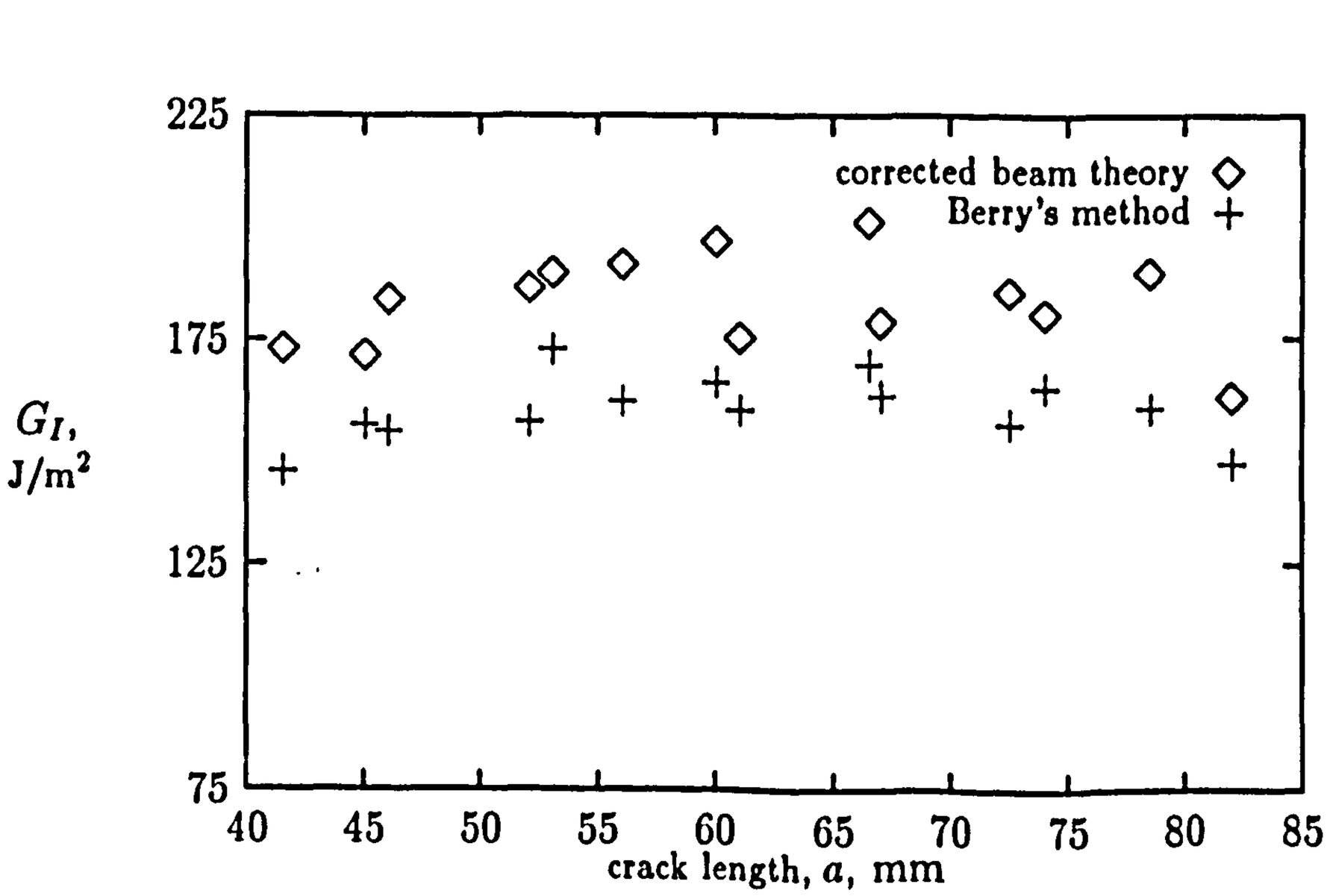


Figure 7.1: Compilation of global energy release rate values for two typical DCB experiments.

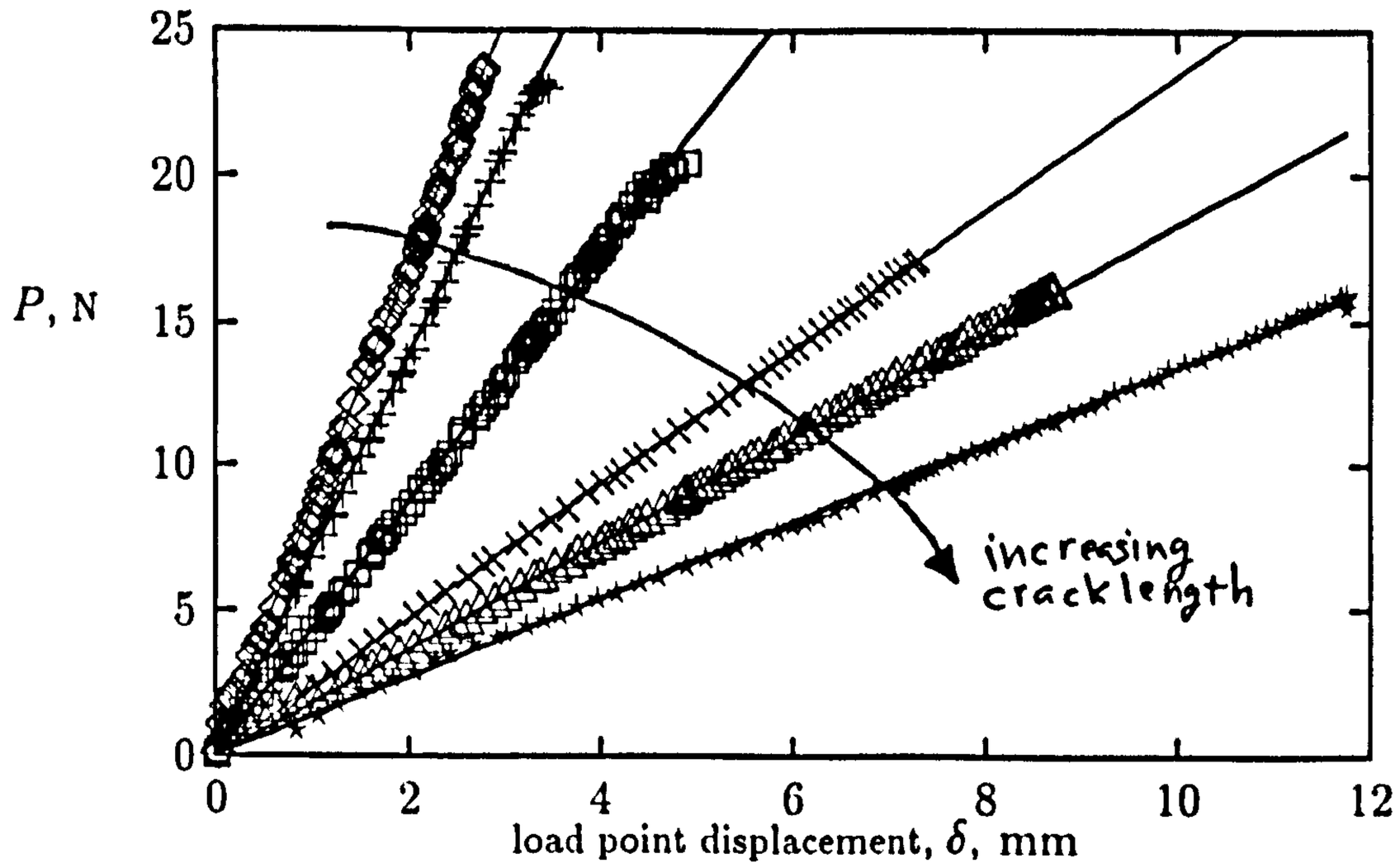


Figure 7.2: Load/displacement plot for DCB validation experiment.

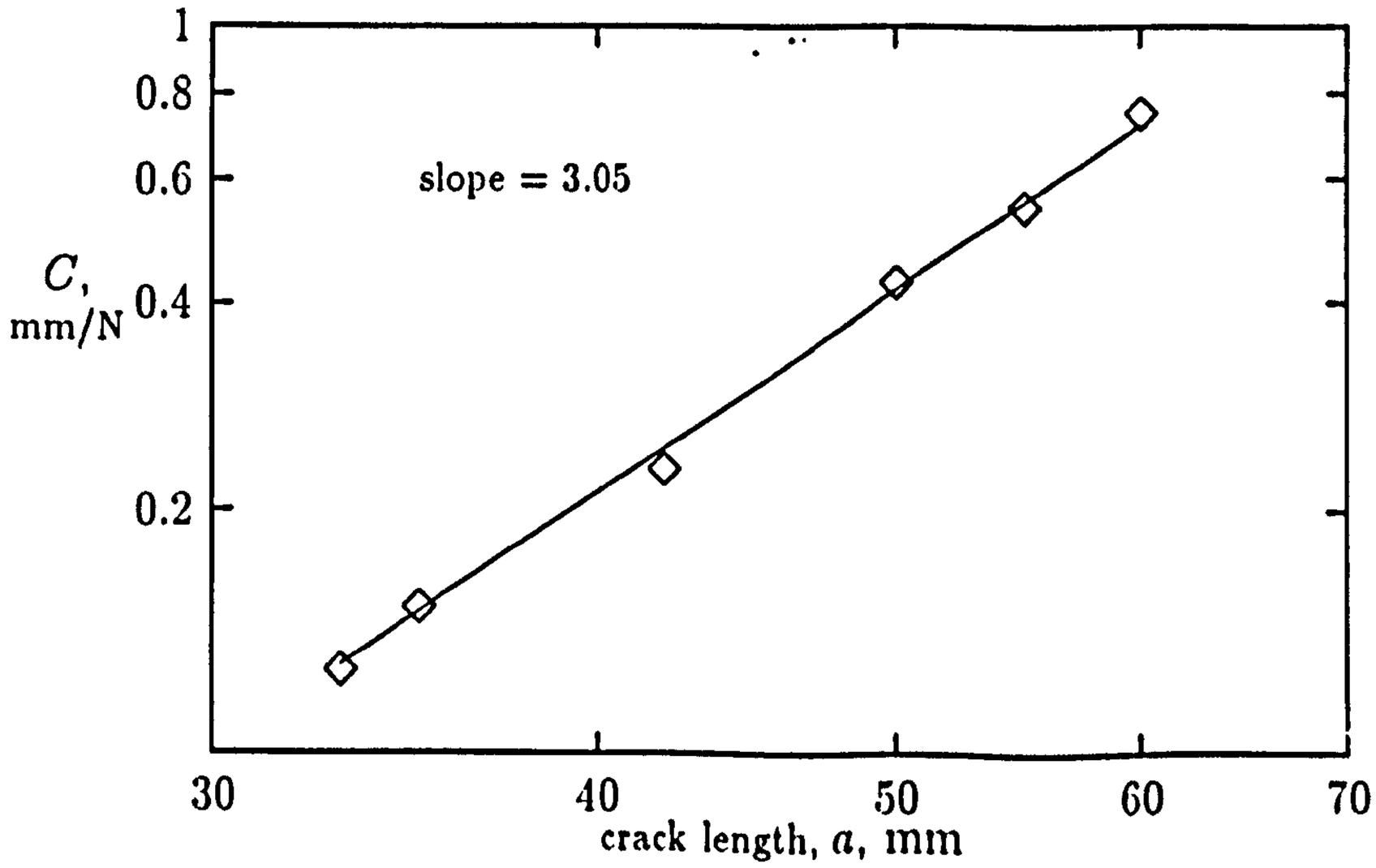


Figure 7.3: log-log plot of compliance versus crack length for DCB calibration (Berry's method).

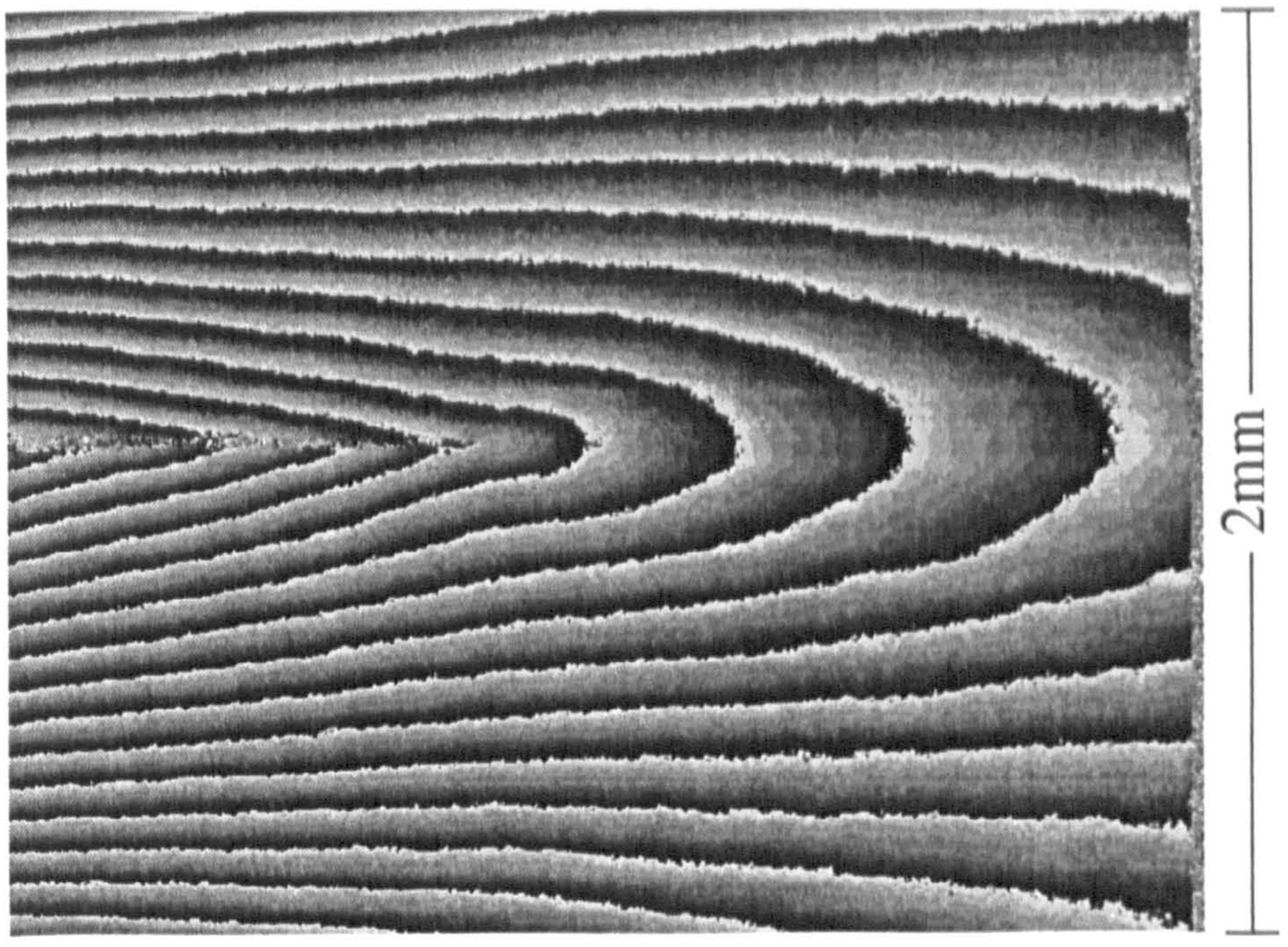


Figure 7.4: Typical wrapped u-field interferogram for a DCB experiment.

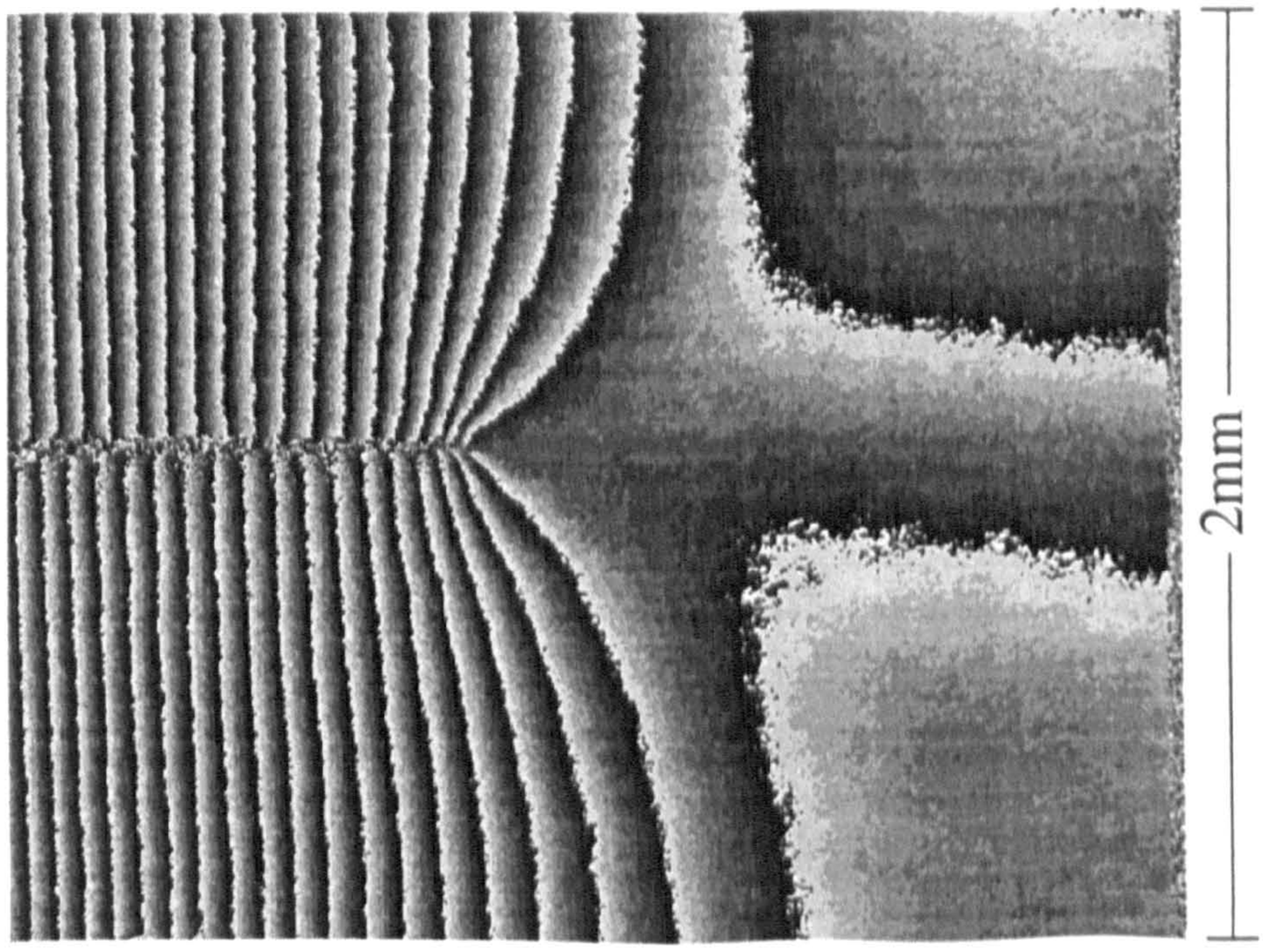


Figure 7.5: Typical wrapped v-field interferogram for a DCB experiment.

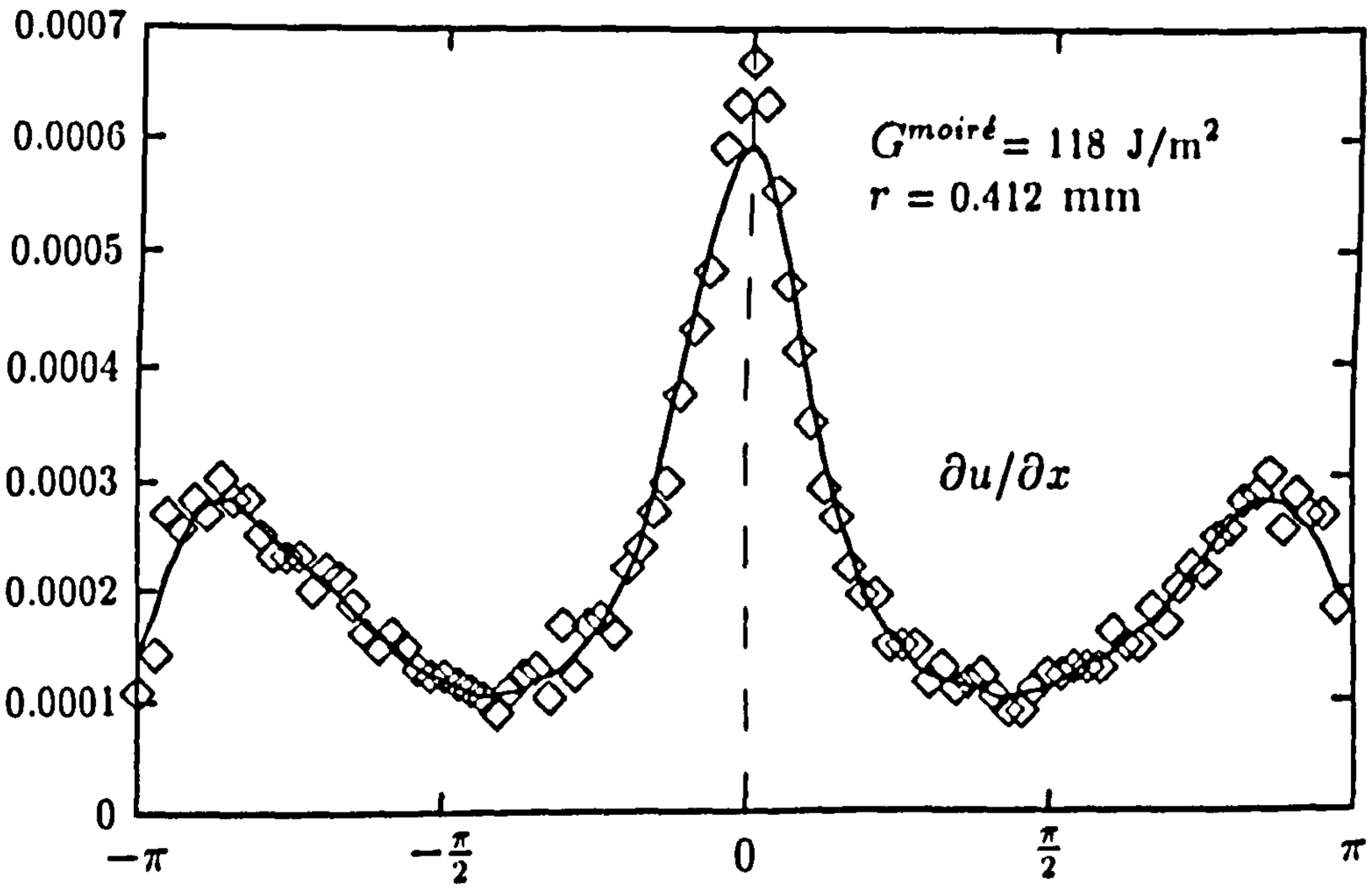


Figure 7.6: $\partial u/\partial x$ on a radial contour for DCB specimen.

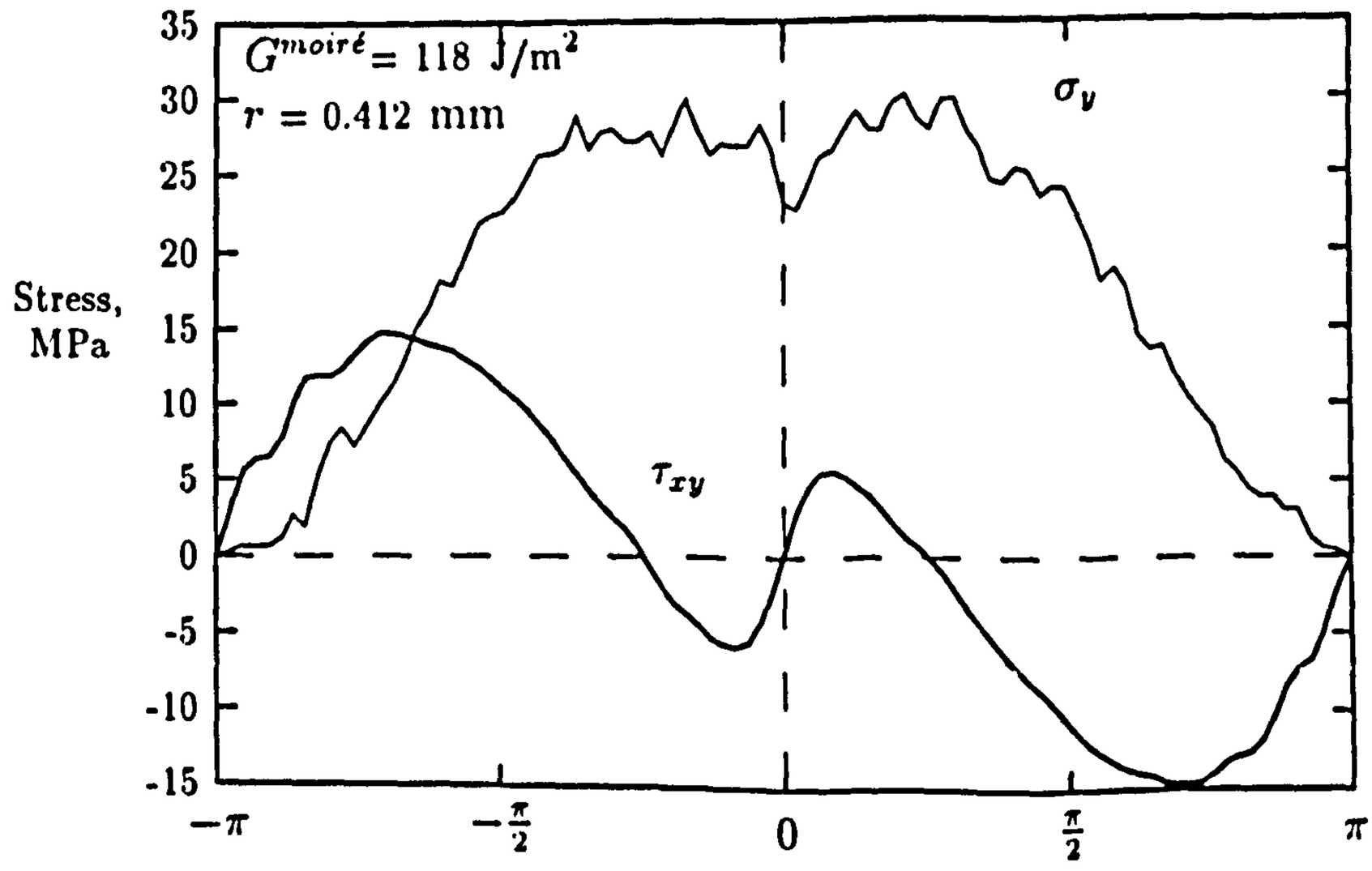


Figure 7.7: σ_y and τ_{xy} on a radial contour for DCB specimen.

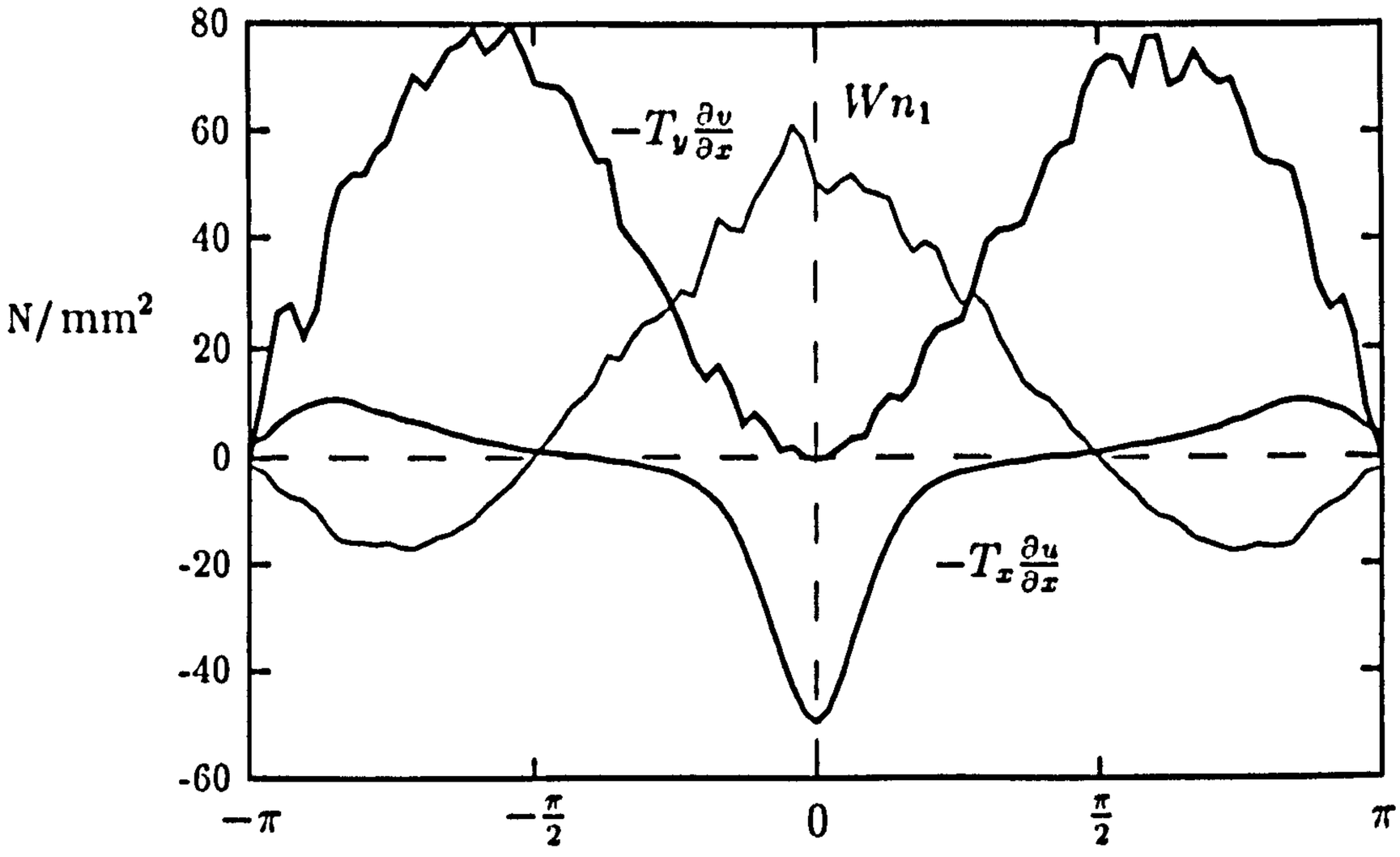


Figure 7.8: Individual contributions in the evaluation of J for a typical DCB contour.

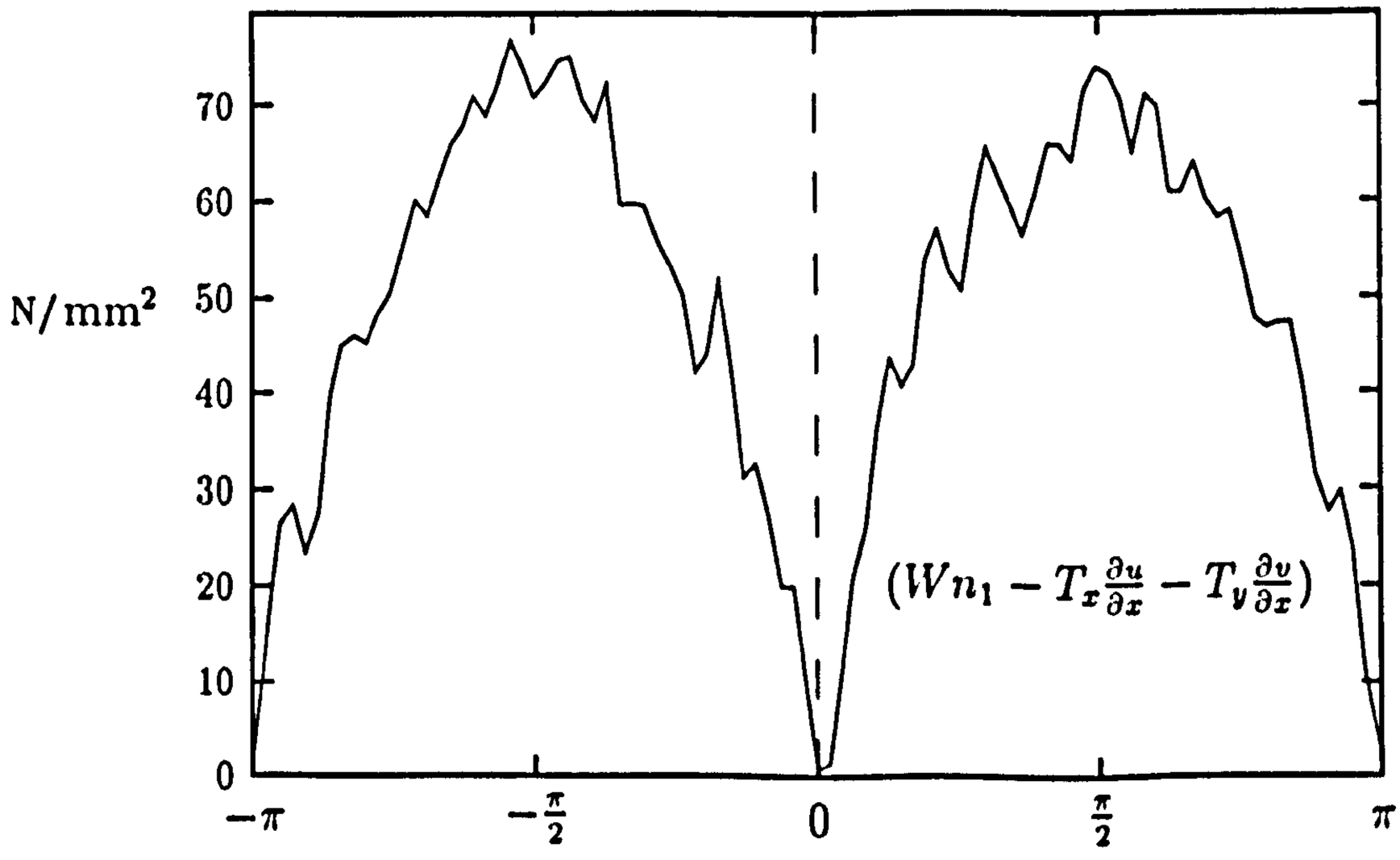


Figure 7.9: The integrand of the J-integral as a function of θ for a typical DCB contour.

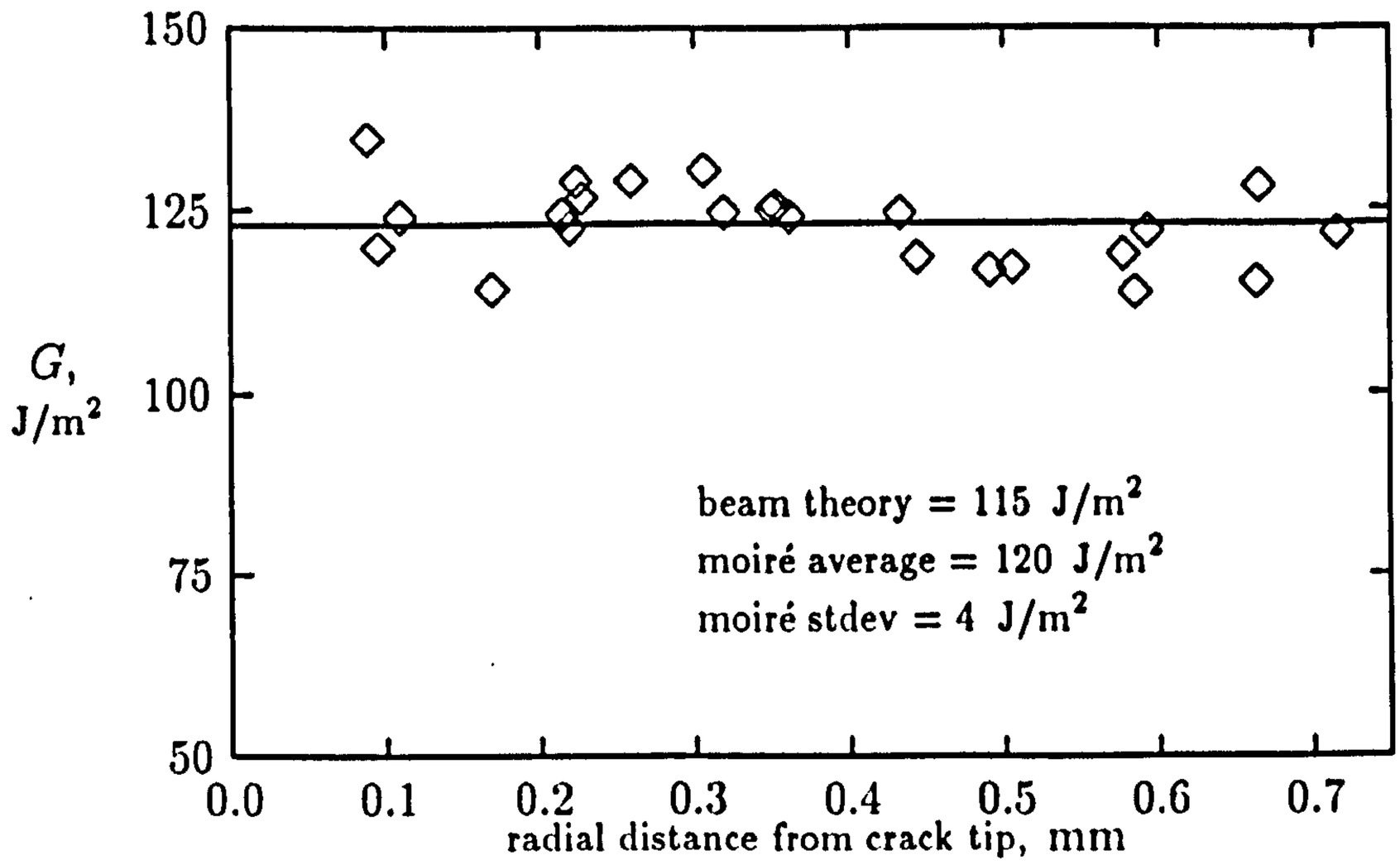


Figure 7.10: J values for contours of different radii for a typical DCB specimen.

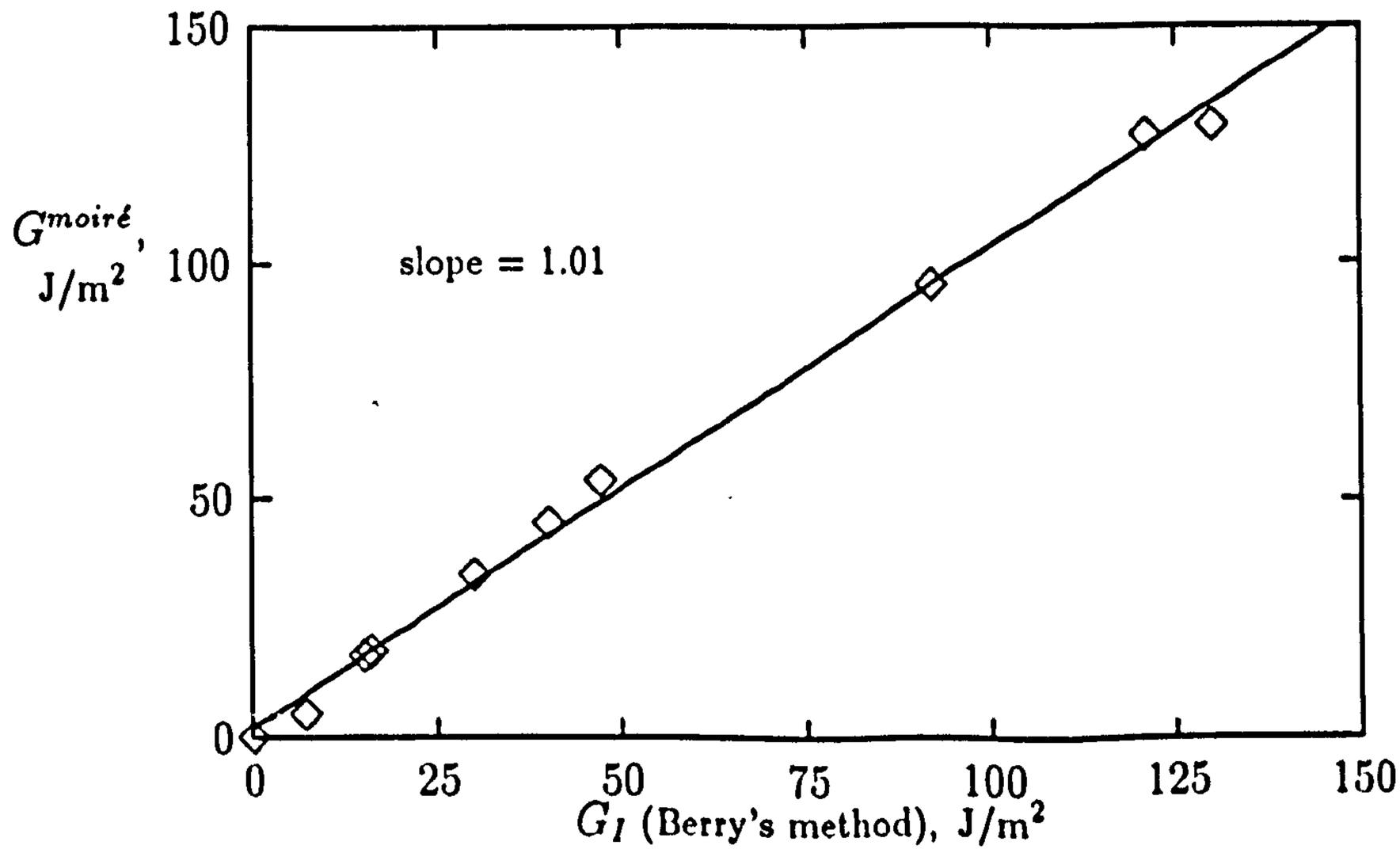


Figure 7.11: Correlation of $G^{moiré}$ versus global values for the DCB validation experiment.

7.1.2 End Notched Flexure Experiments

Far fewer ENF experiments were performed during the course of this work. This was mainly because the ENF test is unstable, so that only one critical energy release rate could be obtained per specimen. Consequently, relatively few statistics were generated for the ENF experiments. For this reason, the development of the moiré method was accomplished using the DCB test.

In total, only three complete ENF tests were performed. These specimens were all precracked using the standard procedure, and the G values obtained were 350, 450 and 425 J/m^2. The fracture surfaces of these specimens was similar to those produced during the DCB experiments. It was possible to distinguish between the fracture surface created due to the initial mode I precracking, the mode II precrack region and the remainder of the fracture surface produced by the catastrophic failure during the actual experiment.

The load/deflection curves obtained during the compliance calibration for the specimen used in the validation exercise are shown in Figure 7.12. The curves are shown as solid lines, because plotting the individual data points would have obscured the plot. Also shown on the plot are the least-squares best fit lines to each curve.

The compliance values, obtained as the slopes of the best fit straight lines, were used to produce the plot shown in Figure 7.13. From this plot, the empirical correction factor, m, was found to be $2.31(10)^{-7}$. There was a much wider variation in the value of m obtained for the three complete ENF experiments. The other two values were $2.54(10)^{-7}$ and $3.1(10)^{-7}$.

A typical load/deflection plot for a ENF specimen that was loaded to fracture is presented in Figure 7.14. The 'jagged' nature of the curve above the load of 250 N is a result of the increased friction between the load screw and the travelling platform. In all of the ENF experiments, no non-linear region of the load/deflection curves was detected. The specimens deformed in a linear manner until the sudden initiation of delamination.

The moiré fringe patterns, however, revealed a very different situation. When the load was increased to as low as 75 % of the critical value, definite growth of the delamination would be indicated by the motion of fringes. The fringe movement was not always accompanied with the growth of the crack at the edge under

observation, suggesting that three-dimensional changes to the delamination front were occurring.

Typical u- and v-field wrapped fringe patterns for an ENF specimen are shown in Figures 7.15 and 7.16. As with the DCB fringe pattern examples, these were produced using 1200 lines/mm epoxy diffraction gratings.

The most notable characteristic of these fringe patterns is the large amount of shear strain at the crack tip. Even for small loads, the number of fringes in the region of the crack tip would very quickly exceed the resolution limit of the video camera. This caused not only a problem in the unwrapping, but it also made precise location of the crack tip difficult.

The v-field fringe pattern also indicates that there is a fairly significant opening of the delamination. A simple fringe count reveals approximately 4 μm of opening displacement at a distance of 1 mm behind the crack tip. This was most likely an artifact of the teflon starter film, which was removed prior to the precracking.

The relative contributions of the three terms in the integrand of the J-integral are plotted as a function of position around a circular contour, at a radius of 0.4 mm in Figure 7.17. The sum of the three terms are shown in Figure 7.18. From these two figures, it is clear that most of the contribution to the J-integral occurs due to the strain energy stored in the material as a result of the large shear strains.

The individual contours for the ENF experiments were affected by noise to the same degree as those for the DCB contours. Thus, it was necessary to perform the same averaging of the individual J values as was done for the DCB $G^{moiré}$ measurements.

For the ENF validation exercise, three different crack lengths were used, and a correlation of the $G^{moiré}$ values with the conventional values obtained by Berry's method are shown in Figure 7.19.

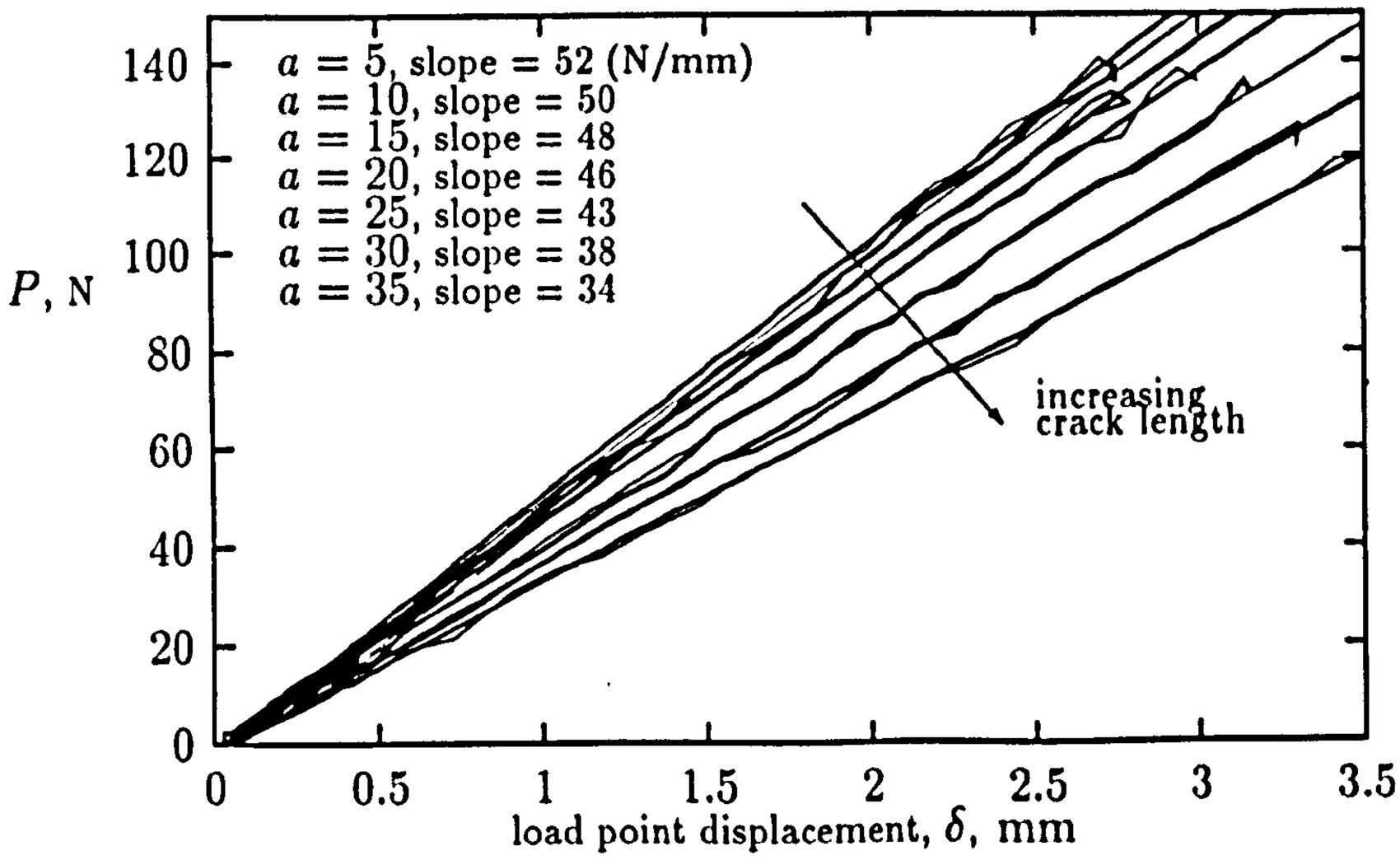


Figure 7.12: Load/displacement plot for ENF compliance calibration.

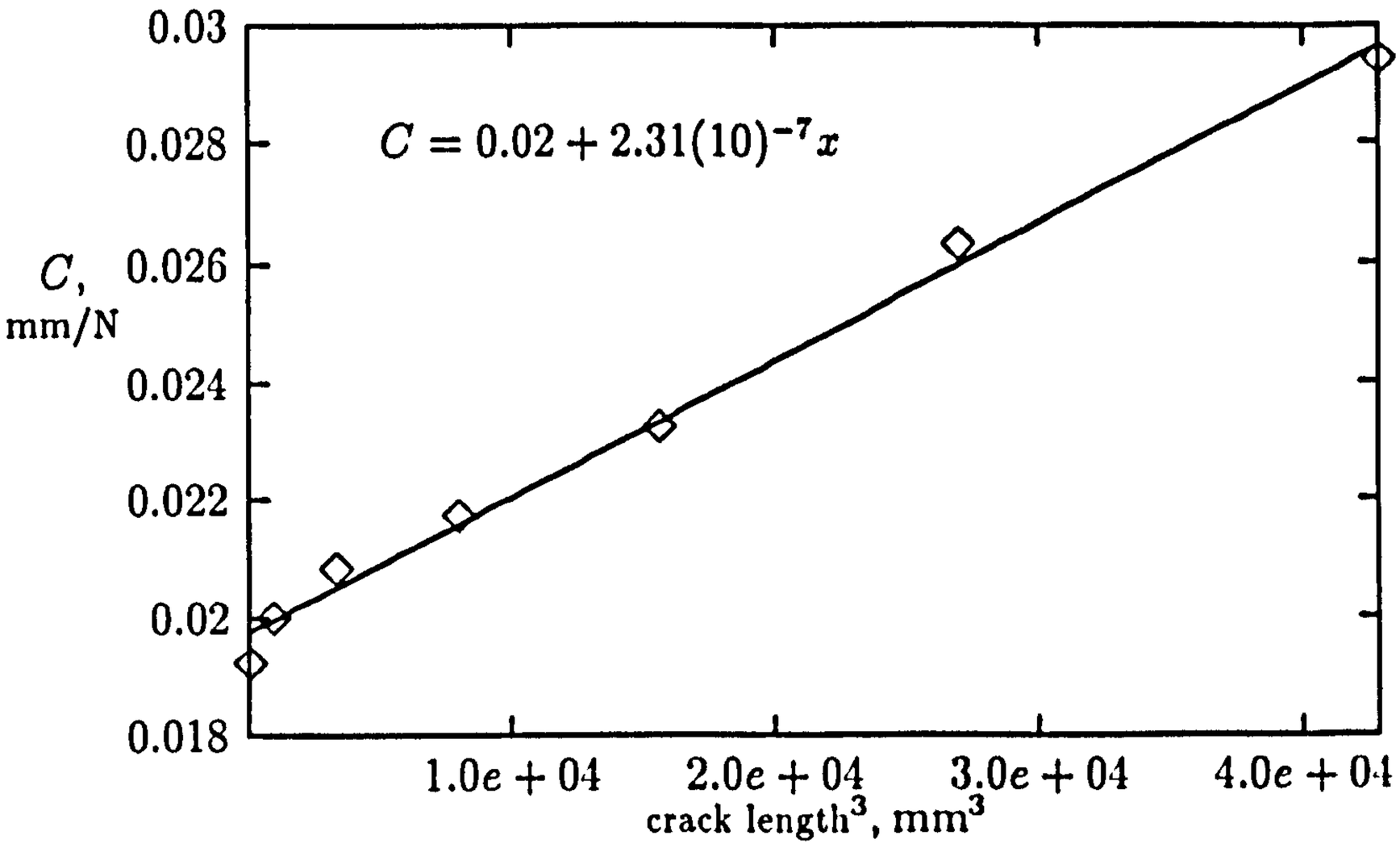


Figure 7.13: Compliance versus a^3 for ENF compliance calibration.

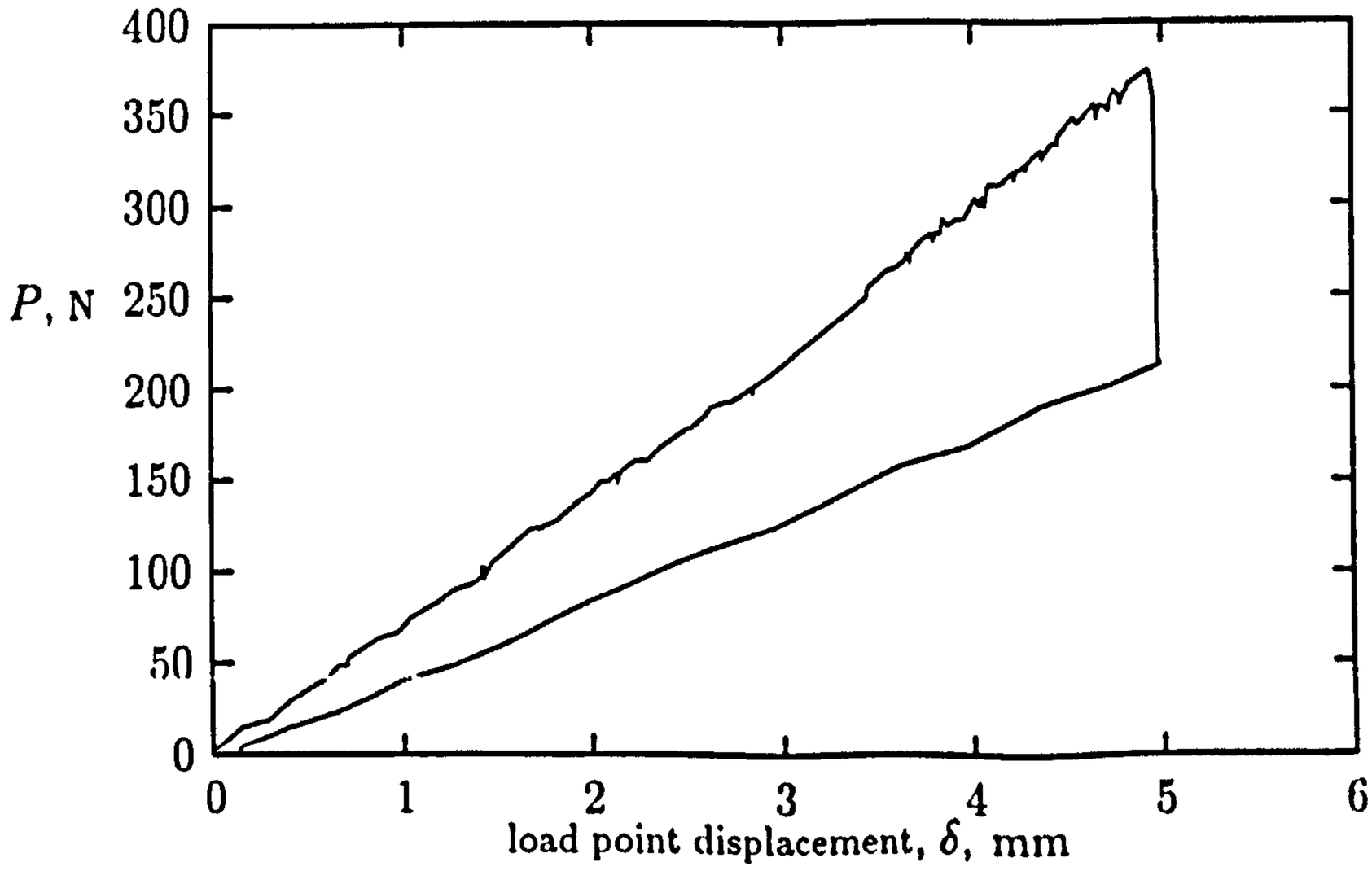


Figure 7.14: Load/displacement plot for a typical ENF experiment.

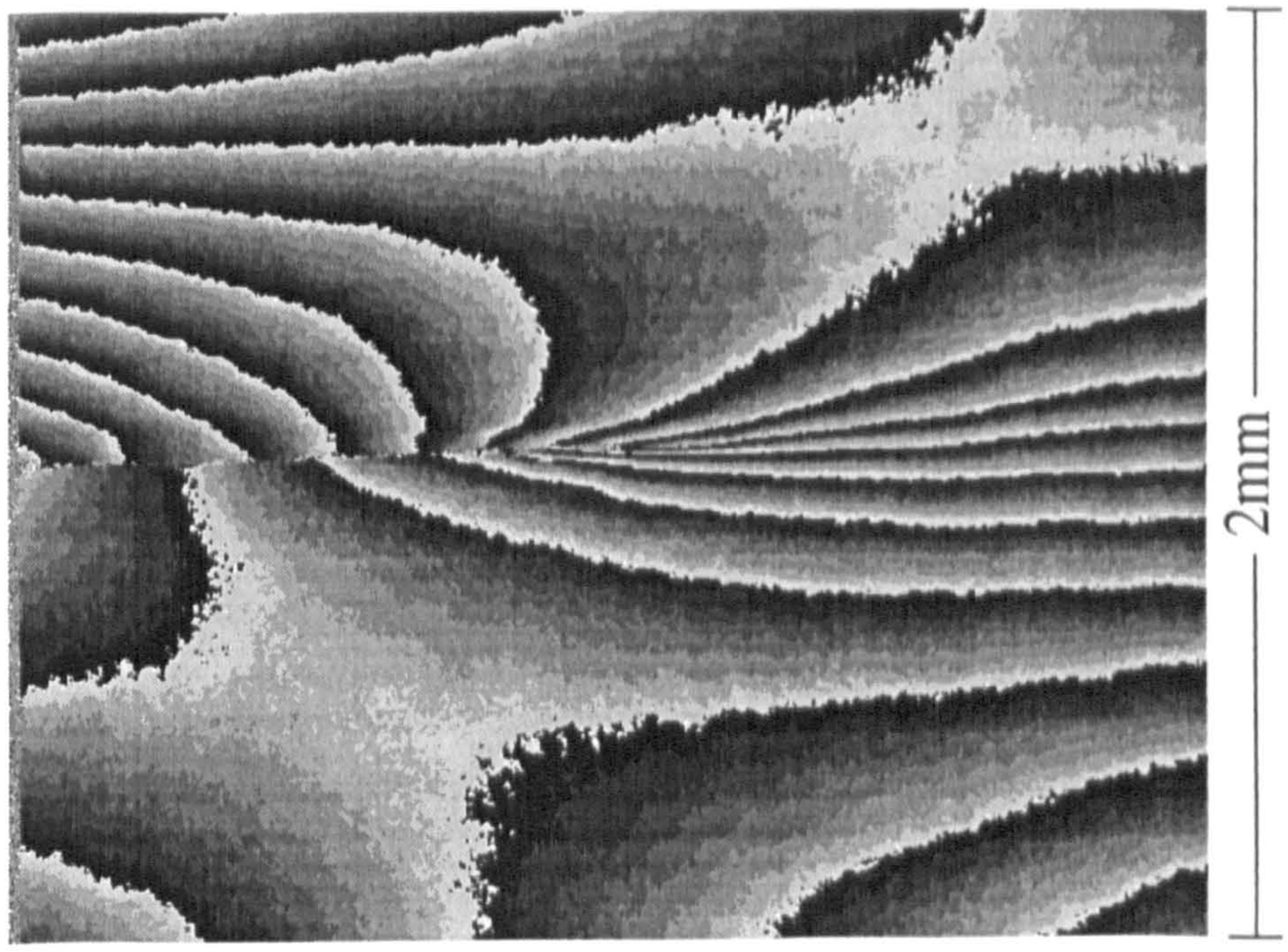


Figure 7.15: Typical wrapped u-field interferogram for an ENF experiment.

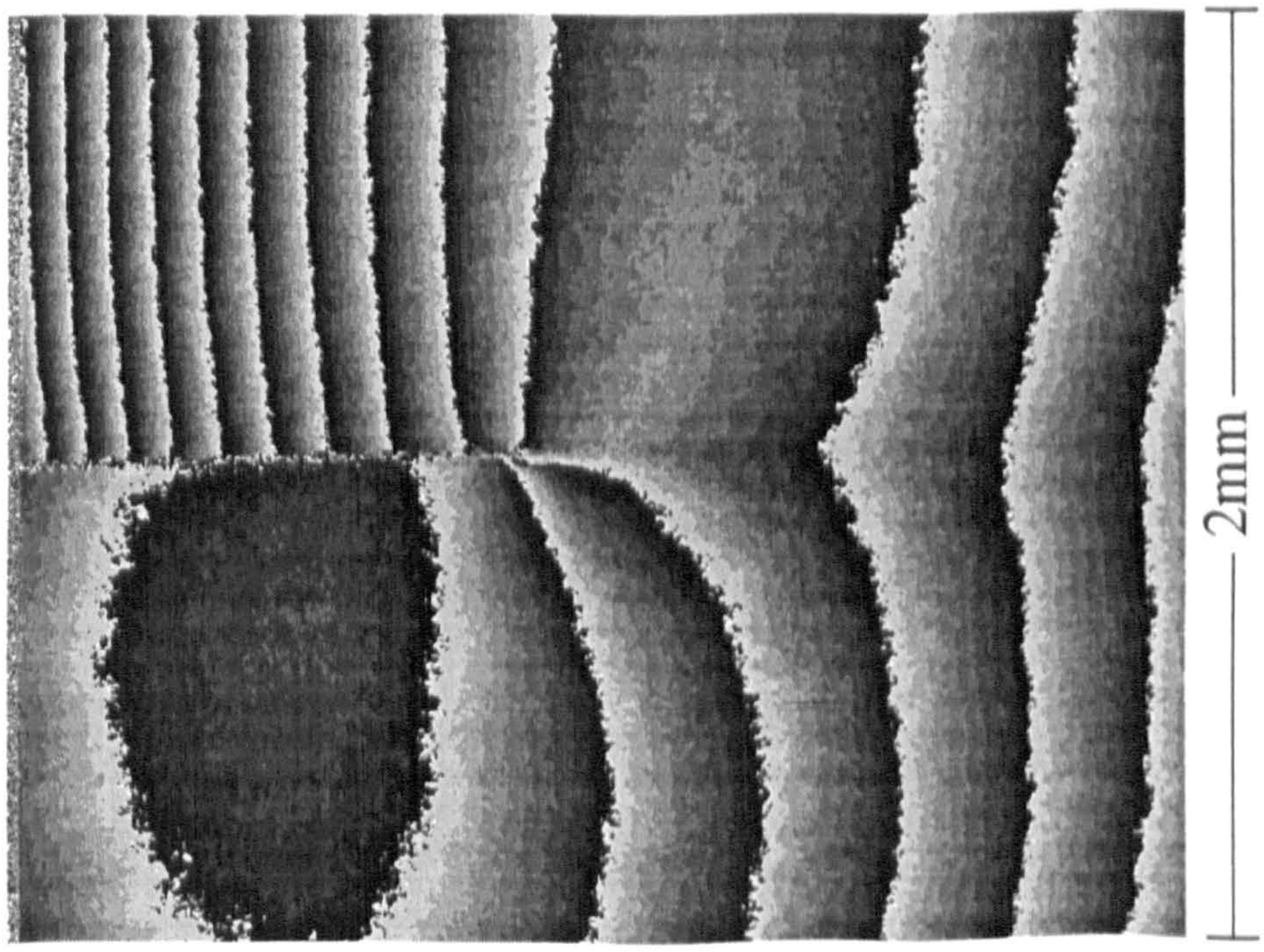


Figure 7.16: Typical wrapped v-field interferogram for an ENF experiment.

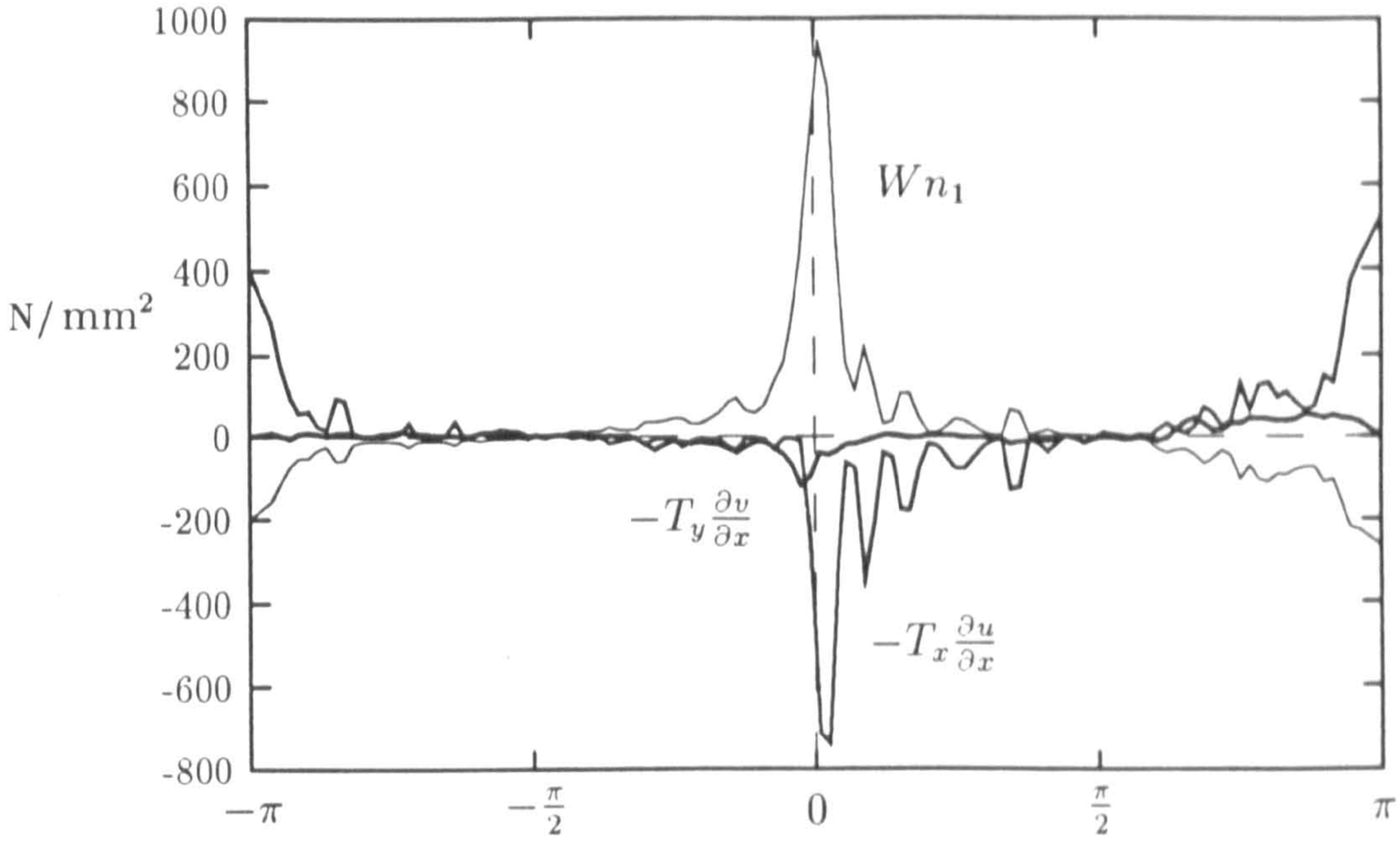


Figure 7.17: Individual contributions in the evaluation of J for a typical ENF contour.

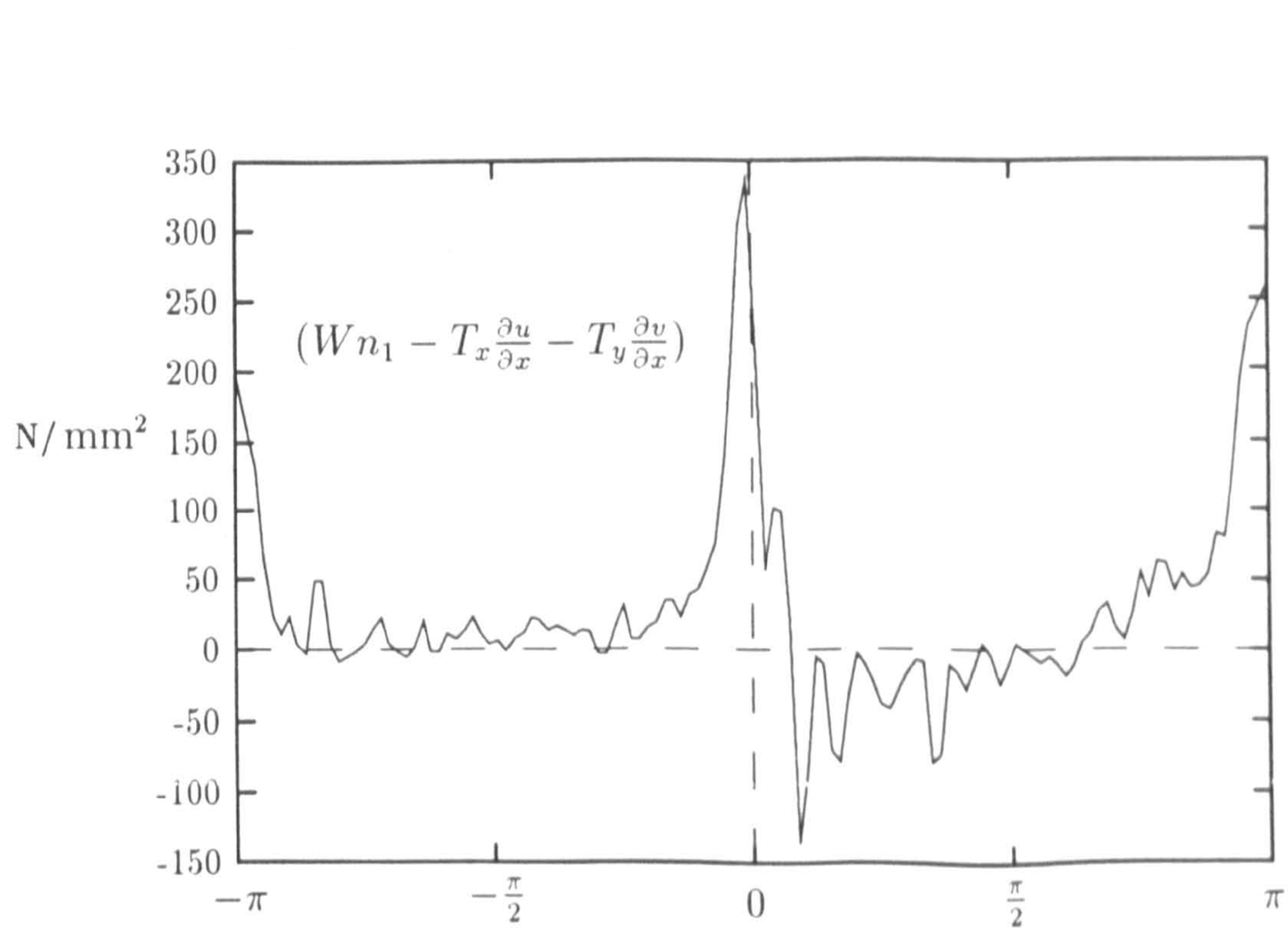


Figure 7.18: The integrand of the J-integral as a function of θ for a typical ENF contour.

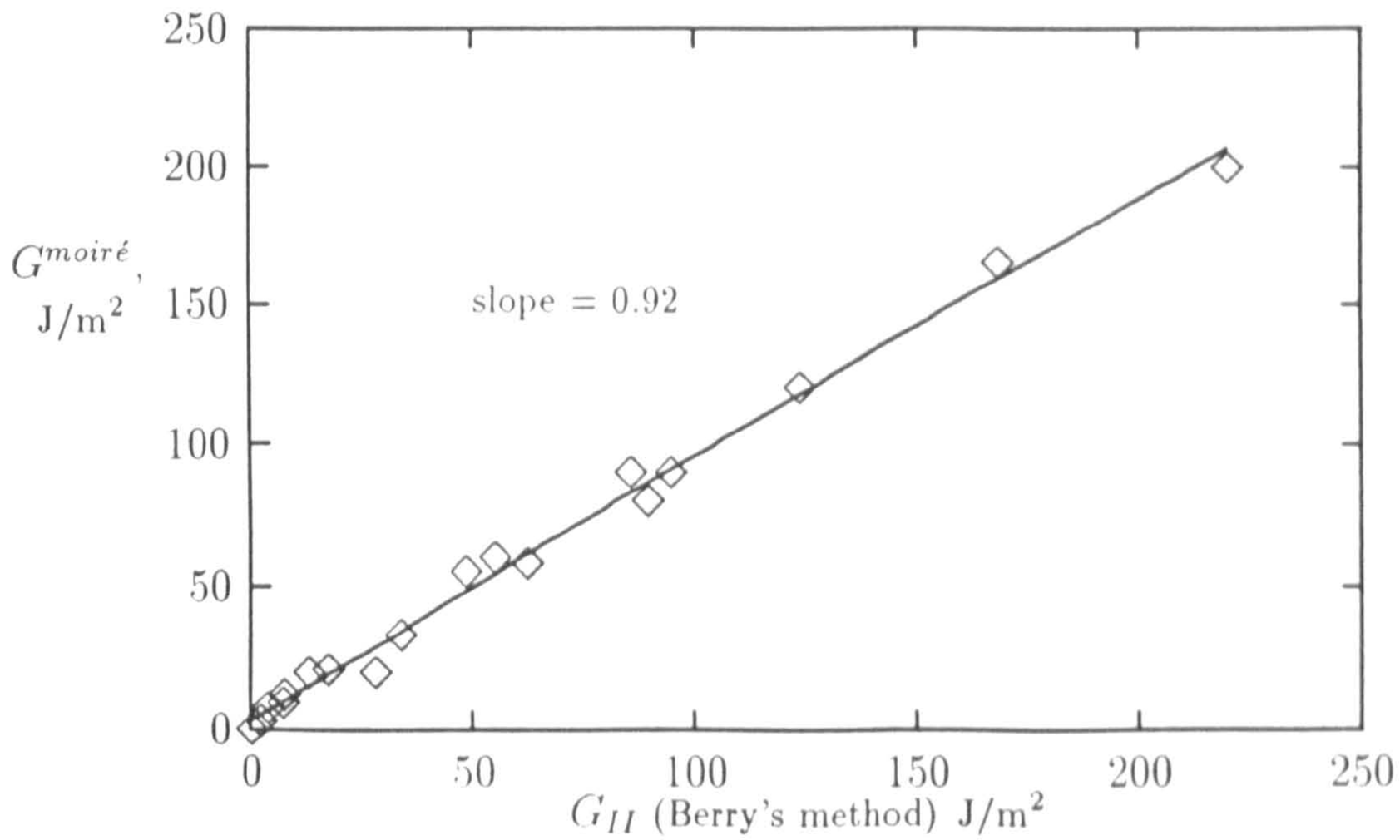


Figure 7.19: Correlation of $G^{moiré}$ versus global values for the ENF validation experiment.

7.1.3 Mixed Mode Bending Experiments

The MMB rig developed for this work was redesigned, as discussed in Section 3.2.4. Only one complete set of experiments was performed after the redesign, so no statistics about the MMB test were generated. Nevertheless, general agreement was found between the moiré measurements of the energy release rate and the values obtained using the conventional application of beam theory.

The specimen tested in the MMB rig were precracked using the standard procedure. Three different nominal mixed mode ratios were chosen, and a separate test was performed for each ratio. It was only necessary to change the distance, c, (as defined in Figure 3.9) as the other dimensions of the rig were held constant between tests.

Four sub-critical pairs of fringe patterns were acquired for each mixed mode ratio, and the corresponding loads were recorded. The maximum loads were kept well below the estimated critical values so that no delamination growth would occur. In this way, the three tests were performed on the same specimen at the same crack length.

The loads obtained during the experiments were used to determine G_I, G_{II} (and hence G_T and Ψ). Both simple beam theory and corrected beam theory were used, and the results are presented in Tables 7.2 through 7.4. The corrections were most prominent for the last test, where the nominal ratio of G_I/G_{II} was 4. However, as can be seen in the tables, there was little difference between the corrected and uncorrected values.

Typical u- and v-field wrapped fringe patterns for an MMB specimen are shown in Figures 7.20 and 7.21. These fringe patterns were not actually acquired during the MMB validation exercise. They were produced from one of the better photoresist diffraction gratings, with a frequency of 2000 lines/mm (each fringe corresponds to 0.25 μm of displacement).

The moiré results and the G_T values obtained using the conventional data reduction methods are compared in Figure 7.22, where individual moiré results are plotted as a function of P^2. The straight solid lines indicate the corrected beam theory results. Error bars indicate the standard deviations of the individual $G^{moiré}$ values.

The agreement between the moiré values was quite good for the predominantly

mode II tests. However, some discrepancy in the trend of the energy release rate values with respect to P^2 can be seen for the test conducted with $G_I/G_{II} = 4.51$. Since this was the only complete set of results generated using the redesigned MMB rig, it is not clear whether the discrepancy arose from the moiré measurement technique, or from the conventional data analysis. It is perhaps significant that at higher G_I/G_{II} ratios, there is more rotation of the loading bar.

P	Beam theory ($\Psi = 0.650$)			Corrected ($\Psi = 0.679$)		
	G_I	G_{II}	G_T	G_I	G_{II}	G_T
10	1.1	1.7	2.8	1.1	1.7	2.8
20	4.4	6.7	11.1	4.6	6.7	11.3
30	9.8	15.1	24.9	10.3	15.2	25.5
35	13.3	20.5	33.9	14.0	20.7	34.7

Table 7.2: Global energy release rate values for MMB experiment ($c = 45\,\mathrm{mm}$).

P	Beam theory ($\Psi = 1.99$)			Corrected ($\Psi = 2.08$)		
	G_I	G_{II}	G_T	G_I	G_{II}	G_T
10	5.6	2.8	8.4	5.9	2.8	8.7
15	12.6	6.3	18.9	13.2	6.4	19.6
17	15.2	7.6	22.9	16.0	7.7	23.7
19	20.2	10.1	30.3	21.2	10.2	31.5

Table 7.3: Global energy release rate values for MMB experiment ($c = 75\,\mathrm{mm}$).

P	Beam theory ($\Psi = 4.32$)			Corrected ($\Psi = 4.51$)		
	G_I	G_{II}	G_T	G_I	G_{II}	G_T
9	20.7	4.8	25.6	21.8	4.8	26.7
12	36.9	8.5	45.4	38.8	8.6	47.4
14	46.7	10.8	57.5	49.1	10.9	60.0
15	53.9	12.5	66.3	56.7	12.6	69.3

Table 7.4: Global energy release rate values for MMB experiment ($c = 140\,\mathrm{mm}$).

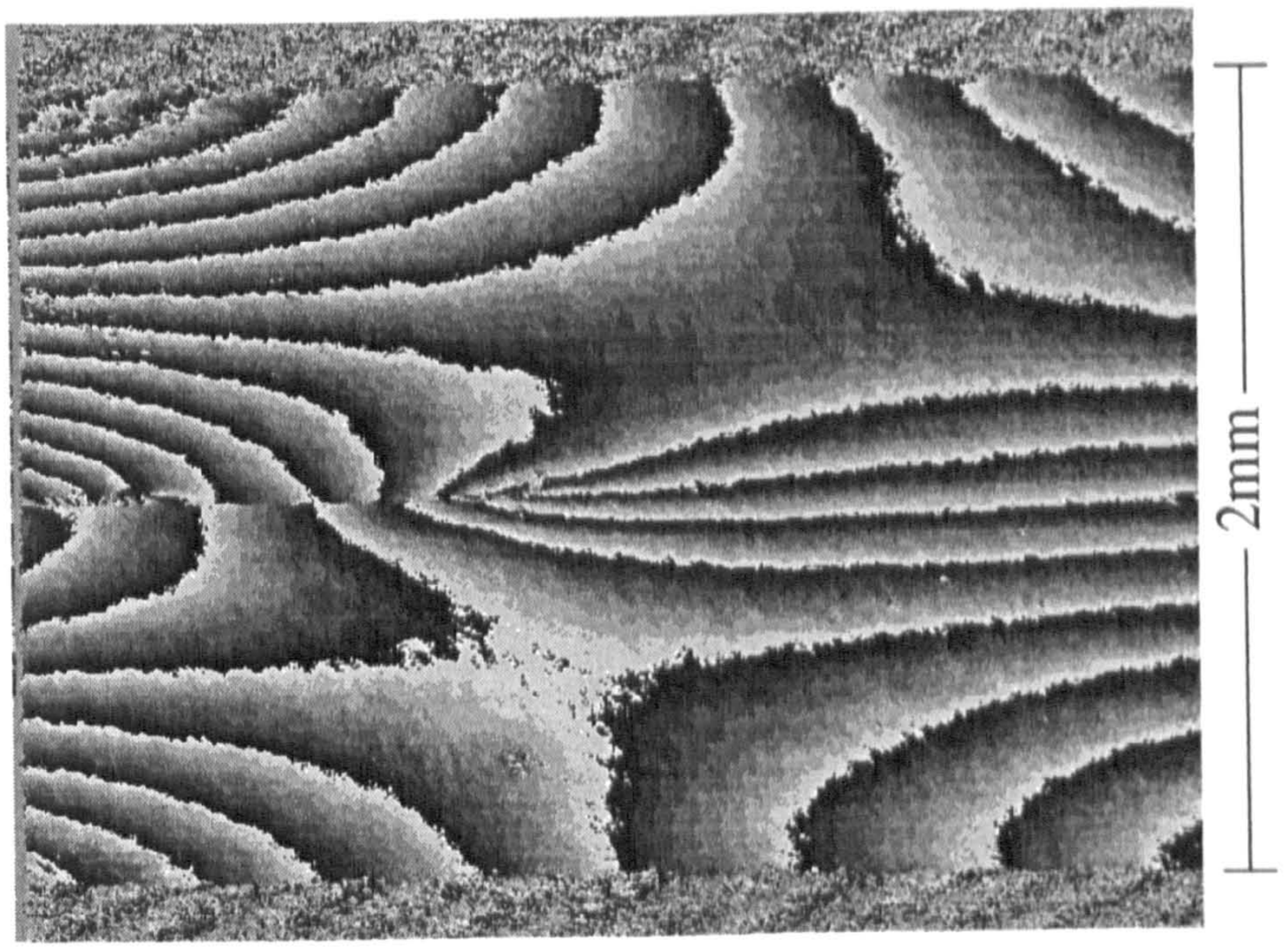


Figure 7.20: Typical wrapped u-field interferogram for a MMB experiment.

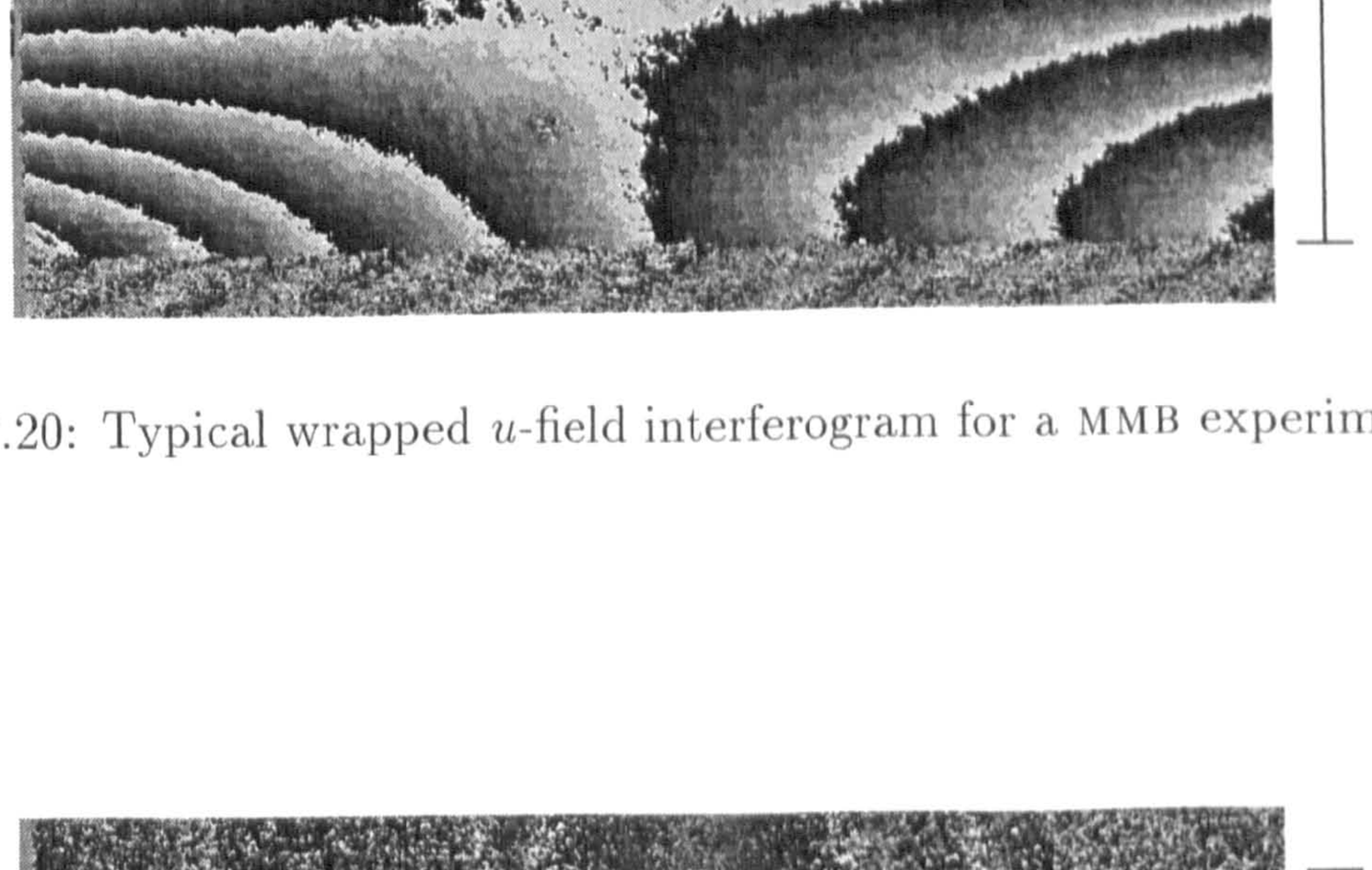


Figure 7.21: Typical wrapped v-field interferogram for a MMB experiment.

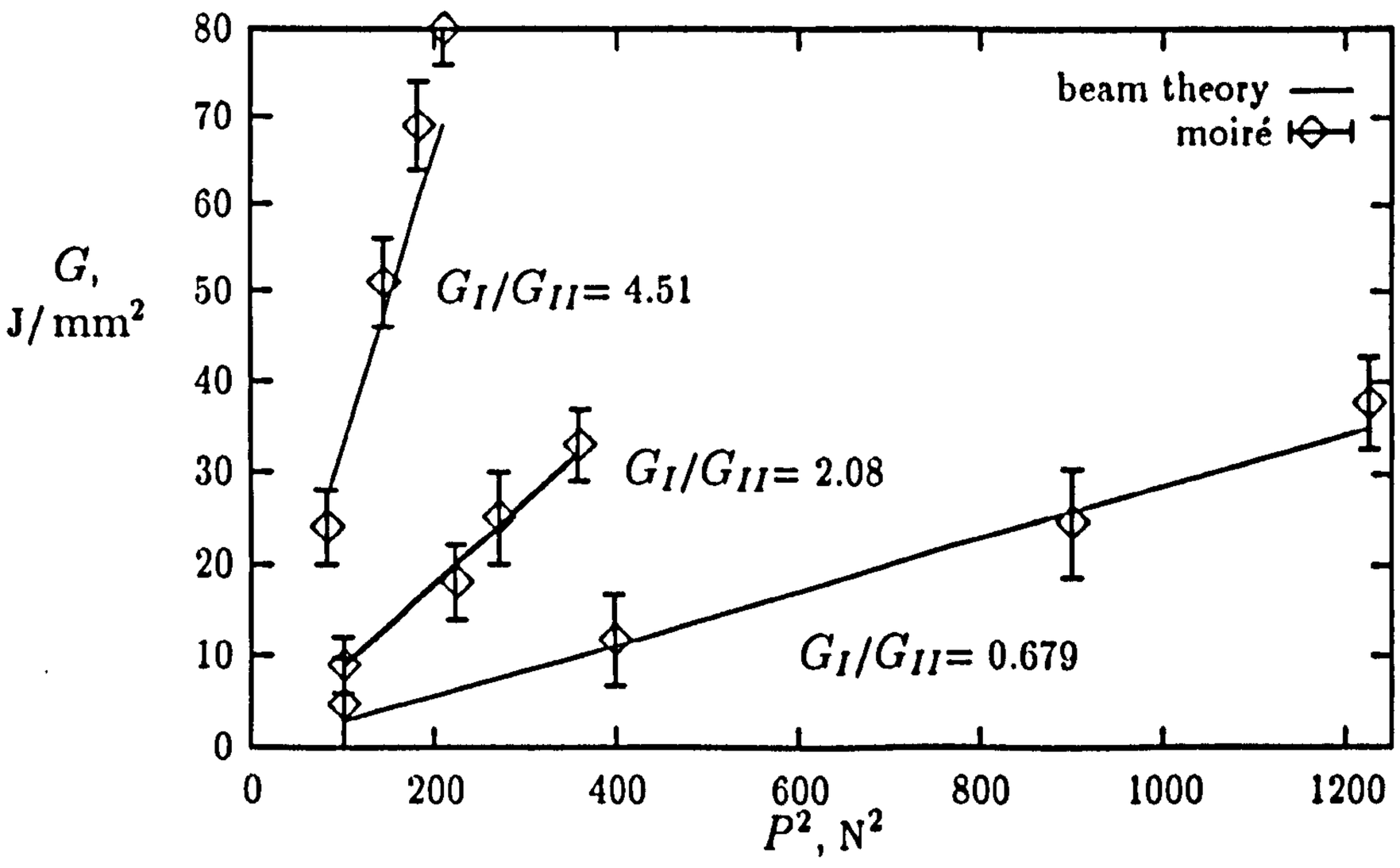


Figure 7.22: $G^{moiré}$ versus P^2 for the MMB validation experiments.

7.2 Extension to Multidirectional Laminates

Two modifications to the J-integral evaluation algorithm were necessary before it could be used for the multidirectional laminates. A provision to allow for rectangular contours, instead of circular ones, was required since the contours would need to be kept within the laminae immediately flanking the delamination. By using rectangular contours, more flexibility was provided for, especially when a neighbouring lamina was only 0.125 mm in thickness.

The second modification made it possible to distinguish between two different materials, each on either side of the delamination. The approach adopted assumes that the delamination lies exactly between the two different materials, by assigning all points with a y-coordinate greater than zero to be material 'A', while all points with a y-coordinate less than zero were designated as material 'B'.

The biggest difficulty encountered during the testing of the multidirectional laminates was that there was no way to control *how* the delamination was going to propagate. If the delamination was not horizontal and located along the bi-material interface, there was no means of determining the J-integral, hence no way of measuring the energy release rate.

As an example, the u- and v-field wrapped fringe patterns from a symmetric $[0_2/+45/0_2/-45/0/90]_s$ specimen are presented in Figures 7.23 and 7.24. The specimen was loaded as in a double cantilever beam test, and the delamination was originally in the centre of the 90°layers. During the precracking, however, the delamination shifted to the upper 90°/0° interface.

During subsequent stable growth, the delamination would propagate for approximately 0.4–0.8 mm and then would shift to the opposite 90°/0° interface. This pattern of growth was observed in two different symmetric $[0_2/+45/0_2/-45/0/90]_s$ specimens. Figures 7.23 and 7.24 characterize this delamination growth behaviour.

The load/displacement plot during the growth of the delamination was markedly different from that for the unidirectional laminates. The growth was not step-wise incremental, rather, the delamination appeared to wander between the upper and lower interfaces bounding the two ply thick 90° layer. Fibre bridging was visually observed in the wake of the delamination.

The previous fringe patterns illustrate the additional complications of applying the moiré technique to multidirectional laminates. However, it was possible to capture the delamination in a position suitable for a J-integral evaluation, and this forms the subject for the first example presented below.

7.2.1 Double Cantilever Beam Experiment

By carefully propagating the delamination shown in the previous figure, it was possible to produce a horizontal *delamination long enough to apply the J-integral* evaluation. Figure 7.25 shows the u-field wrapped fringe pattern after the delamination was grown approximately 4 mm.

An additional optics problem with testing the multidirectional specimens was that a much greater magnification was needed to adequately image the region of interest. Instead of utilizing a 1:4 image ratio, it was necessary to use more than twice this magnification, resulting in an imaging ratio of approximately 1:10. While this ratio is near the maximum for this lens, the most significant problem was not the quality of the image, rather, it was the intensity of the light at the camera.

A compromise had to be made with respect to the magnification of the fringe pattern. At higher magnification, the area of interest would be much larger, so more pixels could be involved in the data acquisition process. However, due to the limited amount of laser power available (30 mW) the additional magnification was accompanied by a decrease in fringe contrast. Needless to say, the fringe patterns acquired at the higher levels of magnification were much more noisy, and more difficult to process. As a result, a greater attention to the interpretation of the raw results from the J-integral evaluation was required.

For these reasons, fringe patterns were acquired at three different levels of magnification, for the same loading condition. The first was at the maximum possible magnification according to the imaging system, an imaging ratio of 1:15, (the 512 pixels in the vertical direction corresponded to about 0.75 mm). These fringes in these images had an average peak-to-peak intensity modulation of only 20-30 gray levels. The second pair of images was taken at a more 'comfortable' imaging ratio of 1:10 (the 512 pixels in the vertical direction corresponded to about 1.0 mm). The fringes in these images had a peak-to-peak modulation of

about 120-140 gray levels. The third pair of images were recored at 'normal' magnification, roughly 1:4. (The u-field fringe pattern for the third set of images is shown in Figure 7.25.)

The data for the first pair of images did not yield reliable J-integral measurements. However, the second pair of images acquired at 1:10 did provide reliable information, albeit the image processing required more attention. The three individual terms for the J-integral along a selected rectangular contour are shown in Figure 7.26. The two material regions are indicated, and it can be seen that there is a significant change in the terms across the interface. The sum of the three terms is presented in Figure 7.27.

To arrive at the J-integral measurement for the DCB $[0_2/+45/0_2/-45/0/90]_s$ specimen, it was again necessary to average the individual results from various contours. For the case of rectangular contours, however, this was not as straightforward as with the circular contours. This was purely an implementation detail—it was a lot easier to systematically vary a single radius, than it would have been to vary the five individual legs of a rectangular contour. Essentially twenty different contours were chosen interactively for each measurement, and these were used to generate an average J of $180\,\mathrm{J/m^2}$, with a standard deviation of $25\,\mathrm{J/m^2}$. As a rough comparison, the energy release rate according to conventional beam theory was determined to be $250\,\mathrm{J/m^2}$.

7.2.2 End Notched Flexure Experiment

The next experiment that was performed with the $[0_2/+45/0_2/-45/0/90]_s$ laminate was an ENF loading of an asymmetric specimen with the delamination originally between layers 10–11. The u-field fringe pattern for this specimen is shown in Figure 7.28.

The specific of this experiment was to capture fringe patterns at four successive loadings, so that the energy release rate could be determined as a function of applied load. To accomplish this, the crack tip indicated by the letter 'A' in Figure 7.28 was centred in the field of view, using an imaging ratio of approximately 1:12. (The efficiency of this grating was slightly better then the grating used in the previous experiment.) Fringe patterns were acquired at the four different loads, and the results of the J-integral evaluations are shown in Table 7.5.

After the experiment, but before the specimen was unloaded from the maximum load of 160 N, the magnification was changed back to the 'normal' imaging ratio of 1:4. This was done so that a pair of images could be recorded that showed the entire specimen (Figure 7.28). At this time, it was discovered that an additional crack must have formed during the loading of the specimen, indicated by the letter 'B' in Figure 7.28.

It is argued that this additional delamination effectively 'shielded' the delamination that was under observation from the applied loading. Unfortunately, this second delamination was not discovered until after the experiment, or it too could have been the subject of detailed J-integral evaluation.

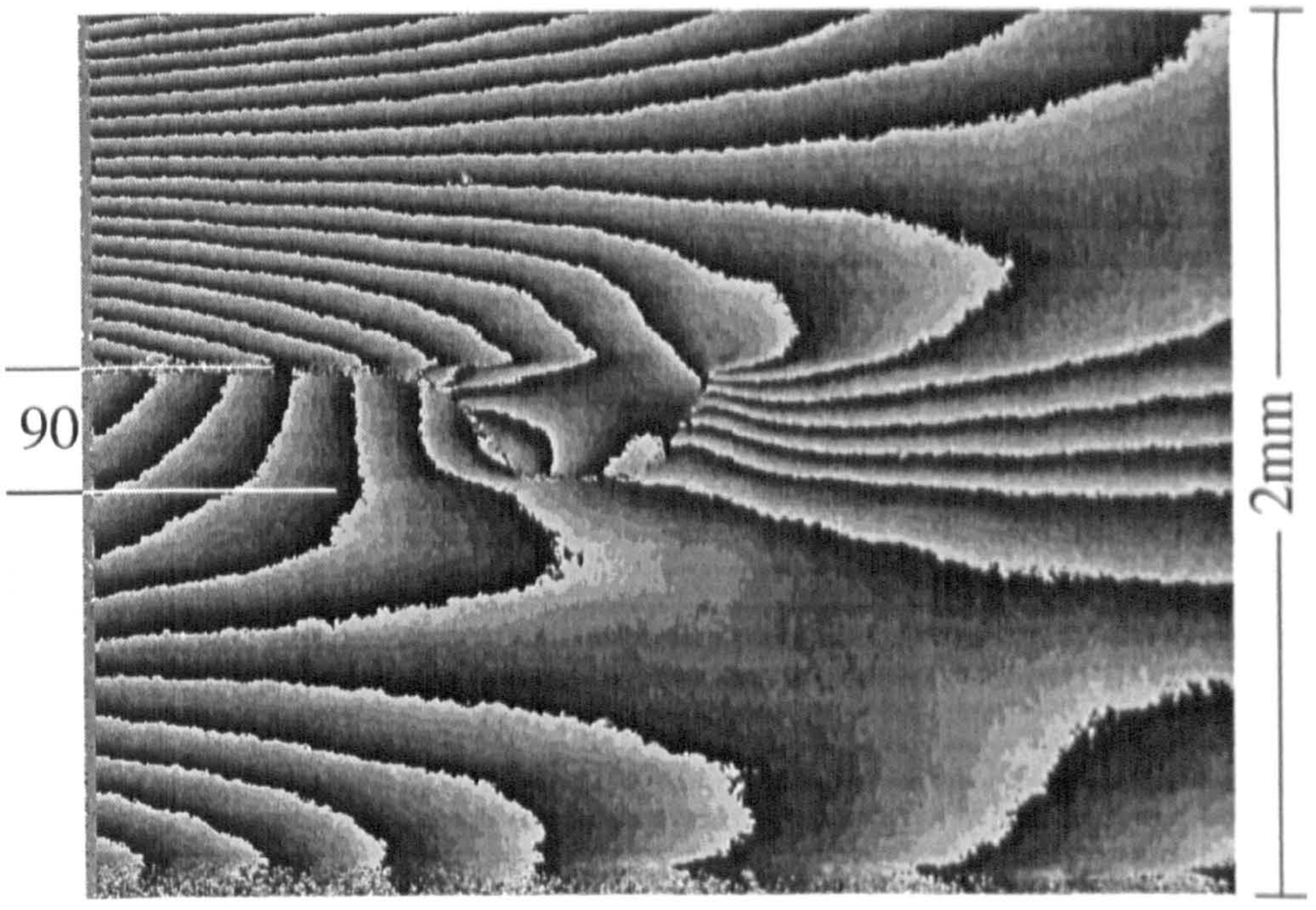


Figure 7.23: u-field fringe fringe pattern for a multilayered ENF specimen with wandering crack in the 90° layer.

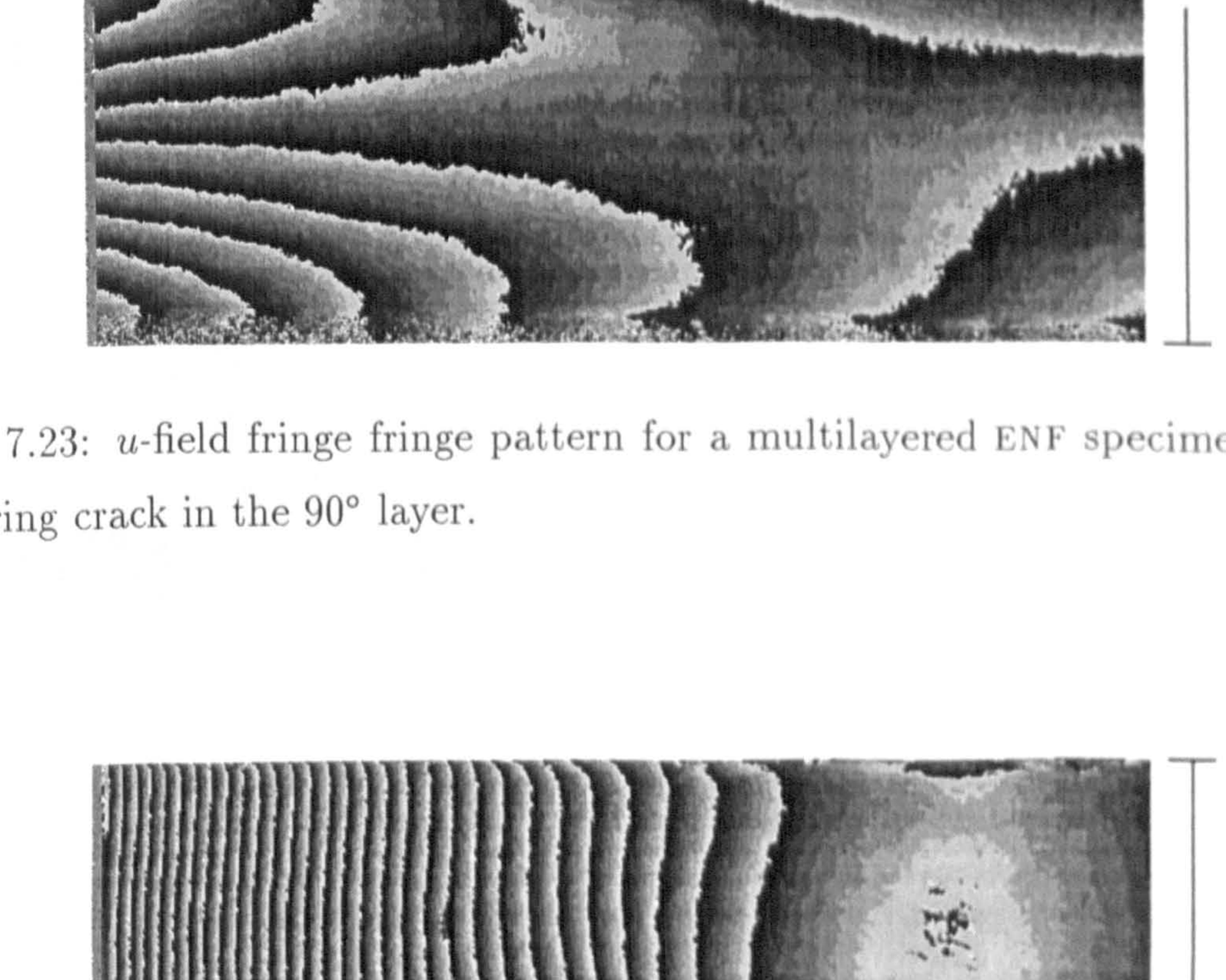


Figure 7.24: v-field fringe fringe pattern for a multilayered ENF specimen with wandering crack in the 90° layer.

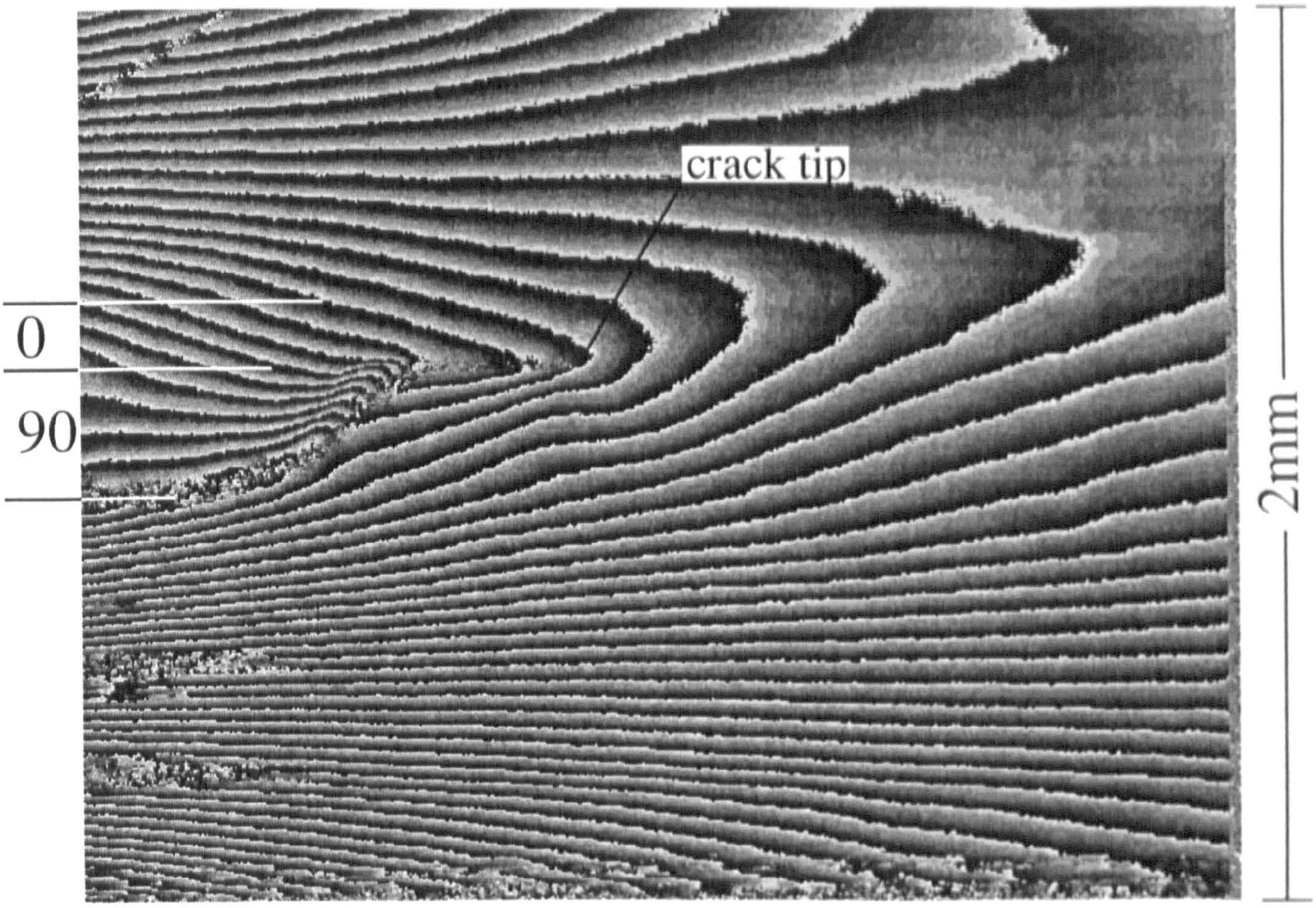


Figure 7.25: u-field fringe pattern for multilayered DCB specimen.

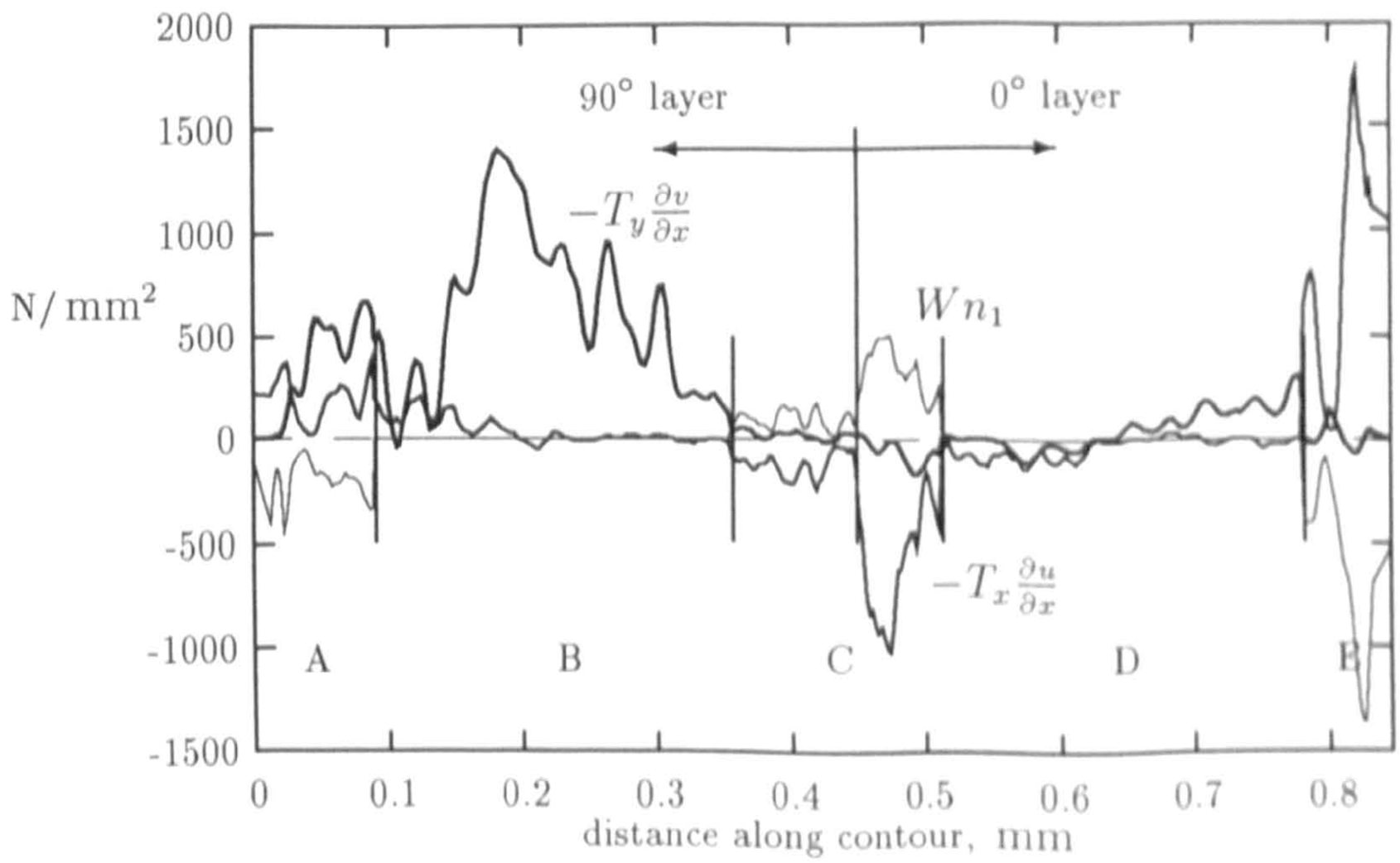


Figure 7.26: Individual contributions in the evaluation of the J-integral for the multilayered DCB specimen.

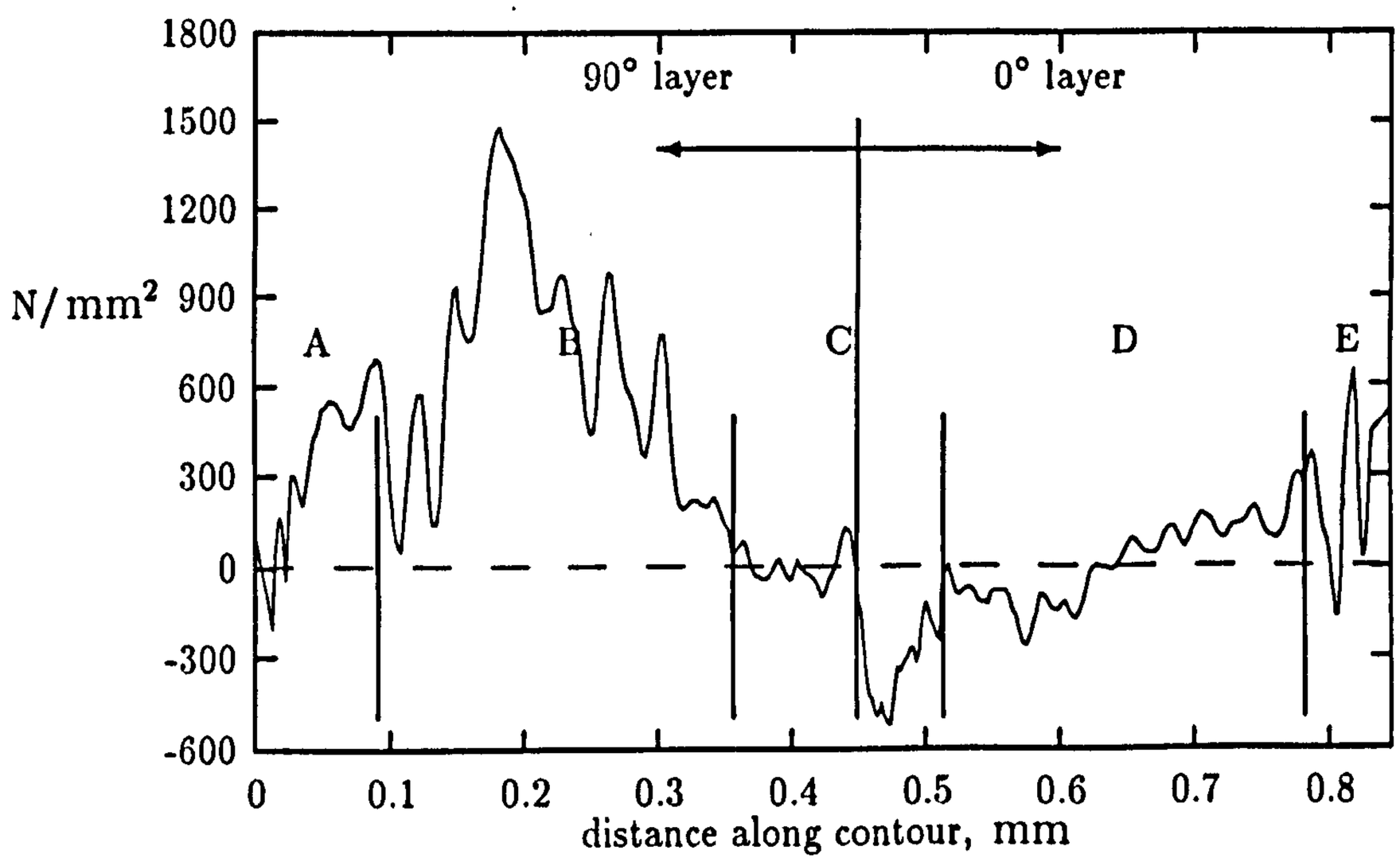


Figure 7.27: The integrand of the J-integral as a function of θ for the multilayered DCB specimen.

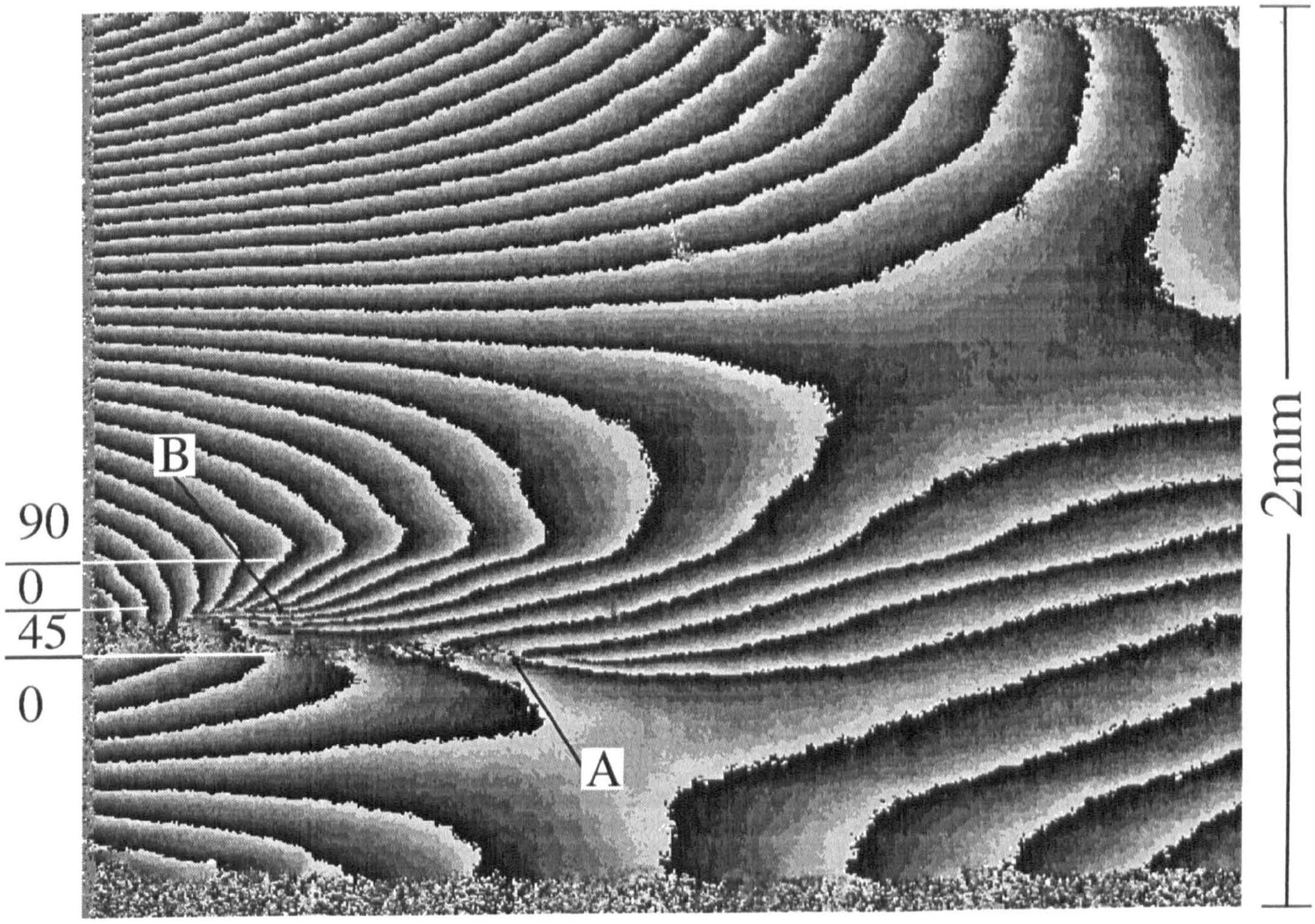


Figure 7.28: u-field fringe pattern for a multilayered ENF specimen illustrating multiple cracking in the 45° lamina.

P, N	G, J/m^2
75	63±8
125	75±12
145	86±9
160	75±11

Table 7.5: J-contour results for the ENF multilayered specimen.

Chapter 8

Discussion and Conclusions

The purpose of this work was to develop an alternative method for measuring the energy release rate in composite laminates using a photomechanics approach. Naturally, an important responsibility associated with this goal was validating the new photomechanics technique against conventional methods. However, delamination testing of laminated composites is a complicated activity, demanding its own separate study. Hence it was necessary to acquire the capacity and appreciation for delamination testing *while* the photomechanics technique was being developed.

During the course of the project, many important issues regarding the mechanics of delamination were encountered. It was not always possible to investigate these issues in sufficient detail. Regardless, some relevant observations about delamination testing were recorded over the course of the project and these are summarized in Section 8.1.

For the successful implementation of the moiré method, it was necessary to develop an image processing system capable of extracting J measurements from a large number of fringe patterns. Such a task naturally encompasses a wide range of issues diffraction grating technology to image processing techniques and software design. A summary of these is presented in Section 8.2.

In Section 8.3, the advantages and limitations of the the moiré method for determining G in composite laminates are discussed. A particular emphasis is placed on recommendations for further research.

8.1 Conventional Delamination Testing

8.1.1 General Comments

It was discovered early in the project that the conditioning of the specimens played a major role in their fracture behaviour. This was no surprise, as it is well known that the mechanical properties of polymeric materials are dependent on such factors as moisture content and temperature history.

For example, during the development of the photoresist grating technique, several specimens were subjected to repeated washing and drying cycles. When they were eventually tested as DCB specimens, they all produced energy release rates on the order of only $70\,J/m^2$. This is in contrast to 'properly' conditioned specimens, which consistently had critical energy release rates around $150\,J/m^2$.

Apart from the obvious environmental effects, there was relatively little scatter in the energy release rates for the DCB experiments. This was not the case, however, with the ENF experiments. There were several possible sources for this lack of consistency in the energy release rates for the ENF experiments.

The most significant factor with respect to the propagation of an ENF delamination was determined to be the *nature* of the *initial* delamination. Obviously, precracked delaminations would propagate at lower energy release rates than those starting from the teflon inserts because of the more natural delamination front. The G values obtained directly from the teflon starter inserts, however, were actually more consistent than any of the methods of precracking could produce.

Careful inspection of NDE scans of precracked ENF specimens revealed that there was a significant variation in the final shape of the delamination front across the width of the specimen. Even with the carefully controlled fatigue precracking procedure, the delamination fronts would need to be inspected using the NDE to insure that they were straight.

The sensitive dependence on the initial delamination profile was probably made more significant by the brittle fracture behaviour of the epoxy matrix. The unidirectional carbon-fibre/epoxy specimens tested during the course of the project always produced smooth fracture surfaces under either mode I or mode II delamination growth. Furthermore, there was never any indication of fibre

bridging in the wake of the delamination.

Another consequence of the brittle fracture behaviour of the specimens was that the empirical correction factors did not play a major role in the reduction of the global test data. For example, the correction factor, n for the DCB experiments was almost identical to the uncorrected beam theory result (compare 3.01 and 3.00). Since the delaminations would propagate at such low load levels, there were no large deformations or rotations to correct for.

The monitoring of the u- and v-field fringe patterns during the experiments provided an unprecedented view of the delamination process. Moiré interferometry fringe patterns are directly related to the displacements on the specimen surface, so the displacement discontinuity produced by a delamination can be located with great precision. By observing the fringe patterns in real time, it was also possible to detect very small deformations and/or rotations of the specimen in the vicinity of the delamination front.

Movement of the fringes was observed well below the loads required to propagate delaminations for many of the specimens tested. This was particularly the case with the ENF experiments. Furthermore, the fringe movement would not always be accompanied by an observed extension of the delamination. This suggests that either the material near the delamination front had undergone plastic deformation while remaining continuous, or that internal delamination growth had occurred.

8.1.2 Mixed Mode Delamination Testing

The redesigned MMB rig provided a very good loading arrangement for the mixed mode testing of laminates. The bearing mounted rollers completely eliminated the problem of friction, and a high degree of repeatability was possible in the loading and unloading of the specimens.

The main advantage of the MMB rig for general mixed mode experiments is that it utilizes a single test rig for a wide range of possible mixed mode ratios. However, it is not clear that from the single experiment performed for this work if the MMB data reduction procedures are capable of providing a consistent conversion between the test parameters (applied loading, geometry of the rig, etc.) and the energy release rate values, G_I and G_{II}.

The good agreement found between the moiré and the global energy release rate values for the two lower G_I/G_{II} ratios in the MMB experiment and the general agreement found in the DCB and ENF results seems to indicate that the slight discrepancy at the higher G_I/G_{II} ratio is an artifact of the actual MMB testing procedure. It is perhaps significant that at higher G_I/G_{II} ratio there is more rotation of the loading bar.

Although no definitive conclusions can be drawn at this time, the ability to measure G using the moiré method has at least provided an independent experimental means to study the MMB test rig in more detail.

It is believed that the mixed mode delamination behaviour of the specimens tested for this work was governed predominately by G_I. At most, it seems that G_{II} would only become important when $G_I/G_{II} \rightarrow 0$.

8.2 Photomechanics

8.2.1 Diffraction Gratings for Moiré Interferometry

Photoresist diffraction gratings produced directly on the specimen surface offer the possibility of near-zero thickness gratings. This is important when discontinuities in the displacement and/or strain fields are to be studied.

A substantial amount of time was spent developing the capability of producing photoresist diffraction gratings directly onto thin edges of carbon-fibre/epoxy specimens. A technique for producing high-efficiency 2000 lines/mm photoresist diffraction gratings was perfected, however, there were two main difficulties.

The holographic exposure time for a single crossed grating was on the order of one hour with the 14 mW laser. This was just too much time to make only one diffraction grating. It was also found to be very difficult to properly polish the carbon-fibre/epoxy specimens. Photoresist diffraction gratings were used in the DCB and ENF validation experiments, but for the above reasons, it was necessary to resort to conventional grating technologies for the remainder of the experiments.

8.2.2 Fourier Fringe Analysis

The Fourier method has the advantage of requiring only one interferogram of each field for the analysis. This makes it particularly attractive for dynamic applications, or instances where vibrations make interferometric measurements difficult. However, because it is based on performing a discrete Fourier transform, the method introduces significant errors near borders and discontinuities.

Early in this project, the technique of optimal filtering was developed to try to overcome these problems. The optimal filter concept was shown to be effective and easy to implement.

However, despite the power of the Fourier method and the improvements afforded by the optimal filter, there still remained grave difficulties associated with edge effects and spectral leakage. This is because the filtering associated with the Fourier method is applied on a global basis—no explicit information about the spatial description of the signal is retained in frequency space. Problems arise when the frequency space representation of the fringe pattern is modified (filtered), and inverse transformed. Regions of the fringe pattern that were necessarily reduced in intensity to enforce cyclic continuity become increasingly sensitive to noise, both random and systematic, and leakage from other parts of the fringe pattern.

Discontinuous fringe patterns arise in many practical circumstances, and their analysis is complicated by a number of factors. Obvious problems arise in coping with 'dead-zones' or regions that contain no data, such as within holes or cracks. In addition, strong transitions in the relative density and orientation of fringes surrounding discontinuities pose further difficulties for the Fourier filtering approach, by increasing the bandwidth of the signal of interest.

Much experience in this project has shown that even with the most judicious implementation, the Fourier method is *not* capable of providing satisfactory measurements near boundaries and discontinuities. For this reason, the alternative technique of phase shifting was implemented.

8.2.3 Phase Shifting Moiré Interferometry

Phase shifting interferometry is both a data collection procedure and an analysis method. It utilizes computer controlled optical components to record fringes

patterns that can be processed in a relatively simple manner to produce noise-reduced displacement maps.

The whole process of collecting data and producing full-field displacement maps (as opposed to displacement contour maps) can be performed on one computer in a matter of minutes, (even seconds, provided the necessary hardware is available). However, it was determined during the course of this work that the process should be divided into two distinct steps—acquisition and processing.

The main reason behind the segregation of the processes is that the computer hardware and user interaction required for each step is distinctly different. Furthermore, it is not necessary nor even advantageous to perform the complete processing of an image before another is acquired. Making a firm distinction between the two steps allows for the development of two independent pieces of software (possibly on different machines) that can be tailored appropriately to each task.

The most significant practical problem encountered during the data collection step was due to mechanical vibrations. Therefore, a numerical study was performed to determine the precise effect of vibration-induced errors on the phase shifting data collection procedure. An additional investigation was then performed to determine the best way to cope with vibrations. The results of this analysis were used to implement a novel phase shifting data collection procedure.

The software responsible for processing the wrapped fringe patterns was designed to be fast and easy to use. Strong graphics support was provided by utilizing an object-oriented window system library, and time consuming image processing algorithms were executed as batch processes. The J-integral routines were developed independently of the main software system and were eventually implemented as user defined sub-routines.

8.3 Energy Release Rates by Moiré Interferometry

The moiré method of measuring the energy release rate was developed for the purpose of validating sub-critical energy release rates produced from computer simulations. The agreement found between the moiré results and the conventional

data reduction methods for the DCB and ENF experiments clearly indicated the validity of the new approach.

8.3.1 General Comments

During the initial validation experiments, it was necessary to insure that the load cell reading was correctly nulled. If there was a discrepancy between the actual load on the specimen and the recorded load, it would result in a constant offset between the moiré predictions of the energy release rate and those obtained from the conventional methods. This is easily appreciated since the moiré method provides an *absolute* measurement of the energy release rate, without recourse to global parameters such as crack length and load.

Because the moiré method represents an absolute measurement of the energy release rate, depending only on the u- and v-field fringe patterns, it is immune to many of the commonly corrected-for side effects of delamination testing (such as large displacement and end-block effects). Furthermore, it will exhibit no mode mixity dependence, which makes an it an ideal tool for mixed mode fracture investigations.

It should be pointed out that while the moiré method was developed specifically for determining energy release rates for delaminations in composites, it is a general technique that could be applied to other types of specimens. However, there are two serious limitations with the moiré method as a general purpose tool.

Obtaining critical energy release rates will be a problem, unless fringe patterns can be captured *at* or *very near* critical loads. Otherwise, the J values extracted from the fringe patterns will underestimate the true critical value of the fracture parameter. It is possible that the unidirectional specimens tested for this work may have been a difficult material system for which to determine critical G values. The fact that the specimens were so brittle and susceptible to delamination growth, may have contributed to the difficulties in capturing fringe patterns near the critical loads.

The second limitation of the method is that only the energy release rate at the surface of the specimen can be measured. Actually, this is not so much a limitation, as it is an opportunity to obtain energy release rates for complicated three-dimensional delaminations (albeit for only at the surface).

Knowing the value of the energy release rate on the edge of the specimen, however, does not provide any information as to what the energy release rate may be along other points along the delamination front on the interior of the specimen. Apart from the agreement found in the DCB, ENF and MMB validation experiments, there is no proof that the energy release rates obtained using the moiré method were representative of the through-thickness G distribution. However, it is worth pointing out that edges present particular problems in computational mechanics—validation at an edge is therefore a good test for the simulation methods.

8.3.2 Extension to Multidirectional Laminates

The bulk of this work was concerned with actually developing and validating the moiré method using the unidirectional laminates. The two experiments performed on the multidirectional laminates were presented as a preliminary demonstration of the new method's extended capabilities.

In order to apply the moiré method to the multidirectional specimens it was necessary to acquire fringe patterns at higher magnifications than for the unidirectional experiments. This was to enable contours to be taken within laminae immediately flanking the delamination. A consequence of the additional magnification, however, was that there was less light available at the camera. This resulted in lower contrast fringe patterns, and hence, larger errors in the determination of the J-integral.

Regardless, the initial energy release rate results for the multidirectional laminates were encouraging, despite the fact that there was no way to know for sure if the values obtained were 'correct'. This situation is to be expected since the moiré method was developed to provide a capability that did not previously exist in experimental form. There are several measures that should be taken, however, to provide a more sound basis for the application of the method to multidirectional laminates.

A more careful study of the strains and stresses at the interface between the two materials is needed. The J-integral may be perfectly valid for this type of situation *in theory*, but the calculation of strains and stresses may introduce errors that overwhelm the true solution. One way to perform such an analysis would be

to study the displacement and strain fields from a 'model' material interface with a reasonable mathematical solution. Such a study may reveal that the present approach yields suitable results with respect to the interface stress fields. This would mean that the contours used to evaluate J could be taken more freely, and allowed to cross an arbitrary number of material interfaces.

Also in experimental terms, the use of diffraction gratings of higher efficiency and a more powerful laser would undoubtably improve the process. The laser power represents no particular technical problem and it is understood that as a result of this project, tailor made diffraction gratings have been sourced for use in future work.

Bibliography

[1] R.M. Jones. *Mechanics of Composite Materials.* McGraw-Hill, 1975.

[2] C.T. Herakovich. Edge effects and delamination failures. *Journal of Strain Analysis*, 24(4), 1989.

[3] R.B. Pipes and N.J. Pagano. Interlaminar stress in composite laminates under uniform axial extension. *Journal of Composite Materials*, 4:538–548, 1970.

[4] A.A. Griffith. The phenomena of flow and rupture in solids. *Philosophical Transactions of the Royal Society of London*, A221:163–197, 1921.

[5] G.R. Irwin. Structural aspects of brittle fracture. *Applied Materials Research*, 3:65–81, 1964.

[6] R.E. Smelser and M.E. Gurtin. On the J-integral for bimaterial bodies. *International Journal of Fracture*, 13:RCR 382–384, 1977.

[7] D. Post. Moiré interferometry. In A. S. Kobyashi, editor, *Handbook on Experimental Mechanics*, chapter 7, pages 314–383. Prentice-Hall, Inc., 1987. Society of Experimental Mechanics.

[8] R.B. Pipes and I.M. Daniel. Moiré analysis of the interlaminar shear edge effect in laminated composites. *Journal of Composite Materials*, 6:255–259, April 1971.

[9] T.G.F. Gray, J. McKelvie, P.M. MacKenzie, and C.A. Walker. Interferometric measurement of J for arbitrary geometry and loading. *International Journal of Fracture*, 24:R 109–114, 1984.

[10] B.S.-J. Kang and A.S. Kobayashi. *J* resistance curves in aluminium SEN specimens using moiré interferometry. *Experimental Mechanics*, 28(2):154–158, 1988.

[11] M. Dadkah and A.S. Kobayashi. HRR field of a moving crack, an experimental analysis. *Engineering Fracture Mechanics*, 34(1):253–262, 1989.

[12] D. Post, R. Czarnek, and D. Joh. Shear strains in a graphite/PEEK beam by moiré interferometry with carrier fringes. *Experimental Mechanics*, 27(3):246–249, 1987.

[13] B. Han. Higher sensitivity moiré interferometry for micromechanics studies. *Optical Engineering*, 31(7):1517–1526, July 1992.

[14] K.M. Liechti. Moiré of crack opening interferometry in adhesive fracture mechanics. *Experimental Mechanics*, 25:255–261, September 1985.

[15] K. Takahashi, K. Arakawa, and M. Ishiguma. Interlaminar crack tip opening displacement of CFRP studied by moiré interferometry, 1992. Proceedings of the 1992 BCS/EACM Conference on Composite Materials, Amsterdam.

[16] J.P. Berry. Determination of fracture surface energies by the cleavage technique. *Journal of Applied Physics*, 34(1):62–68, January 1963.

[17] D.F. Devitt, R.A. Schapery, and W.L. Bradley. A method for determining the mode I delamination fracture toughness of elastic and viscoelastic composite materials. *Journal of Composite Materials*, 14:270–285, October 1980.

[18] K.E. Bisshopp and D.C. Drucker. Large deflections of cantilever beams. *Quarterly of Applied Math*, III:272, 1945.

[19] W.D. Bascom, J.L. Bitner, R.J. Moulton, and A.R. Siebert. The interlaminar fracture of organic-matrix woven reinforced composites. *Composites*, pages 9–18, January 1980.

[20] D.J. Wilkins, J.R. Eisenmann, R.A. Camin, and W.S. Margolis. Characterizing delamination growth in graphite-epoxy. In K.L. Reifsneider, editor, *Damage in Composite Materials*, pages 168–183. American Society for Testing and Materials, Philadelphia, 1982. ASTM STP 775.

[21] J.M. Whitney, C.E. Browing, and W. Hoogsteden. A double cantilever beam test for characterizing mode I delamination of composite materials. *Journal of Reinforced Plastics and Composites*, pages 297–313, October 1982.

[22] S.P. Timeshenko and J.N. Goodier. *Theory of Elasticity.* McGraw-Hill Book Company, 1934.

[23] P.E. Keary, L.B. Ilcewicz, C. Schaar, and J. Trostle. Mode I fracture toughness of composites using slender DCB specimens. *Journal of Composite Materials*, 19:164–177, March 1985.

[24] J.G. Williams. Large displacement and end block effects in the DCB interlaminar test in mode I and II. *Journal of Composite Materials*, 21:330–347, April 1987.

[25] J.G. Williams. The analysis of interlaminar fracture in uniaxial fibre-polymer composites. *Proceedings of the Royal Society of London*, A 427:173–199, 1990.

[26] J.G. Williams. The fracture mechanics of delamination tests. *Journal of Strain Analysis*, 24(4):207–214, 1989.

[27] M.F. Kanninen. An augmented double cantilever beam model for studying crack propagation and arrest. *International Journal of Fracture*, 9(1):83–91, March 1973.

[28] C.E. Browning, F.L. Abrams, and J.M. Whitney. A four point shear test for graphite/epoxy composites. In C.E. Browning, editor, *Composite Materials: Quality Assurance and Processing*, pages 54–74. American Society for Testing and Materials, Philadelphia, 1983. ASTM STP 797.

[29] A.J. Russell and K.N. Street. Moisture and temperature effects on the mixed mode delamination fracture of unidirectional graphite/epoxy. In W.S. Johnson, editor, *Delamination and Debonding of Materials*, pages 349–370. American Society for Testing and Materials, Philadelphia, 1985. ASTM STP 876.

[30] L.A. Carlsson, J.W. Gillespie, Jr., and R.B. Pipes. On the analysis and design of the end-notched flexure(ENF) specimen for mode II testing. *Journal of Composite Materials*, 20:594–604, November 1986.

[31] L.A. Carlsson, J.W. Gillespie, Jr., and B.R. Trethewey. Mode II interlaminar fracture of graphite/epoxy and graphite/peek. *Journal of Reinforced Plastics and Composites*, 5:106–123, 1986.

[32] Y. Wang and J.G. Williams. Corrections for mode II fracture toughness specimens of composite materials. *Composites Science and Technology*, 43:251–256, 1992.

[33] C.R. Corletto and W.L. Bradley. Mode II delamination fracture toughness of unidirectional graphite/epoxy composites. In P.A. Lagace, editor, *Composite Materials: Fatigue and Fracture*, pages 201–221. American Society for Testing and Materials, Philadelphia, 1989. ASTM STP 1012.

[34] R.A. Jurf and R.B. Pipes. Interlaminar fracture of composite materials. *Journal of Composite Materials*, 16:386–394, September 1982.

[35] M. Arcan, Z. Hashin, and A. Voloshin. A method to produce uniform plane-stress states with applications to fiber-reinforced materials. *Experimental Mechanics*, pages 141–146, April 1978.

[36] T.K. O'Brien. Characterization of delamination onset and growth in a composite laminate. In K.L. Reifsneider, editor, *Damage in Composite Materials*, pages 140–167. American Society for Testing and Materials, Philadelphia, 1982. ASTM STP 775.

[37] T.K. O'Brien. Mixed mode strain-energy-release rate effects on edge delamination of composites. In D.J. Wilkins, editor, *Effects of Defects in Composite Materials*, pages 125–142. American Society for Testing and Materials, Philadelphia, 1983. ASTM STP 836.

[38] E.F. Rybicki, T.D. Hernandez, Jr., J.D. Deibler, R.C. Knight, and S.S. Vinson. Mode I and mixed mode energy release rate values for delamination of graphite/epoxy test specimens. *Journal of Composite Materials*, 21:105–123, February 1987.

[39] J.G. Williams. On the calculation of energy release rates for cracked laminates. *International Journal of Fracture*, 30:101–119, 1988.

[40] J.G. Williams. On the analysis of mixed mode failure. *International Journal of Fracture*, 54:269–291, 1992.

[41] J.R. Reeder and J.H. Crews, Jr. Mixed mode bending method for delamination testing. *AIAA Journal*, 28(7), July 1990.

[42] A.A. Aliyu and I.M. Daniel. Effects of strain rate on delamination fracture toughness of graphite/epoxy. In W.S. Johnson, editor, *Delamination and Debonding of Materials*, pages 336–348. American Society for Testing and Materials, Philadelphia, 1985. ASTM STP 876.

[43] J.R. Reeder and J.H. Crews, Jr. Nonlinear analysis and redesign of the mixed mode bending delamination test. Technical report, National Aeronautics and Space Administration, Langely Research Center, Hampton Virginia, 1986. NASA Technical Memorandum 102777.

[44] F.X. de Charentenay, J.M. Harry, Y.J. Prel, and M.L. Benzeggagh. Characterizing the effect of delamination defect by mode I delamination test. In D.J. Wilkins, editor, *Effects of Defects in Composite Materials*, pages 84–103. American Society for Testing and Materials, Philadelphia, 1983. ASTM STP 836.

[45] P. Davies. Interlaminar fracture testing of composites. Technical report, European Group on Fracture: Polymers and Composites Task Group, April 1990. Protocol for joint round robin EGF/ASTM/JIS.

[46] P. Davies et al. Round-robin interlaminar fracture testing of carbon-fibre/epoxy and PEEK composites. *Composites Science and Technology*, 43:129–136, 1992.

[47] V.J. Parks. Geometric moiré. In A. S. Kobyashi, editor, *Handbook on Experimental Mechanics*, chapter 6, pages 282–313. Prentice-Hall, Inc., 1987. Society of Experimental Mechanics.

[48] J.D. Gaskill. *Linear Systems, Fourier Transforms, and Optics.* John Wiley & Sons, 1978.

[49] E. Hecht. *Optics.* Addison-Wesly Publishing Co., 1988.

[50] M. Takeda, H. Ina, and S. Kobayashi. Fourier transform method of fringe pattern analysis for computer-based topography and interferometry. *Journal of the OpticalSociety of America*, 72(1):156–160, January 1982.

[51] D.J. Bone, H.-A. Bachor, and J. Sandeman. Fringe pattern analysis using a 2-d Fourier transform. *Applied Optics*, 25(10):1653–1660, May 1986.

[52] J.M. Huntley and J.E. Field. High resolution moiré photography: Application to dynamic stress analysis. *Optical Engineering*, 28(8):926–933, August 1989.

[53] D.J. Bone. Fourier fringe analysis: The two-dimensional phase unwrapping problem. *Applied Optics*, 25(30):3627–3632, September 1991.

[54] W.H. Press, B.P. Flannery, S.A. Teukolsky, and W.T. Vetterling. *Numerical Recipes in C.* Cambridge University Press, Cambridge, 1988.

[55] J.E. Greivenkamp and J.H. Bruning. Phase shifting interferometry. In Daniel Malacara, editor, *Optical Shop Testing*, chapter 14, pages 501–598. John Wiley & Sons, New York, 1986. Series in Pure and Applied Optics.

[56] K. Creath. Phase-measurement interferometry techniques. In E. Wolf, editor, *Progress in Optics*, volume XXVI, pages 349–393. Elsevier Science Publishers, Amsterdam, 1986.

[57] P. Carré. Installation et utilisation du compateur photoelectrique et interferential du bureau international des poids et mesures. *Metrologia*, 2(13):13–23, 1966.

[58] J.E. Greivenkamp. Generalized data reduction for heterodyne interferometry. *Optical Engineering*, 23(4):350–352, July 1984.

[59] Y.Y. Cheng and J.C. Wyant. Phase shifter calibration in phase shifting interferometry. *Applied Optics*, 24(18):3049–3052, September 1985.

[60] J. Schwider, R. Burrow, K.-E. Elssner, J. Grzanna, R. Spolaczyk, and K. Merkel. Digital wavefront measuring interferometry: Some systematic error sources. *Applied Optics*, 22(21):3421–3432, November 1983.

[61] C. Ai and J.C. Wyant. Effect of piezoelectric transducer nonlinearity on phase shift interferometry. *Applied Optics*, 26(6):1112–1116, March 1987.

[62] P. Hariharan, B.F. Oreb, and T. Eiju. Digital phase-shifting interferometry: A simple error-compensating phase calculation algorithm. *Applied Optics*, 26(13):2504–2505, July 1987.

[63] J. Schwider. Phase shifting interferometry: Reference phase error reduction. *Applied Optics*, 28(18):3889–3892, September 1989.

[64] C.P. Brophy. Effect of error correlation on the computed phase of phase shifting interferometry. *Journal of the OpticalSociety of America*, 7(4):537–541, April 1990.

[65] K. Kinnstaetter, A.W. Lohmann, J. Schwider, and N. Streibl. Accuracy of phase shifting interferometry. *Applied Optics*, 27(24):5082–5089, December 1988.

[66] J.J. Snyder. Algorithm for fast digital analysis of interference fringes. *Applied Optics*, 19(8):1223–1225, April 1980.

[67] G. Lai and T. Yatagi. Generalized phase shifting interferometry. *Journal of the OpticalSociety of America*, 8(5):822–827, May 1991.

[68] G.D. Lassahn, J.K. Lassahn, P.L. Taylor, and V.A. Deason. Multiphase fringe analysis with unknown phase shifts. personal communication, 1993.

[69] C.T. Farrell and M.A. Player. Phase step measurement and variable step algorithms in phase shifting interferometry. *Measurement Science and Technology*, 3:953–958, March 1992.

[70] K. Okada, A. Sato, and J. Tsujiuchi. Simultaneous calculation of phase distribution and scanning phase shift in phase shifting interferometry. *Optics Communications*, 84(3,4):118–124, July 1991.

[71] J. Schwider. Advanced evaluation techniques in interferometry. In E. Wolf, editor, *Progress in Optics*, volume XXVIII, pages 273–359. Elsevier Science Publishers, Amsterdam, 1988.

[72] R.M. Goldstein, H.A. Zebker, and C.L. Werner. Satellite radar interferometry: Two dimensional phase unwrapping. *Radio Science*, 23(4):713–720, July-August 1988.

[73] J.M. Huntley. Noise-immune phase unwrapping algorithm. *Applied Optics*, 28(15):3268–3270, August 1989.

[74] R.W. Hamming. *Digital Filters.* Prentice Hall, Engelwood Cliffs, NJ, 1983.

[75] *C.H. Reinsch. Smoothing by spline functions. Numerische Mathematik, 10:177–183, 1967.*

[76] G.E. Forsythe, M.A. Malcolm, and C.B. Moler. *Computer Methods for Mathematical Computations.* Prentice–Hall, Inc., Englewood Cliffs, NJ, 1977.

Acknowldgements

The author is indebted to David Nash and Douglas Brennan for their kind help and state-of-the-art resources. The technical assistance of John Low and Ronnie McKenzie was also appreciated, and a special thanks is due to Jack Sivewright for his assistance throughout the course of the project.

This research would not have been possible without the opportunity provided by Dr. J. McKelvie. His encouragement and patient guidance provided the ideal environment in which this work was performed. The author is forever greatful.