

UNCLASSIFIED

2 Rec'd 21 Jul 89

SECURITY CLASSIFICATION OF THIS PAGE

REPORT DOCUMENTATION PAGE

Form Approved OMB No. 0704-0188

1a. REPORT SECURITY CLASSIFICATION UNCLASSIFIED 1b. RESTRICTIVE MARKINGS

AD-A211 565

3. DISTRIBUTION/AVAILABILITY OF REPORT Approved for public release; distribution unlimited.

5. MONITORING ORGANIZATION REPORT NUMBER(S) AFOSR-TR-89-1181

6a. NAME OF PERFORMING ORGANIZATION California Institute of Technology

6b. OFFICE SYMBOL (if applicable) E

7a. NAME OF MONITORING ORGANIZATION The Air Force Office of Scientific Research

6c. ADDRESS (City, State, and ZIP Code) 1201 E. California Boulevard Pasadena, CA 91125

7b. ADDRESS (City, State, and ZIP Code) Building 410 Bolling AFB, Washington DC 20332-6448

8a. NAME OF FUNDING / SPONSORING ORGANIZATION AFOSR

8b. OFFICE SYMBOL (if applicable) NA

9. PROCUREMENT INSTRUMENT IDENTIFICATION NUMBER Grant No. AFOSR-84-0254

8c. ADDRESS (City, State, and ZIP Code) AFOSR/NA Bolling AFB DC 20332-6448

10. SOURCE OF FUNDING NUMBERS PROGRAM ELEMENT NO. 61100F PROJECT NO. 2302 TASK NO. B2 WORK UNIT ACCESSION NO.

11. TITLE (Include Security Classification) AN EXPERIMENTAL AND ANALYTICAL PROGRAM TO DEVELOP CRACK TIP FRACTURE CRITERIA (U)

PERSONAL AUTHOR(S) DBS W. G. KNAUSS, C. R. SCHULTHEISS

13a. TYPE OF REPORT Final Report

13b. TIME COVERED FROM 840901 TO 880301

14. DATE OF REPORT (Year, Month, Day) July 19, 1989

15. PAGE COUNT 169

16. SUPPLEMENTARY NOTATION

Table with 3 columns: FIELD, GROUP, SUB-GROUP

18. SUBJECT TERMS (Continue on reverse if necessary and identify by block number)

19. ABSTRACT (Continue on reverse if necessary and identify by block number) Please see Reverse

DTIC ELECTE AUG 21 1989 S E D

20. DISTRIBUTION/AVAILABILITY OF ABSTRACT UNCLASSIFIED/UNLIMITED SAME AS RPT. DTIC USERS

21. ABSTRACT SECURITY CLASSIFICATION UNCLASSIFIED

22a. NAME OF RESPONSIBLE INDIVIDUAL George K. Harkness

22b. TELEPHONE (Include Area Code) (202) 767-0463

22c. OFFICE SYMBOL NA

Abstract:

The large plastic deformations at the tip of a crack in a ductile heat treatment of 4340 steel are studied numerically and experimentally. The numerical simulation of the experiment uses a small strain, incremental plasticity law, with a power law hardening behavior. Both the in-plane and out-of-plane deformations are measured on the same specimen at the same time.

The experimental technique of moire interferometry is used to measure the in-plane displacements. This technique is described in detail, including an analysis of the effect of out-of-plane rotations on the use of the technique. A three beam interferometer and a four beam interferometer are compared.

The out-of-plane displacements are measured with a Twyman-Green interferometer. The analysis of this data is still underway and will be forwarded when completed.

The analysis of the data from the experiment is analyzed through a digital image processor. The hardware and software are briefly described. (Further information is available in previous reports submitted under this contract.)

The numerical model is described in detail. The material properties are determined from a uniaxial test on specimens taken from the same bar as the fracture specimens and with identical heat treatment. A numerical model of the fracture specimen having 7581 degrees of freedom is used to provide data for comparison between 400 N (linear behavior) up to 75000 N (catastrophic failure of the steel specimen) in fourteen steps. This model characterizes the crack as a notch having a similar size as the notch in the steel fracture specimen. Data from a previous study is also available, but at only five load levels; this model uses a sharp crack.

Comparison between the experiment and the numerical model is shown at two load steps, 4000 N, where plastic deformation is contained within about 0.01 cm, and 30000 N, where the plastic deformation is on the order of 0.2 cm.

Final Technical Report to
The Air Force Office Of Scientific Research
On

AN EXPERIMENTAL AND ANALYTICAL PROGRAM
TO DEVELOP CRACK TIP FRACTURE CRITERIA

Grant No. AFOSR-84-0254

by

C. R. Schultheisz
W. G. Knauss

June 26, 1989

Graduate Aeronautical Laboratories
California Institute of Technology
Pasadena, CA 91125

AFOSR-TR-89-1181

Accession For	
NTIS GRA&I	<input checked="" type="checkbox"/>
DTIC TAB	<input checked="" type="checkbox"/>
Unannounced	<input type="checkbox"/>
Justification	
By _____	
Distribution/	
Availability Codes	
Dist	Avail and/or Special
A-1	



Abstract:

The large plastic deformations at the tip of a crack in a ductile heat treatment of 4340 steel are studied numerically and experimentally. The numerical simulation of the experiment uses a small strain, incremental plasticity law, with a power law hardening behavior. Both the in-plane and out-of-plane deformations are measured on the same specimen at the same time.

The experimental technique of moire interferometry is used to measure the in-plane displacements. This technique is described in detail, including an analysis of the effect of out-of-plane rotations on the use of the technique. A three beam interferometer and a four beam interferometer are compared.

The out-of-plane displacements are measured with a Twyman-Green interferometer. The analysis of this data is still underway and will be forwarded when completed.

The analysis of the data from the experiment is analyzed through a digital image processor. The hardware and software are briefly described. (Further information is available in previous reports submitted under this contract.)

The numerical model is described in detail. The material properties are determined from a uniaxial test on specimens taken from the same bar as the fracture specimens and with identical heat treatment. A numerical model of the fracture specimen having 7581 degrees of freedom is used to provide data for comparison between 400 N (linear behavior) up to 75000 N (catastrophic failure of the steel specimen) in fourteen steps. This model characterizes the crack as a notch having a similar size as the notch in the steel fracture specimen. Data from a previous study is also available, but at only five load levels; this model uses a sharp crack.

Comparison between the experiment and the numerical model is shown at two load steps, 4000 N, where plastic deformation is contained within about 0.01 cm, and 30000 N, where the plastic deformation is on the order of 0.2 cm.

Table of Contents

1. Introduction	1
2. Moire Interferometry	5
2.1 Preview	5
2.2 Moire Interferometry and Geometric Moire	7
2.3 Moire Interferometry Allowing for Arbitrary Changes in the Orientation of the Grating	14
2.4 Experimental Apparatus	45
2.5 Orthogonal Displacement Measurements Using a Three Beam Moire Interferometer	69
3. Digital Image Processing	88
3.1 Introduction	88
3.2 Hardware and Software	89
3.3 Fringe Tracing Approach	91
4. Plastic Fracture	94
4.1 Elastic and Plastic Constitutive Relations	94
4.2 Uniaxial Test Results	103
4.3 Finite Element Analysis	112
4.4 Three Point Bend Experiment	130
5. Comparison of Results	136
References	166

SECTION 1

Introduction

A crack in a body dominates the nature of the stress and strain fields near its tip, and within the framework of two-dimensional linear elastostatics, the distributions of stress and strain are independent of the geometry of the specimen and loading except for a multiplying factor, the *stress intensity factor*, K , and completely independent of the specimen composition. This notion of crack tip autonomy, which leads to a universal square root singularity at the crack tip, has, in the past, left only the calculation of the stress intensity factor, K , as the main effort of fracture mechanics research [24,26]. For a problem which has only load boundary conditions, the stress intensity factor is a function only of the geometry and magnitude of the applied loading, and is independent of material properties. The displacement field does depend on the material properties, and for a problem including displacement boundary conditions the stress intensity factor is also a function of the material properties.

The stress intensity factor does describe the conditions near the crack tip for many applications when the applied loads are small, such as for fatigue, but it is clear that it does not represent the correct solution in most cases. The K -field solution is derived for two-dimensional problems, plane strain or plane stress. These two cases may be thought of as the limiting cases of infinitely thick and infinitesimally thin bodies. While the two-dimensional solutions for the in-plane displacement field quantities are identical to each other except for a factor depending on the Poisson's

ratio ν of the material, the out-of plane displacement are zero in the plane strain case and unbounded in the plane stress case. This implies that for bodies of finite thickness the field quantities will differ from the two-dimensional K -field because of three-dimensional effects.

It is recognized that the stress-strain response at the tip of the crack must be non-linear. Since the stresses predicted within the linear theory of elasticity are singular, the material at the crack tip must act in a non-linear way to mitigate the effect of the stress singularity through plastic deformation. Thus, while the stress intensity factor has been an important tool of fracture mechanics investigations up to now, it is clear that it is only a starting point for the characterization of the fracture process. In an effort to emphasize the realities of the fracture process, this study investigates the plastic deformation at the tip of a crack in a ductile material, in this case, a ductile heat treatment of 4340 steel.

There are several analytical models of the nonlinear plastic behavior at the crack tip, from the simple Dugdale-Barenblatt cohesive zone model to the power law hardening model of Hutchinson [6] and Rice and Rosengren [20]. Along with high speed computers have come numerous finite element and other numerical approximations. These models make assumptions about the constitutive nature of materials and attempt to predict the nature of the fracture process on the basis of those assumptions. Such assumptions must be evaluated experimentally, by comparing the analytical or numerical results with the real world. Actual displacement, strain and stress distributions must be determined experimentally and compared with the predictions to test the validity of the models.

The current study has been undertaken to determine an experimental technique capable of measuring physical quantities in a region surrounding a crack tip, to use

that technique to evaluate the numerical simulations, and to make a comparison between the experimentally measured quantities and the quantities predicted from a state of the art three-dimensional plastic numerical model. The experimental measurements in this study are obtained through a technique called moire interferometry developed by Post [18,19]. This optical technique measures relative in-plane displacements over an entire field of view. The displacements are revealed as fringes which represent contours of constant in-plane displacement parallel to one axis of an imposed coordinate system. Both orthogonal components of in-plane displacement are measured simultaneously over the same field on one side of the specimen, while the out-of-plane displacement component can be measured from the reverse side using a Twyman-Green interferometer.

There are many size scales of importance in this problem, and the measurement technique must be able to encompass all of them. There is an external or *large* scale, within which the field quantities are well described by linear elastostatics. There is the scale of the fracture process zone, which, for the ductile material in this study which introduces plastic yielding, can become large compared with the dimensions of the specimen (on the order of one centimeter). Additionally, there is the scale of the sensitivity of the measurement technique, which must be fine enough to detect discrepancies between the experiment and analytical or numerical predictions.

One limitation on the use of the moire interferometer is that it is sensitive to both in-plane and out-of-plane displacement gradients. However, if the in-plane displacement gradients are large, the effect of the out-of-plane displacement gradients may be neglected. The effect of the out-of-plane displacement gradients is most apparent in the formation of a shadow spot around the crack tip, due to limitations in the recording optics. Large in-plane displacements will themselves result in a

shadow spot, also due to limitations of the recording optics, but this occurs at a higher load level than the shadow spot due to the out-of-plane deformation.

SECTION 2

Moire Interferometry

2.1 Preview

The initial section is a discussion of the technique of moire interferometry, and its relation to geometric moire techniques. The equations describing the fringes in geometric moire will be given; these equations indicate that the fringes represent contours of constant in-plane displacement. The equations describing the fringes in moire interferometry will be compared to those for geometric moire.

Second, the governing equations for a two beam interferometer will be derived for the general problem of interference between two diffracted plane waves. This derivation will then be made more specific to the case of moire interferometry. Planar deformation will be discussed, and the correspondence between the equations describing the fringes in moire interferometry and the equations describing fringes observed in geometrical moire will be shown, leading to the conclusion that the fringes observed in moire interferometry can be taken as contours of constant relative in-plane displacement for planar deformation. The effects of non-planar deformation will be investigated under the assumption of small rotations of the surface away from the original plane of the grating.

Third, the four beam interferometer used in the experiments will be described, as will the procedure for aligning the interferometer during an experiment. The requirements for coherence, polarization and matching of beam intensity will be touched upon. The optical collection apparatus will be discussed, and the theoretical limits of the measurement system will be estimated.

Fourth, the performance of a three beam interferometer will be compared with that of the four beam interferometer. The governing equations describing the fringes for the three beam interferometer will be derived for the case of edit planar deformation.

2.2 Moire Interferometry and Geometric Moire

Moire interferometry is an experimental technique developed by Dr. Daniel Post at Virginia Polytechnic Institute and State University for measuring in-plane displacements at the surface of an object [9,17,18,19]. This technique relies upon the interference between two laser beams diffracted from the specimen surface to infer the in-plane deformations at the surface. The sensitivity of the measurement is on the order of the wavelength of light.

The equations describing the fringes observed using the moire interferometer will be derived. It will be shown that the equations describing the fringes in moire interferometry are identical to the equations describing the fringes in geometric moire when the deformation remains strictly in a plane. (Moire interferometry is insensitive to translations, but it is sensitive to rotations of the specimen surface away from the original plane of the surface.) This identity between the two sets of equations indicates that the fringes observed in moire interferometry can be taken as contours of constant relative in-plane displacement as they are in geometric moire, even though the fringes are generated by very different mechanisms.

The fringes observed in moire interferometry are generated by interfering light waves which intersect at a small angle. This angle varies with the in-plane displacement through the change in the state of a diffraction grating on the surface of the specimen. However, the angle may also vary due to rotations of the surface out of the plane, as will be the case in fracture problems, which may have sizable out-of-plane displacement gradients near the tip of the crack. These rotations will introduce additional terms in the equations describing the fringes, which, if significant will mean that the fringes can no longer be viewed as displacement contours. The equations describing the fringes will be derived allowing for rotations

which change the plane of the grating in an attempt to estimate the magnitude of the error introduced if the fringes are taken as contours of displacement. Alternatively, this investigation will help to determine how corrections could be made if the out-of-plane displacement gradients are known through some other method, as, for example, from measurements using the Twyman-Green interferometer.

It is desirable to be able to view the fringes observed using the moire interferometer as contours of displacement, because then similar data reduction methods can be used on data generated from either the Twyman-Green interferometer or the moire interferometer. The fringes observed in a Twyman-Green interferometer represent contours of constant relative out-of-plane displacement, if the slope of the surface is not too large. In the Twyman-Green interferometer, a path length difference caused by out-of-plane deformation causes interference between two laser beams, one reflected from the specimen and one reflected from a flat reference mirror. Along a given fringe, the phase difference of the interfering light waves is the same, indicating that the path length difference is also the same, which indicates that the out-of-plane displacement is constant along that fringe.

2.2.1 Geometric Moire

The use of geometric moire methods requires a grating consisting of regularly spaced lines inked or scribed upon the specimen. A similarly spaced set of lines is inked or scribed upon a transparent medium; this is called the reference or master grating. These lines are usually straight lines parallel to one axis, but they can be concentric circles or radial lines. The reference grating is usually identical to the specimen grating before deformation, but sometimes the specimen grating has a spacing which is an integral multiple of the reference grating. The fringes in geometric moire are

patterns which result when the deforming specimen grating is viewed through the reference grating. [2.3.4.23] The spacing and inclination of the fringes are related to the pitch p_0 of the reference grating, the pitch p of the specimen grating and the angle θ between the lines of the two gratings. The spacing of the fringes is δ , which is found as

$$\delta = \frac{pp_0}{[p^2 + p_0^2 - 2pp_0 \cos \theta]^{\frac{1}{2}}} \quad (2.2.1)$$

The inclination of the fringes is described by the angle ϕ , which is the angle between the fringes and the lines of the reference grating.

$$\tan \phi = \frac{p_0 \sin \theta}{p_0 \cos \theta - p} \quad (2.2.2)$$

The increment of displacement between adjacent fringes is p_0 , and the direction of the displacement measured is perpendicular to the lines of the reference grating.

2.2.2 Moire Interferometry

If the fringes generated by the moire interferometer do correspond to displacement contours, the fringes will be described by equations which are identical to the equations describing fringes generated in geometrical moire. In the case of the moire interferometer, the equations describing the fringes will be written in terms of the initial (undeformed) state of the grating (which is retained by the arrangement of the incident laser beams) and in terms of the the current state of the grating. The grating itself is a periodic surface height variation which has an initial wavelength of w_0 and a current wavelength of w . Let the angle between the current grating ridges relative to the initial grating ridges be γ . The fringe spacing δ in the case of strictly planar deformation is found to be

$$\delta = \frac{ww_0}{2[w^2 + w_0^2 - 2ww_0 \cos \gamma]^{\frac{1}{2}}} \quad (2.2.3)$$

Note that there is a fringe multiplication factor of two in the denominator of equation (2.2.3) compared with equation (2.2.1). This produces twice as many fringes as expected from equation (2.2.1), or, equivalently, halves the displacement increment between adjacent contours, so that while the displacement increment between the fringes of equation (2.2.1) is p_0 , the displacement increment between the fringes of equation (2.2.3) is $w_0/2$. The angle ϕ between the fringes and the lines of the grating in its initial state is described by an equation identical to equation (2.2.2):

$$\tan \phi = \frac{w_0 \sin \gamma}{w_0 \cos \gamma - w} \quad (2.2.4)$$

The displacement increment between adjacent fringes is $w_0/2$, and the direction of the displacement measured is perpendicular to the ridges of the diffraction grating in its initial state.

The knowledge of the initial state of the grating is retained in the laser beams which are diffracted to produce the interference fringes observed in moire interferometry. These laser beams are aligned to the initial grating state and held constant while the grating deforms.

It has been suggested that a direct analogy can be drawn between fringes of geometrical moire and the fringes in moire interferometry. [9,19] The incoming laser beams are aligned to the initial grating state in such a way that the interference between the beams produces a standing wave with a spacing of $w_0/2$ (where w_0 is the initial grating spacing). This interference pattern can be thought of as a virtual reference grating, and the fringes seen using the moire interferometer can be thought of as due to moire patterns between the current grating at spacing w and at an angle of γ to the virtual reference grating. The fringe multiplication factor arises from the fact that the reference grating is spaced at exactly half of the initial

grating spacing, and therefore is spaced at approximately half the deformed grating spacing (at least for small deformations), as described in by Livnat and Post [9] and by the author [21].

Although the analogy between geometric moire and moire interferometry is apt when the deformation remains strictly in the plane of the original grating, rotations of the plane will introduce angular changes in the diffracted laser beams which are not related to the in-plane displacement of the grating. This will introduce changes in the equations governing the spacing and orientation of the fringes. The equations describing the fringe spacing and orientation will be derived allowing for arbitrary rotations of the grating in an effort to determine the magnitude of the error introduced by assuming that the fringes represent contours of purely in-plane displacement.

2.2.3 Assumptions Underlying the Governing Equations

Moire interferometric fringes are the result of the interference between two sets of light waves which travel from the surface of the specimen to the observer at slightly different angles. The two light beams originate at the same laser source and follow paths of equal length to the specimen, but they arrive at the specimen from different directions. The surface of the specimen is a reflective diffraction grating, also called a phase grating, since it diffracts the light by altering its phase and not its amplitude as a transparent grating with inked lines would.

Initially, the incoming laser beams are arranged to intersect the phase grating at equal but opposite angles relative to the surface normal in such a way that the first diffraction order beams exit along the normal to the specimen surface to the

observer. Assuming a perfect grating and perfectly planar coherent light waves, this initial configuration produces a null interference pattern, since the phase difference of the diffracted light waves is constant over the field of view. If the grating is now stretched or shrunk to have a new wavelength, but the plane of the diffraction does not change, the laser beams will be diffracted away from the specimen normal by equal but opposite amounts (due to the symmetry of the incoming laser beams with respect to the surface normal). Since the exiting beams now intersect at an angle, the phase difference between the light waves will vary across the field of view, and the interference will be constructive in some places, destructive in others. This interference pattern can be used to infer the new wavelength of the grating and thus the in-plane deformation at the specimen surface.

If the plane of the grating does change, either through a rigid rotation or through out-of-plane deformation gradients, the angles of the diffracted beams will have an additional deflection which does not depend on the in-plane deformation.

The equations governing the formation of the fringe patterns observed in moire interferometry are derived under the assumption that the interference occurs between two coherent plane light waves of equal amplitude and similar polarization, which means that each plane wave can be described by a unit vector (or ray) normal to the plane wave (as in geometric optics). The light waves intersect at an angle and interfere to form a volume of parallel light and dark bands which are stationary in space. These bands (fringes) are perpendicular to the vector difference of the unit normal light rays. The spacing of the fringes depends on the wavelength of the light and the angle of intersection since the relative phase information of the two waves can be determined from the wavelength and the angle of intersection as long as plane wavefronts are assumed. Since the fringes observed in moire interferometry

are analogous to the fringes observed in generalized moire methods or the fringes generated in a Twyman-Green interferometer in that they represent contours of constant displacement, the absolute phase information of the waves is not important, since a shift in the phase of one or the other (by a shift in path length for example) will only shift the amplitude of the fringe without altering the spacing or shape of the contours. In other words, whether the fringe is light or dark is not important as long as the fringe spacing and inclination remains the same. This statement emphasizes one of the important facts about interferometry in general, namely that the displacements measured at a point are only known in relation to the displacements at other points in the field of view. In order to make absolute measurements, the absolute displacement of one point in the field must be known from some other method.

This derivation of the equations describing the fringes in moire interferometry is similar to that by Livnat and Post [9] in approach, although they restricted their derivation to planar deformation, and the approach was arrived at independently. The work by Livnat and Post offers some interesting comparisons between moire, interferometry and geometric moire methods. The equations for the case of planar deformation and the correspondence between moire interferometry and geometric moire were also derived in a slightly different way by the author [21] and in a very different manner by Pirodda [16]. The paper by Pirodda contains interesting descriptions of several interferometric measurement techniques.

2.3 The Governing Equations of Moire Interferometry Allowing for Arbitrary Changes in the Orientation of the Grating

2.3.1 Diffraction Equations

One type of diffraction grating which acts in reflection is called a phase grating because it alters the phase of a light wave periodically across the width of the light wave. This is in contrast to an amplitude grating, which alters the amplitude of a light wave periodically across its width. A phase grating is simply a periodic height variation, arranged in peaks and valleys. Ordinarily, the height variation is much smaller than the distance between peaks, while this spacing between peaks, the grating wavelength, is on the order of the wavelength of light.

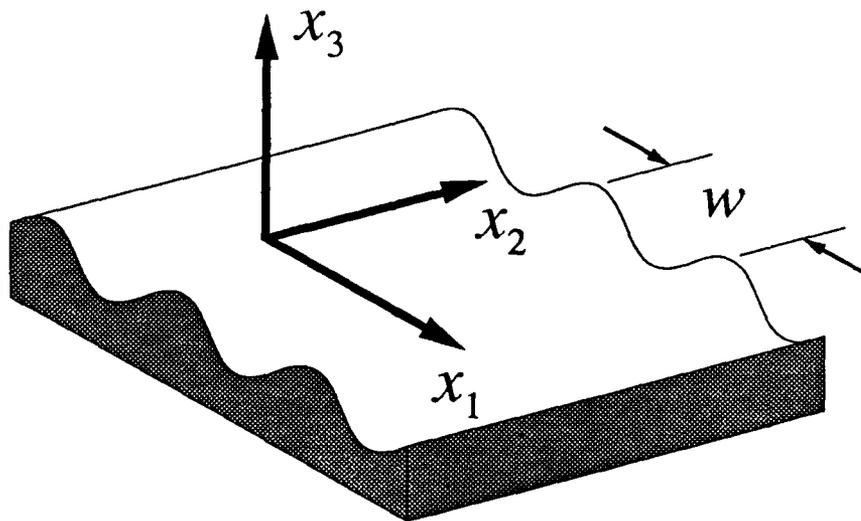


FIGURE 1. The Grating Coordinate Frame. The x_2 coordinate axis remains parallel to the ridges of the grating during deformation. The x_3 axis remains normal to the mean grating surface. The current grating wavelength is w .

Consider such a phase grating of wavelength w . Attach a right-handed orthogonal coordinate system to the grating so that the x_1 - x_2 plane coincides with the mean grating surface such that the x_2 axis is parallel to the ridges of the grating. This coordinate frame moves with the grating, so the x_2 axis always remains parallel to the current ridges of the grating and the x_3 axis always remains perpendicular to the mean grating surface. See Figure 1.

To calculate the direction of a diffracted light ray (of any order) from an incident light ray which is directed at the grating from an arbitrary angle, the incident light ray must be resolved into components with respect to the coordinate frame attached to the grating as describe above and in Figure 1. For instance, if unit vector \mathbf{a} represents the incident ray and unit vector $\mathbf{b}^{(n)}$ represents the n^{th} order ray diffracted from the incident ray, the equations relating the components of \mathbf{a} to the components of $\mathbf{b}^{(n)}$ are:

$$a_1 + b_1^{(n)} = \frac{n\lambda}{w} \quad (2.3.1)$$

$$a_2 + b_2^{(n)} = 0 \quad (2.3.2),$$

where a_α and $b_\alpha^{(n)}$ are the components of \mathbf{a} and $\mathbf{b}^{(n)}$ in the grating coordinate frame and λ is the wavelength of light used. Note that while the diffracted light rays are directed along $\mathbf{b}^{(n)}$, the incident light rays are actually directed along $-\mathbf{a}$. This means that both a_3 and $b_3^{(n)}$ are nonnegative, and since \mathbf{a} and $\mathbf{b}^{(n)}$ are unit vectors, equations (2.3.1) and (2.3.2) uniquely determine $\mathbf{b}^{(n)}$ from any given \mathbf{a} . Equations (2.3.1) and (2.3.2) are given in this form to correspond to the form usually presented in the literature concerning diffraction gratings following Guild [3,4].

2.3.2 Interference of Plane Light Waves

Let the unit vectors \mathbf{u} and \mathbf{v} represent the normals of plane light waves which interfere to form the fringes observed in moire interferometry. These plane light waves interfere so as to create a standing wave pattern of parallel light and dark fringes as shown in Figure 2. The fringes are surfaces along which the phase difference between the two sets of light waves is constant. The planes of constant phase difference are normal to the difference vector $\mathbf{z} = \mathbf{u} - \mathbf{v}$ and are spaced a distance δ from one dark band to the next dark band. (Alternatively, one could say that the standing wave has a wavelength of δ .) Let 2β be the (acute) angle between \mathbf{u} and \mathbf{v} , then δ is given by:

$$\delta = \frac{\lambda}{2|\sin \beta|} \quad (2.3.3)$$

where λ is the wavelength of light used.

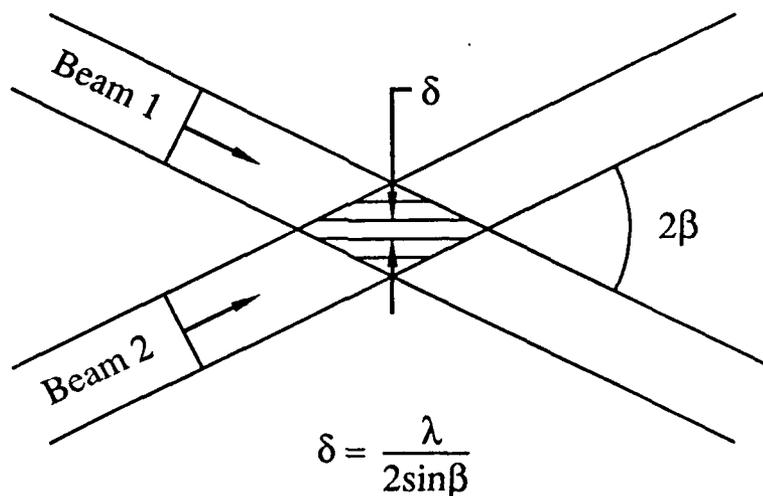


FIGURE 2. Standing interference pattern formed between collimated laser beams crossing at angle 2β . The light has a wavelength of λ . The interference pattern has a wavelength δ .

The spacing and orientation of the fringe bands can be specified by a single vector \mathbf{d} which contains all the information about the fringes. The vector \mathbf{d} is normal to the fringe planes and has a length equal to the distance between the fringe planes. Since δ is the distance between the fringe planes and the vector \mathbf{z} ($= \mathbf{u} - \mathbf{v}$) is normal to the fringe planes, the vector \mathbf{d} is determined by

$$\mathbf{d} = \delta \frac{\mathbf{z}}{|\mathbf{z}|} = \frac{\lambda}{2|\sin \beta|} \frac{\mathbf{z}}{|\mathbf{z}|} \quad (2.3.4)$$

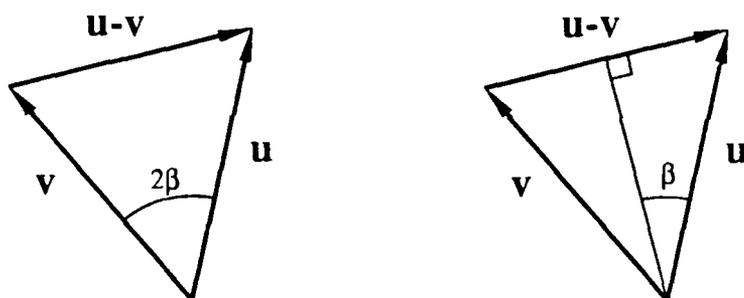


FIGURE 3. Unit vectors representing the normals to diffracted plane light waves. It is seen that $|2 \sin \beta| = |\mathbf{u} - \mathbf{v}|$. The spacing of the interference fringes is $\delta = \lambda / (|2 \sin \beta|)$.

From Figure 3, it is apparent that since \mathbf{u} and \mathbf{v} are unit vectors,

$$2|\sin \beta| = |\mathbf{u} - \mathbf{v}| = |\mathbf{z}| \quad (2.3.5)$$

Therefore,

$$\delta = \frac{\lambda}{|\mathbf{z}|} \quad (2.3.6)$$

and

$$\mathbf{d} = \frac{\lambda}{|\mathbf{z}|^2} \mathbf{z} \quad (2.3.7)$$

2.3.3 Geometry

In moire interferometry, it is necessary to calculate the directions of diffracted rays originating from two different incident rays. These incident rays are fixed in space while the grating is allowed to undergo an arbitrary change in orientation and line spacing.

Consider a coordinate frame which is fixed in space. Denote the fixed axes by X_1, X_2, X_3 . The grating coordinate system, x_1, x_2, x_3 as shown in Figure 1, is initially aligned with the fixed system. As the grating deforms, the lines of the grating rotate so that the grating coordinate frame no longer coincides with the fixed frame. Additionally, the plane of the grating is allowed to rotate with respect to the fixed axes. The change in the orientation of the grating can be described through the use of Euler angles. First let the grating rotate an angle γ_3 about the x_3 (or X_3) axis. Next let the grating rotate an angle γ_2 about the new position of the grating x_2 axis. Finally, let the grating rotate an angle γ_1 about the new position of the x_1 axis to its final position.

The relation between the components of any vector \mathbf{a} resolved in the $X_1X_2X_3$ coordinate frame and the components of \mathbf{a} resolved in the final position of the $x_1x_2x_3$ (grating) coordinate system are determined through the successive matrix operation:

$$\begin{pmatrix} a_1 \\ a_2 \\ a_3 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos \gamma_1 & \sin \gamma_1 \\ 0 & -\sin \gamma_1 & \cos \gamma_1 \end{pmatrix} \begin{pmatrix} \cos \gamma_2 & 0 & -\sin \gamma_2 \\ 0 & 1 & 0 \\ \sin \gamma_2 & 0 & \cos \gamma_2 \end{pmatrix} \begin{pmatrix} \cos \gamma_3 & \sin \gamma_3 & 0 \\ -\sin \gamma_3 & \cos \gamma_3 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} A_1 \\ A_2 \\ A_3 \end{pmatrix} \quad (2.3.8)$$

where a_1, a_2, a_3 are the components of \mathbf{a} in the $x_1x_2x_3$ frame and A_1, A_2, A_3 are the components of \mathbf{a} in the $X_1X_2X_3$ frame.

This transformation is necessary, because the diffraction equations are written in terms of components in the grating frame while the incoming laser beams are specified in components in the fixed frame.

2.3.4 Moire Interferometry

The light rays represented by \mathbf{u} and \mathbf{v} are diffracted from two fixed incident light rays. The two fixed incident rays will be represented by unit vectors \mathbf{j} and \mathbf{k} . Again, it is emphasized that these vectors are directed outward from the original grating surface, in the opposite direction from the incoming light waves. These two unit vectors have components J_1, J_2, J_3 and K_1, K_2, K_3 , respectively, in the globally fixed $X_1X_2X_3$ frame, and components j_1, j_2, j_3 and k_1, k_2, k_3 , respectively, in the grating fixed $x_1x_2x_3$ frame.

Let unit vector \mathbf{u} represent the diffracted ray of order 1 from the incident ray represented by \mathbf{j} , and let unit vector \mathbf{v} represent the diffracted ray of order -1 from the incident ray represented by \mathbf{k} . Let these vectors have components u_1, u_2, u_3 and v_1, v_2, v_3 , respectively, in the grating fixed $x_1x_2x_3$ frame. Then from equations (2.3.1) and (2.3.2),

$$u_1 = \frac{\lambda}{w} - j_1 \quad (2.3.9)$$

$$u_2 = -j_2 \quad (2.3.10)$$

$$v_1 = -\frac{\lambda}{w} - k_1 \quad (2.3.11)$$

$$v_2 = -k_2 \quad (2.3.12)$$

and since \mathbf{u} and \mathbf{v} are unit vectors,

$$u_3 = [1 - (u_1)^2 - (u_2)^2]^{\frac{1}{2}} \quad (2.3.13)$$

$$v_3 = [1 - (v_1)^2 - (v_2)^2]^{\frac{1}{2}} \quad (2.3.14)$$

In the experimental arrangement as used in this study, the two fixed incident rays represented by \mathbf{j} and \mathbf{k} lie in the X_1 - X_3 plane and are symmetric with respect to the X_2 - X_3 plane. In the global $X_1X_2X_3$ coordinate frame, these vectors have a simple representation:

$$\begin{Bmatrix} J_1 \\ J_2 \\ J_3 \end{Bmatrix} = \begin{Bmatrix} \sin \alpha \\ 0 \\ \cos \alpha \end{Bmatrix}, \quad \begin{Bmatrix} K_1 \\ K_2 \\ K_3 \end{Bmatrix} = \begin{Bmatrix} -\sin \alpha \\ 0 \\ \cos \alpha \end{Bmatrix}, \quad -\frac{\pi}{2} \leq \alpha \leq \frac{\pi}{2} \quad (2.3.15)$$

The components of \mathbf{j} and \mathbf{k} in the $x_1x_2x_3$ frame are then found using equation (2.3.8) as

$$\begin{Bmatrix} j_1 \\ j_2 \\ j_3 \end{Bmatrix} = \begin{Bmatrix} \sin \alpha \cos \gamma_3 \cos \gamma_2 - \cos \alpha \sin \gamma_2 \\ -\sin \alpha \sin \gamma_3 \cos \gamma_1 + \sin \alpha \cos \gamma_3 \sin \gamma_2 \sin \gamma_1 + \cos \alpha \cos \gamma_2 \sin \gamma_1 \\ \sin \alpha \sin \gamma_3 \sin \gamma_1 + \sin \alpha \cos \gamma_3 \sin \gamma_2 \cos \gamma_1 + \cos \alpha \cos \gamma_2 \cos \gamma_1 \end{Bmatrix} \quad (2.3.16)$$

$$\begin{Bmatrix} k_1 \\ k_2 \\ k_3 \end{Bmatrix} = \begin{Bmatrix} -\sin \alpha \cos \gamma_3 \cos \gamma_2 - \cos \alpha \sin \gamma_2 \\ \sin \alpha \sin \gamma_3 \cos \gamma_1 - \sin \alpha \cos \gamma_3 \sin \gamma_2 \sin \gamma_1 + \cos \alpha \cos \gamma_2 \sin \gamma_1 \\ -\sin \alpha \sin \gamma_3 \sin \gamma_1 - \sin \alpha \cos \gamma_3 \sin \gamma_2 \cos \gamma_1 + \cos \alpha \cos \gamma_2 \cos \gamma_1 \end{Bmatrix} \quad (2.3.17)$$

By using equations (2.3.9) through (2.3.14), the components of the unit vectors \mathbf{u} (which represents the diffracted ray of order 1 from the incident ray represented by \mathbf{j}) and \mathbf{v} (which represents the diffracted ray of order -1 from the incident ray represented by \mathbf{k}) in the grating coordinate frame can be found as

$$u_1 = \frac{\lambda}{w} - \sin \alpha \cos \gamma_3 \cos \gamma_2 + \cos \alpha \sin \gamma_2 \quad (2.3.18)$$

$$u_2 = \sin \alpha \sin \gamma_3 \cos \gamma_1 - \sin \alpha \cos \gamma_3 \sin \gamma_2 \sin \gamma_1 - \cos \alpha \cos \gamma_2 \sin \gamma_1 \quad (2.3.19)$$

$$u_3 = [1 - (u_1)^2 - (u_2)^2]^{\frac{1}{2}} \quad (2.3.20)$$

$$v_1 = -\frac{\lambda}{w} + \sin \alpha \cos \gamma_3 \cos \gamma_2 + \cos \alpha \sin \gamma_2 \quad (2.3.21)$$

$$v_2 = -\sin \alpha \sin \gamma_3 \cos \gamma_1 + \sin \alpha \cos \gamma_3 \sin \gamma_2 \sin \gamma_1 - \cos \alpha \cos \gamma_2 \sin \gamma_1 \quad (2.3.22)$$

$$v_3 = [1 - (v_1)^2 - (v_2)^2]^{\frac{1}{2}} \quad (2.3.23)$$

The components of the difference vector $\mathbf{z} = \mathbf{u} - \mathbf{v}$ in the grating coordinate frame are then

$$z_1 = 2 \left(\frac{\lambda}{w} - \sin \alpha \cos \gamma_3 \cos \gamma_2 \right) \quad (2.3.24)$$

$$z_2 = 2(\sin \alpha \sin \gamma_3 \cos \gamma_1 - \sin \alpha \cos \gamma_3 \sin \gamma_2 \sin \gamma_1) \quad (2.3.25)$$

$$z_3 = [1 - (u_1)^2 - (u_2)^2]^{\frac{1}{2}} - [1 - (v_1)^2 - (v_2)^2]^{\frac{1}{2}} \quad (2.3.26)$$

Because the fringe planes are perpendicular to the vector \mathbf{z} , their spacing δ is found from equation (2.3.6) as

$$\delta = \frac{\lambda}{|\mathbf{z}|} \quad (2.3.27)$$

so that the orientation and spacing of the fringes is described by the normal vector \mathbf{d} as in equation (2.3.7), by

$$\mathbf{d} = \frac{\lambda}{|\mathbf{z}|^2} \mathbf{z} \quad (2.3.28)$$

2.3.5 Initial Alignment

Before any deformation occurs, i.e., while the grating axes and the global axes coincide, the angle α is adjusted so that there are no fringes visible in the field of view, which means the fringe spacing δ is infinite. This angle α remains constant throughout the deformation. The angle α depends only on the initial wavelength of the grating, and, in effect, α acts to retain the initial grating wavelength to compare it with the current deformed grating wavelength.

Recall that the initial grating wavelength is w_0 ; this is known in advance. Also, γ_1, γ_2 and γ_3 are zero initially. The two incident laser beams are arranged so that the diffracted beams lie along the x_3 axis, meaning $u_1 = u_2 = 0$ and $u_3 = v_3 = 1$.

It follows from equations (2.3.18) through (2.3.23), making the indicated substitutions

$$u_1 = 0 = \frac{\lambda}{w_0} - \sin \alpha \quad (2.3.29)$$

$$u_2 = 0 \quad (2.3.30)$$

$$u_3 = 1 \quad (2.3.31)$$

$$v_1 = 0 = -\frac{\lambda}{w_0} + \sin \alpha \quad (2.3.32)$$

$$v_2 = 0 \quad (2.3.33)$$

$$v_3 = 1 \quad (2.3.34)$$

which requires that

$$\sin \alpha = \frac{\lambda}{w_0} \quad (2.3.35)$$

This relation does not change during the deformation.

There is no similar simple representation for $\cos \alpha$, so the terms $\sin \alpha$ and $\cos \alpha$ will be retained unless the substitution helps to simplify an equation. It should be understood, however, that α is a known quantity which is related to the undeformed grating wavelength w_0 and to the wavelength of light used λ .

2.3.6 Observed Fringes

In general, the fringe surfaces described by the vector \mathbf{d} are not directly observed, because the fringe planes are not aligned with the optical axis of the recording device. The optical system used to record the fringes is aligned along the fixed X_3 axis and focuses on a plane parallel to the X_1 - X_2 plane; it will thus record the fringe planes where they intersect a plane parallel to the X_1 - X_2 plane. Thus the optical system will not in general record the true fringe wavelength δ , but instead an apparent fringe wavelength δ^* . This apparent fringe wavelength corresponds to the spacing between the intersections of the fringe planes with planes parallel to the X_1 - X_2 plane.

Let $\mathbf{E}_1, \mathbf{E}_2$ and \mathbf{E}_3 be unit vectors aligned with the globally fixed X_1, X_2 and X_3 axes, respectively. The observed wavelength and inclination of fringes can be represented by a vector \mathbf{d}^* which lies in the X_1 - X_2 plane and is perpendicular to the unit vector \mathbf{E}_3 , as shown in Figure 4. Knowing the vector \mathbf{d} which contains all the information about the fringes, the vector \mathbf{d}^* can be constructed as the apparent fringe spacing δ^* times a unit vector in the X_1 - X_2 plane in the proper direction.

The vector \mathbf{d}^* will always have a length that is greater or equal to the length of \mathbf{d} , since \mathbf{d} is the shortest vector between two fringe planes. (The vector \mathbf{d} is a projection of the vector \mathbf{d}^* .)

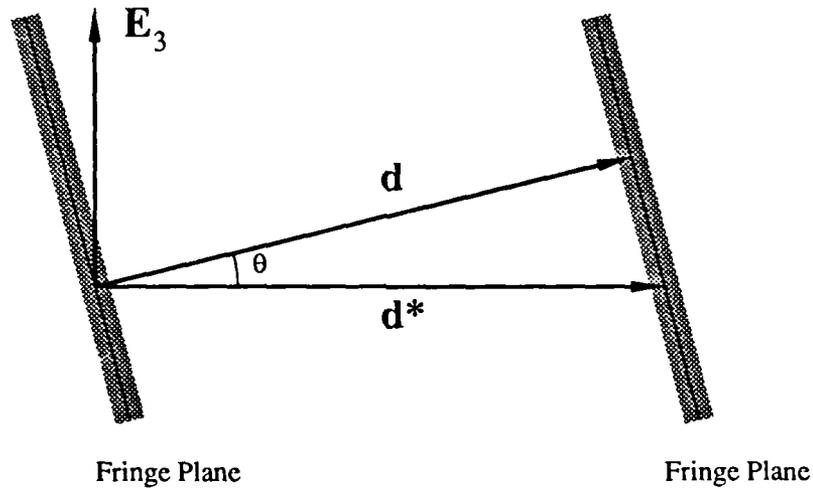


FIGURE 4. The relation between the apparent fringe vector \mathbf{d}^* and the true fringe vector \mathbf{d} . The vector \mathbf{d}^* is perpendicular to the globally fixed direction along \mathbf{E}_3 .

Let θ be the angle between the vectors \mathbf{d} and \mathbf{d}^* .

$$\delta^* |\cos \theta| = \delta \tag{2.3.36}$$

$$\mathbf{d}^* = \delta^* \frac{(\mathbf{d} - (\mathbf{d} \cdot \mathbf{E}_3)\mathbf{E}_3)}{|\mathbf{d} - (\mathbf{d} \cdot \mathbf{E}_3)\mathbf{E}_3|} \tag{2.3.37}$$

The denominator of equation (2.3.37) can be expanded to find

$$\begin{aligned}
 |\mathbf{d} - (\mathbf{d} \cdot \mathbf{E}_3)\mathbf{E}_3| &= [(\mathbf{d} - (\mathbf{d} \cdot \mathbf{E}_3)\mathbf{E}_3) \cdot (\mathbf{d} - (\mathbf{d} \cdot \mathbf{E}_3)\mathbf{E}_3)]^{\frac{1}{2}} \\
 &= [\mathbf{d} \cdot \mathbf{d} - 2(\mathbf{d} \cdot \mathbf{E}_3)^2 + (\mathbf{d} \cdot \mathbf{E}_3)^2]^{\frac{1}{2}} \\
 &= [\mathbf{d} \cdot \mathbf{d} - (\mathbf{d} \cdot \mathbf{E}_3)^2]^{\frac{1}{2}} \\
 &= \frac{\lambda}{|\mathbf{z}|^2} [\mathbf{z} \cdot \mathbf{z} - (\mathbf{z} \cdot \mathbf{E}_3)^2]^{\frac{1}{2}} \tag{2.3.38}
 \end{aligned}$$

where $\mathbf{z} (= \mathbf{u} - \mathbf{v})$ is the difference vector between the unit vectors representing the interfering plane waves. To find δ^* ,

$$\begin{aligned}
 \mathbf{d}^* \cdot \mathbf{d} &= \delta^* \frac{(\mathbf{d} - (\mathbf{d} \cdot \mathbf{E}_3)\mathbf{E}_3)}{[\mathbf{d} \cdot \mathbf{d} - (\mathbf{d} \cdot \mathbf{E}_3)^2]^{\frac{1}{2}}} \cdot \mathbf{d} \\
 &= \delta^* \frac{(\mathbf{d} \cdot \mathbf{d} - (\mathbf{d} \cdot \mathbf{E}_3)(\mathbf{d} \cdot \mathbf{E}_3))}{[\mathbf{d} \cdot \mathbf{d} - (\mathbf{d} \cdot \mathbf{E}_3)^2]^{\frac{1}{2}}} \\
 &= \delta^* [\mathbf{d} \cdot \mathbf{d} - (\mathbf{d} \cdot \mathbf{E}_3)^2]^{\frac{1}{2}} \\
 &= \delta^* \frac{\lambda}{|\mathbf{z}|^2} [\mathbf{z} \cdot \mathbf{z} - (\mathbf{z} \cdot \mathbf{E}_3)^2]^{\frac{1}{2}} \tag{2.3.39}
 \end{aligned}$$

but

$$\mathbf{d}^* \cdot \mathbf{d} = \delta\delta^* |\cos \theta| = \delta^2 = \mathbf{d} \cdot \mathbf{d} = \left(\frac{\lambda}{|\mathbf{z}|} \right)^2 \tag{2.3.40}$$

Therefore,

$$\begin{aligned}
 \delta^* &= \frac{\mathbf{d} \cdot \mathbf{d}}{[\mathbf{d} \cdot \mathbf{d} - (\mathbf{d} \cdot \mathbf{E}_3)^2]^{\frac{1}{2}}} \\
 &= \frac{\lambda}{[\mathbf{z} \cdot \mathbf{z} - (\mathbf{z} \cdot \mathbf{E}_3)^2]^{\frac{1}{2}}} \tag{2.3.41}
 \end{aligned}$$

and so,

$$\begin{aligned}
 \mathbf{d}^* &= \left[\frac{\mathbf{d} \cdot \mathbf{d}}{\mathbf{d} \cdot \mathbf{d} - (\mathbf{d} \cdot \mathbf{E}_3)^2} \right] (\mathbf{d} - (\mathbf{d} \cdot \mathbf{E}_3)\mathbf{E}_3) \\
 &= \left[\frac{\lambda}{\mathbf{z} \cdot \mathbf{z} - (\mathbf{z} \cdot \mathbf{E}_3)^2} \right] (\mathbf{z} - (\mathbf{z} \cdot \mathbf{E}_3)\mathbf{E}_3) \tag{2.3.42}
 \end{aligned}$$

Let ϕ^* be the angle between the observed fringes and the X_2 axis. ϕ^* is also the angle between the vector \mathbf{d}^* and the X_1 axis, so that

$$\tan \phi^* = \frac{\mathbf{d}^* \cdot \mathbf{E}_2}{\mathbf{d}^* \cdot \mathbf{E}_1} = \frac{\mathbf{d} \cdot \mathbf{E}_2}{\mathbf{d} \cdot \mathbf{E}_1} = \frac{\mathbf{z} \cdot \mathbf{E}_2}{\mathbf{z} \cdot \mathbf{E}_1} \quad (2.3.43)$$

To emphasize the equations describing the fringe spacing and inclination, they are written together here again. The observed or apparent fringe spacing δ^* and inclination angle ϕ^* are found to be

$$\delta^* = \frac{\lambda}{[\mathbf{z} \cdot \mathbf{z} - (\mathbf{z} \cdot \mathbf{E}_3)^2]^{\frac{1}{2}}} \quad (2.3.41)$$

$$\tan \phi^* = \frac{\mathbf{z} \cdot \mathbf{E}_2}{\mathbf{z} \cdot \mathbf{E}_1} \quad (2.3.43)$$

Again, \mathbf{z} is the difference vector between the vectors representing the diffracted rays, $\mathbf{z} = \mathbf{u} - \mathbf{v}$, and \mathbf{E}_1 , \mathbf{E}_2 and \mathbf{E}_3 are unit vectors aligned with the globally fixed coordinate frame.

The components of \mathbf{z} are known in the grating coordinate frame. In order to use equations (2.3.41) and (2.3.43), one must find the components of \mathbf{z} in the globally fixed frame, or find the components of the vectors \mathbf{E}_1 , \mathbf{E}_2 and \mathbf{E}_3 in the grating system. The latter course will be followed here. Let \mathbf{e}_1 , \mathbf{e}_2 and \mathbf{e}_3 be unit vectors aligned with the grating x_1 , x_2 and x_3 axes, respectively. \mathbf{E}_1 , \mathbf{E}_2 and \mathbf{E}_3 are the unit vectors aligned with the globally fixed X_1 , X_2 and X_3 axes. Then, again using equation (2.3.8),

$$\begin{aligned} \mathbf{E}_1 = & (\cos \gamma_2 \cos \gamma_3) \mathbf{e}_1 + (\sin \gamma_1 \sin \gamma_2 \cos \gamma_3 - \cos \gamma_1 \sin \gamma_3) \mathbf{e}_2 \\ & + (\cos \gamma_1 \sin \gamma_2 \cos \gamma_3 + \sin \gamma_1 \sin \gamma_3) \mathbf{e}_3 \end{aligned} \quad (2.3.44)$$

$$\begin{aligned} \mathbf{E}_2 = & (\cos \gamma_2 \sin \gamma_3) \mathbf{e}_1 + (\sin \gamma_1 \sin \gamma_2 \sin \gamma_3 + \cos \gamma_1 \cos \gamma_3) \mathbf{e}_2 \\ & + (\cos \gamma_1 \sin \gamma_2 \sin \gamma_3 - \sin \gamma_1 \cos \gamma_3) \mathbf{e}_3 \end{aligned} \quad (2.3.45)$$

$$\mathbf{E}_3 = (-\sin \gamma_2) \mathbf{e}_1 + (\sin \gamma_1 \cos \gamma_2) \mathbf{e}_2 + (\cos \gamma_1 \cos \gamma_2) \mathbf{e}_3 \quad (2.3.46)$$

2.3.7 Special Case: Planar Deformation

Examine first the case of deformation in which the mean grating surface remains parallel to the globally fixed X_1 - X_2 plane. This is the ideal case in which the moire interferometric fringes exactly correspond to the fringes of geometric moire methods as contours of constant in-plane displacement. In this case, $\gamma_1 \equiv 0$ and $\gamma_2 \equiv 0$ and the components of \mathbf{u} and \mathbf{v} are given by:

$$u_1 = \frac{\lambda}{w} - \sin \alpha \cos \gamma_3 \quad (2.3.47)$$

$$u_2 = \sin \alpha \sin \gamma_3 \quad (2.3.48)$$

$$u_3 = [1 - (u_1)^2 - (u_2)^2]^{\frac{1}{2}} \quad (2.3.49)$$

$$v_1 = -\frac{\lambda}{w} + \sin \alpha \cos \gamma_3 \quad (2.3.50)$$

$$v_2 = -\sin \alpha \sin \gamma_3 \quad (2.3.51)$$

$$v_3 = [1 - (v_1)^2 - (v_2)^2]^{\frac{1}{2}} \quad (2.3.52)$$

Notice that $u_1 = -v_1$ and $u_2 = -v_2$, which implies that $u_3 = v_3$. The vector \mathbf{z} ($= \mathbf{u} - \mathbf{v}$) is found to have components in the grating coordinate frame given by

$$z_1 = 2 \left(\frac{\lambda}{w} - \sin \alpha \cos \gamma_3 \right) \quad (2.3.53)$$

$$z_2 = 2(\sin \alpha \sin \gamma_3) \quad (2.3.54)$$

$$z_3 = 0 \quad (2.3.55)$$

In vector form,

$$\mathbf{z} = 2 \left(\frac{\lambda}{w} - \sin \alpha \cos \gamma_3 \right) \mathbf{e}_1 + 2(\sin \alpha \sin \gamma_3) \mathbf{e}_2 \quad (2.3.56)$$

$$\begin{aligned}
 |\mathbf{z}|^2 &= 4 \left[\left(\frac{\lambda}{w} \right)^2 + \sin^2 \alpha \cos^2 \gamma_3 + \sin^2 \alpha \sin^2 \gamma_3 - 2 \frac{\lambda}{w} \sin \alpha \cos \gamma_3 \right] \\
 &= 4 \left[\left(\frac{\lambda}{w} \right)^2 + \sin^2 \alpha - 2 \frac{\lambda}{w} \sin \alpha \cos \gamma_3 \right]
 \end{aligned} \tag{2.3.57}$$

The true fringe spacing δ is

$$\delta = \frac{\lambda}{2 \left[\left(\frac{\lambda}{w} \right)^2 + \sin^2 \alpha - 2 \frac{\lambda}{w} \sin \alpha \cos \gamma_3 \right]^{\frac{1}{2}}} \tag{2.3.58}$$

The vector $\mathbf{d} = (\mathbf{z}/|\mathbf{z}|)\delta$ is

$$\mathbf{d} = \frac{\lambda \left[\left(\frac{\lambda}{w} - \sin \alpha \cos \gamma_3 \right) \mathbf{e}_1 + (\sin \alpha \sin \gamma_3) \mathbf{e}_2 \right]}{2 \left[\left(\frac{\lambda}{w} \right)^2 + \sin^2 \alpha - 2 \frac{\lambda}{w} \sin \alpha \cos \gamma_3 \right]} \tag{2.3.59}$$

Equations (2.3.44) through (2.3.46) reduce to

$$\mathbf{E}_1 = \cos \gamma_3 \mathbf{e}_1 - \sin \gamma_3 \mathbf{e}_2 \tag{2.3.60}$$

$$\mathbf{E}_2 = \sin \gamma_3 \mathbf{e}_1 + \cos \gamma_3 \mathbf{e}_2 \tag{2.3.61}$$

$$\mathbf{E}_3 = \mathbf{e}_3 \tag{2.3.62}$$

Therefore, $\mathbf{d} \cdot \mathbf{E}_3 = 0$, and so the observed fringe spacing δ^* is the same as the true fringe spacing δ in equation (2.3.58),

$$\delta^* = \frac{\lambda}{2 \left[\left(\frac{\lambda}{w} \right)^2 + \sin^2 \alpha - 2 \frac{\lambda}{w} \sin \alpha \cos \gamma_3 \right]^{\frac{1}{2}}} \tag{2.3.63}$$

The inclination of the fringe planes with respect to the X_2 axis is

$$\tan \phi^* = \frac{\left(\frac{\lambda}{w} \right) \sin \gamma_3}{\left(\frac{\lambda}{w} \right) \cos \gamma_3 - \sin \alpha} \tag{2.3.64}$$

Substituting $\sin \alpha = \lambda/w_0$ from equation (2.3.35), one finds for the fringe spacing and inclination

$$\delta^* = \frac{ww_0}{2[w^2 + w_0^2 - 2ww_0 \cos \gamma_3]^{\frac{1}{2}}} \quad (2.3.65)$$

$$\tan \phi^* = \frac{w_0 \sin \gamma_3}{w_0 \cos \gamma_3 - w} \quad (2.3.66)$$

Equations (2.3.65) and (2.3.66) are important because they imply that in the case of planar deformation, the fringes can be interpreted as contours of constant displacement, as shown by the author [21], as well as by Livnat and Post [9] and Pirodda [16]. Note that the fringe spacing and inclination depend only on the initial and current states of the grating and are independent of the wavelength of light used. The constant displacement increment between adjacent fringes is $w_0/2$ as shown by the author [21]. Figures 5 and 6 show the variation of the fringe spacing and inclination for the case where there is no out-of-plane rotation. The choice of $w_0 = 1.667 \mu\text{m}$ matches the grating wavelength used in the experiments.

2.3.8 Approximations for General Deformations

The case of planar deformation provides a starting point for approximating the equations describing the fringes in cases where the deformation does not remain strictly in a plane parallel to the original mean grating surface. These approximations will help show the dominant sources of error caused by the out-of-plane rotations γ_1 and γ_2 . In effect, it is desirable to be able to interpret the fringes as contour lines of displacement, rather than to back out the underlying parameters w and γ_3 . The identity of equations (2.3.65) and (2.3.66) with the the equations for the

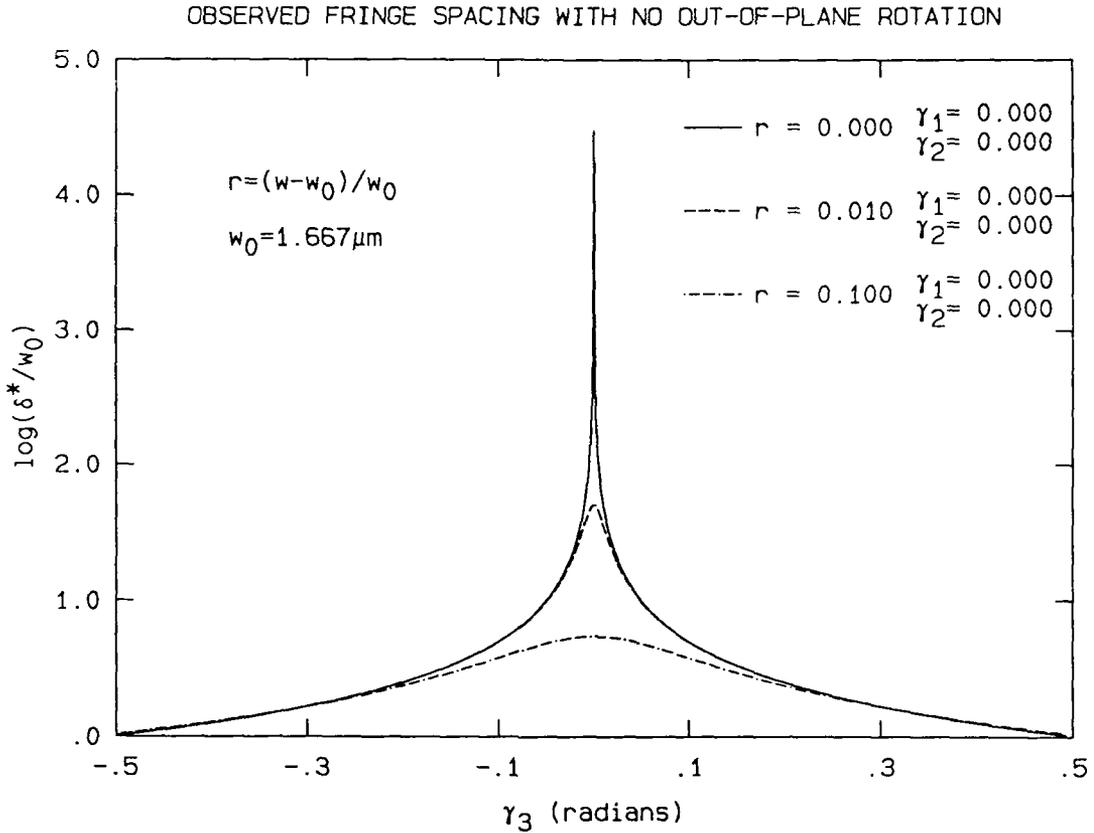


FIGURE 5. The variation of the observed fringe spacing δ^* when the out-of-plane rotations are identically zero. The fringe spacing is always maximum, at $\gamma_3 = 0$. Negative stretch ratios have similar effect on the fringe spacing. $w_0 = 1.667 \mu\text{m}$ corresponds to the grating wavelength used in the experiments.

case of geometric moire means that the fringes do represent contours of displacement when $\gamma_1 = \gamma_2 = 0$. If the effects of these out-of-plane rotations on equations (2.3.43) and (2.3.41) can be neglected, then the fringes can still be interpreted as contours of displacement. In the initial, undeformed configuration both of the diffracted rays are parallel to the global X_3 axis, and which corresponds to the grating x_3 axis, so $u_3 \equiv v_3 \equiv 1$. Assume that as the grating deforms, these components do not differ from unity by very much. (This assumption can be justified on the basis of the

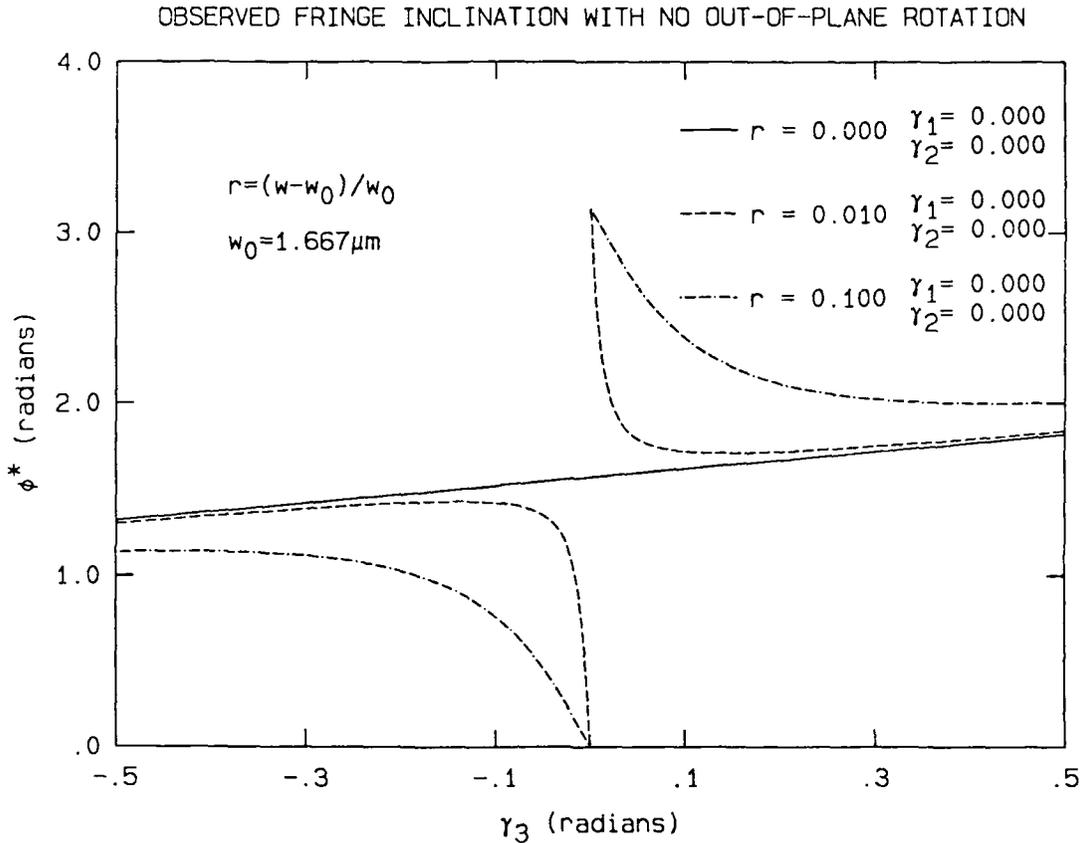


FIGURE 6. The variation of the observed fringe inclination ϕ^* when the out-of-plane rotations are identically zero. Note that the fringe inclination is periodic in π , the abrupt jumps at $\gamma_3 = 0$ are really smooth transitions across branches of the arctangent function. Negative stretch ratios have similar effect on the fringe inclination, but the curves will be in the opposite direction. $w_0 = 1.667 \mu\text{m}$ corresponds to the grating wavelength used in the experiments.

physical limitations of the optics used to record the fringes: large changes in these components that would violate this assumption lead to fringe densities that cannot be recorded.) Using this assumption, u_3 and v_3 can be approximated by two term Taylor expansions about unity:

$$u_3 = 1 - \frac{1}{2}(u_1^2 + u_2^2) \tag{2.3.67}$$

$$v_3 = 1 - \frac{1}{2}(v_1^2 + v_2^2) \quad (2.3.68)$$

Therefore,

$$z_3 = u_3 - v_3 = \frac{1}{2}(v_1^2 + v_2^2 - u_1^2 - u_2^2) \quad (2.3.69)$$

Using equations (2.3.18) through (2.3.23), which spell out the components u_1 , u_2 , v_1 and v_2 ,

$$z_3 = 2 \left[\left(\frac{\lambda}{w} - \sin \alpha \cos \gamma_3 \cos \gamma_2 \right) (-\sin \gamma_2 \cos \alpha) + (\sin \alpha \sin \gamma_3 \cos \gamma_1 - \sin \alpha \cos \gamma_3 \sin \gamma_2 \sin \gamma_1)(\cos \alpha \cos \gamma_2 \sin \gamma_1) \right] \quad (2.3.70)$$

From equations (2.3.24) and (2.3.25) it is seen that

$$\begin{aligned} z_3 &= (-\sin \gamma_2 \cos \alpha)(u_1 - v_1) + (\sin \gamma_1 \cos \gamma_2 \cos \alpha)(u_2 - v_2) \\ &= (-\sin \gamma_2 \cos \alpha)(z_1) + (\sin \gamma_1 \cos \gamma_2 \cos \alpha)(z_2) \end{aligned} \quad (2.3.71)$$

For this approximation,

$$z_1 = 2 \left(\frac{\lambda}{w} - \sin \alpha \cos \gamma_3 \cos \gamma_2 \right) \quad (2.3.72)$$

$$z_2 = 2(\sin \alpha \sin \gamma_3 \cos \gamma_1 - \sin \alpha \cos \gamma_3 \sin \gamma_2 \sin \gamma_1) \quad (2.3.73)$$

$$z_3 = z_1(-\sin \gamma_2 \cos \alpha) + z_2(\sin \gamma_1 \cos \gamma_2 \cos \alpha) \quad (2.3.74)$$

Using the relations for \mathbf{E}_1 , \mathbf{E}_2 and \mathbf{E}_3 in terms of the grating coordinate frame,

$$\begin{aligned} \mathbf{z} \cdot \mathbf{E}_3 &= z_1(-\sin \gamma_2) + z_2(\sin \gamma_1 \cos \gamma_2) \\ &\quad + z_1(\cos \gamma_1 \cos \gamma_2)(-\sin \gamma_2 \cos \alpha) \\ &\quad + z_2(\cos \gamma_1 \cos \gamma_2)(\cos \alpha \cos \gamma_2 \sin \gamma_1) \\ &= z_1(1 + \cos \gamma_1 \cos \gamma_2 \cos \alpha)(-\sin \gamma_2) \\ &\quad + z_2(1 + \cos \gamma_1 \cos \gamma_2 \cos \alpha)(\sin \gamma_1 \cos \gamma_2) \end{aligned} \quad (2.3.75)$$

$$\begin{aligned}
 \mathbf{z} \cdot \mathbf{z} &= z_1^2 + z_2^2 + z_1^2(\cos^2 \alpha \sin^2 \gamma_2) \\
 &\quad - 2z_1 z_2(\cos \alpha \sin \gamma_2)(\cos \alpha \cos \gamma_2 \sin \gamma_1) \\
 &\quad + z_2^2(\cos^2 \alpha \cos^2 \gamma_2 \sin^2 \gamma_1)
 \end{aligned} \tag{2.3.76}$$

$$\begin{aligned}
 (\mathbf{z} \cdot \mathbf{E}_3)^2 &= z_1^2(\sin^2 \gamma_2)(1 + \cos \gamma_1 \cos \gamma_2 \cos \alpha)^2 \\
 &\quad - 2z_1 z_2(\sin \gamma_2 \sin \gamma_1 \cos \gamma_2)(1 + \cos \gamma_1 \cos \gamma_2 \cos \alpha)^2 \\
 &\quad + z_2^2(\sin^2 \gamma_1 \cos^2 \gamma_2)(1 + \cos \gamma_1 \cos \gamma_2 \cos \alpha)^2
 \end{aligned} \tag{2.3.77}$$

$$\begin{aligned}
 \mathbf{z} \cdot \mathbf{z} - (\mathbf{z} \cdot \mathbf{E}_3)^2 &= z_1^2[1 + \sin^2 \gamma_2(\cos^2 \alpha - (1 + \cos \gamma_1 \cos \gamma_2 \cos \alpha)^2)] \\
 &\quad + z_2^2[1 + \sin^2 \gamma_1 \cos^2 \gamma_2(\cos^2 \alpha - (1 + \cos \gamma_1 \cos \gamma_2 \cos \alpha)^2)] \\
 &\quad - 2z_1 z_2[\sin \gamma_2 \sin \gamma_1 \cos \gamma_2(\cos^2 \alpha - (1 + \cos \gamma_1 \cos \gamma_2 \cos \alpha)^2)]
 \end{aligned} \tag{2.3.78}$$

$$\begin{aligned}
 &= z_1^2 + z_2^2 \\
 &\quad + (z_1 \sin \gamma_2 - z_2 \sin \gamma_1 \cos \gamma_2)^2(\cos^2 \alpha - (1 + \cos \gamma_1 \cos \gamma_2 \cos \alpha)^2)
 \end{aligned} \tag{2.3.79}$$

$$\begin{aligned}
 \mathbf{z} \cdot \mathbf{E}_2 &= z_1(\cos \gamma_2 \sin \gamma_3) + z_2(\sin \gamma_1 \sin \gamma_2 \sin \gamma_3 + \cos \gamma_1 \cos \gamma_3) \\
 &+ z_1(-\sin \gamma_2 \cos \alpha)(\cos \gamma_1 \sin \gamma_2 \sin \gamma_3 - \sin \gamma_1 \cos \gamma_3) \\
 &+ z_2(\sin \gamma_1 \cos \gamma_2 \cos \alpha)(\cos \gamma_1 \sin \gamma_2 \sin \gamma_3 - \sin \gamma_1 \cos \gamma_3)
 \end{aligned}
 \tag{2.3.80}$$

$$\begin{aligned}
 \mathbf{z} \cdot \mathbf{E}_1 &= z_1(\cos \gamma_2 \cos \gamma_3) + z_2(\sin \gamma_1 \sin \gamma_2 \cos \gamma_3 - \cos \gamma_1 \sin \gamma_3) \\
 &+ z_1(-\sin \gamma_2 \cos \alpha)(\cos \gamma_1 \sin \gamma_2 \cos \gamma_3 + \sin \gamma_1 \sin \gamma_3) \\
 &+ z_2(\sin \gamma_1 \cos \gamma_2 \cos \alpha)(\cos \gamma_1 \sin \gamma_2 \cos \gamma_3 + \sin \gamma_1 \sin \gamma_3)
 \end{aligned}
 \tag{2.3.81}$$

2.3.9 Small Angle Approximations

As the plane of the grating rotates away from its initial orientation, fringes are generated which are not due to the in-plane deformation alone, but have a part which is due to the rotation of the grating surface. The effect of this rotation can be estimated by approximating the rotations which change the orientation of the grating plane with small angles. This approximation will provide information about the leading sources of discrepancy in the equations describing the fringe spacing and orientation.

If the angles γ_1 and γ_2 (which are measures of how far the plane of the grating has varied from its initial orientation) are restricted to be small so that terms of order γ_α^2 ($\alpha = 1, 2$) can be neglected when compared with 1, the trigonometric functions sin and cos are approximated by:

$$\begin{aligned} \sin \gamma_\alpha &= \gamma_\alpha \\ \cos \gamma_\alpha &= 1 \end{aligned} \quad \alpha = 1, 2$$

The angle γ_3 is not restricted as it is part of the in-plane deformation even though this angle is usually small.

With these approximations,

$$z_1 = 2 \left(\frac{\lambda}{w} - \sin \alpha \cos \gamma_3 \right) \quad (2.3.82)$$

$$z_2 = 2(\sin \alpha \sin \gamma_3 - \gamma_1 \gamma_2 \sin \alpha \cos \gamma_3) \quad (2.3.83)$$

$$z_3 = z_1(-\gamma_2 \cos \alpha) + z_2(\gamma_1 \cos \alpha) \quad (2.3.84)$$

$$\mathbf{z} \cdot \mathbf{z} - (\mathbf{z} \cdot \mathbf{E}_3)^2 = z_1^2 + z_2^2 - (z_1 \gamma_2 - z_2 \gamma_1)^2 (1 + 2 \cos \alpha) \quad (2.3.85)$$

$$\begin{aligned} \mathbf{z} \cdot \mathbf{E}_2 &= z_1(\sin \gamma_3) + z_2(\gamma_1 \gamma_2 \sin \gamma_3 + \cos \gamma_3) \\ &\quad + z_1(-\gamma_2 \cos \alpha)(\gamma_2 \sin \gamma_3 - \gamma_1 \cos \gamma_3) \\ &\quad + z_2(\gamma_1 \cos \alpha)(\gamma_2 \sin \gamma_3 - \gamma_1 \cos \gamma_3) \end{aligned} \quad (2.3.86)$$

$$\begin{aligned} \mathbf{z} \cdot \mathbf{E}_1 &= z_1(\cos \gamma_3) + z_2(\gamma_1 \gamma_2 \cos \gamma_3 - \sin \gamma_3) \\ &\quad + z_1(-\gamma_2 \cos \alpha)(\gamma_2 \cos \gamma_3 + \gamma_1 \sin \gamma_3) \\ &\quad + z_2(\gamma_1 \cos \alpha)(\gamma_2 \cos \gamma_3 + \gamma_1 \sin \gamma_3) \end{aligned} \quad (2.3.87)$$

The third term in equation (2.3.85) can be neglected with respect to the first two terms since it is of the order $\gamma_\alpha \gamma_\beta$ times the first two terms, and is itself identically zero when the sum of the first two terms is zero. ($x^2 + y^2 = 0 \Rightarrow x \equiv y \equiv 0$). So for this approximation

$$\begin{aligned} \mathbf{z} \cdot \mathbf{z} - (\mathbf{z} \cdot \mathbf{E}_3)^2 &= z_1^2 + z_2^2 \\ &= 4 \left[\left(\frac{\lambda}{w} \right)^2 + \sin^2 \alpha - 2 \frac{\lambda}{w} \sin \alpha \cos \gamma_3 \right. \\ &\quad \left. + \gamma_1^2 \gamma_2^2 \sin^2 \alpha \cos^2 \gamma_3 - \gamma_1 \gamma_2 \sin^2 \alpha \sin 2\gamma_3 \right] \quad (2.3.88) \end{aligned}$$

$$\begin{aligned} \mathbf{z} \cdot \mathbf{E}_2 &= z_1 \sin \gamma_3 + z_2 \cos \gamma_3 \\ &\quad + z_2 \gamma_1 \gamma_2 \sin \gamma_3 (1 + \cos \alpha) \\ &\quad - z_2 \gamma_1^2 \cos \gamma_3 \cos \alpha \\ &\quad - z_1 \gamma_2^2 \sin \gamma_3 \cos \alpha \\ &\quad + z_1 \gamma_1 \gamma_2 \cos \gamma_3 \cos \alpha \quad (2.3.89) \end{aligned}$$

$$\begin{aligned} \mathbf{z} \cdot \mathbf{E}_1 &= z_1 \cos \gamma_3 - z_2 \sin \gamma_3 \\ &\quad + z_2 \gamma_1 \gamma_2 \cos \gamma_3 (1 + \cos \alpha) \\ &\quad + z_2 \gamma_1^2 \sin \gamma_3 \cos \alpha \\ &\quad - z_1 \gamma_2^2 \cos \gamma_3 \cos \alpha \\ &\quad - z_1 \gamma_1 \gamma_2 \sin \gamma_3 \cos \alpha \quad (2.3.90) \end{aligned}$$

$$\begin{aligned}
 \frac{\mathbf{z} \cdot \mathbf{E}_2}{2} &= \frac{\lambda}{w} \sin \gamma_3 \\
 &\quad - \gamma_1 \gamma_2 \sin \alpha \cos^2 \gamma_3 \\
 &\quad + \gamma_1 \gamma_2 \sin \alpha \sin^2 \gamma_3 (1 + \cos \alpha) \\
 &\quad - \gamma_1^2 \gamma_2^2 \sin \alpha \sin \gamma_3 \cos \gamma_3 (1 + \cos \alpha) \\
 &\quad - \gamma_1^2 \sin \alpha \sin \gamma_3 \cos \gamma_3 \cos \alpha \\
 &\quad + \gamma_1^3 \gamma_2 \sin \alpha \cos^2 \gamma_3 \cos \alpha \\
 &\quad - \gamma_2^2 \frac{\lambda}{w} \sin \gamma_3 \cos \alpha \\
 &\quad + \gamma_2^2 \sin \alpha \sin \gamma_3 \cos \gamma_3 \cos \alpha \\
 &\quad + \gamma_1 \gamma_2 \frac{\lambda}{w} \cos \gamma_3 \cos \alpha \\
 &\quad - \gamma_1 \gamma_2 \sin \alpha \cos^2 \gamma_3 \cos \alpha
 \end{aligned} \tag{2.3.91}$$

$$\begin{aligned}
 \frac{\mathbf{z} \cdot \mathbf{E}_1}{2} &= \frac{\lambda}{w} \cos \gamma_3 - \sin \alpha \\
 &\quad + \gamma_1 \gamma_2 \sin \alpha \sin \gamma_3 \cos \gamma_3 \\
 &\quad + \gamma_1 \gamma_2 \sin \alpha \sin \gamma_3 \cos \gamma_3 (1 + \cos \alpha) \\
 &\quad - \gamma_1^2 \gamma_2^2 \sin \alpha \cos^2 \gamma_3 (1 + \cos \alpha) \\
 &\quad + \gamma_1^2 \sin \alpha \sin^2 \gamma_3 \cos \alpha \\
 &\quad - \gamma_1^3 \gamma_2 \sin \alpha \sin \gamma_3 \cos \gamma_3 \cos \alpha \\
 &\quad - \gamma_2^2 \frac{\lambda}{w} \cos \gamma_3 \cos \alpha \\
 &\quad + \gamma_2^2 \sin \alpha \cos^2 \gamma_3 \cos \alpha \\
 &\quad - \gamma_1 \gamma_2 \frac{\lambda}{w} \sin \gamma_3 \cos \alpha \\
 &\quad + \gamma_1 \gamma_2 \sin \alpha \sin \gamma_3 \cos \gamma_3 \cos \alpha
 \end{aligned} \tag{2.3.92}$$

$$\begin{aligned}
 \frac{\mathbf{z} \cdot \mathbf{E}_2}{2} &= \frac{\lambda}{w} \sin \gamma_3 \\
 &\quad - \gamma_1 \gamma_2 \sin \alpha \cos 2\gamma_3 (1 + \cos \alpha) \\
 &\quad + \gamma_1 \gamma_2 \frac{\lambda}{w} \cos \gamma_3 \cos \alpha \\
 &\quad - \gamma_1^2 \gamma_2^2 \sin \alpha \sin \gamma_3 \cos \gamma_3 (1 + \cos \alpha) \\
 &\quad + (\gamma_2^2 - \gamma_1^2) \sin \alpha \sin \gamma_3 \cos \gamma_3 \cos \alpha \\
 &\quad + \gamma_1^3 \gamma_2 \sin \alpha \cos^2 \gamma_3 \cos \alpha \\
 &\quad - \gamma_2^2 \frac{\lambda}{w} \sin \gamma_3 \cos \alpha
 \end{aligned} \tag{2.3.93}$$

$$\begin{aligned}
 \frac{\mathbf{z} \cdot \mathbf{E}_1}{2} &= \frac{\lambda}{w} \cos \gamma_3 - \sin \alpha \\
 &\quad + \gamma_1 \gamma_2 \sin \alpha \sin 2\gamma_3 (1 + \cos \alpha) \\
 &\quad - \gamma_1 \gamma_2 \frac{\lambda}{w} \sin \gamma_3 \cos \alpha \\
 &\quad - \gamma_1^2 \gamma_2^2 \sin \alpha \cos^2 \gamma_3 (1 + \cos \alpha) \\
 &\quad + (\gamma_1^2 - \gamma_2^2) \sin \alpha \sin^2 \gamma_3 \cos \alpha \\
 &\quad - \gamma_1^3 \gamma_2 \sin \alpha \sin \gamma_3 \cos \gamma_3 \cos \alpha \\
 &\quad - \gamma_2^2 \left(\frac{\lambda}{w} \cos \gamma_3 - \sin \alpha \right) \cos \alpha
 \end{aligned} \tag{2.3.94}$$

In equation (2.3.93), the last four terms can be neglected with respect to the first three, however, in equation (2.3.94) the first five terms must be kept to capture the behavior of the denominator in all cases. So for this level of approximation,

$$\frac{\mathbf{z} \cdot \mathbf{E}_2}{2} = \frac{\lambda}{w} \sin \gamma_3 - \gamma_1 \gamma_2 \sin \alpha \cos 2\gamma_3 (1 + \cos \alpha) + \gamma_1 \gamma_2 \frac{\lambda}{w} \cos \gamma_3 \cos \alpha \tag{2.3.95}$$

$$\begin{aligned} \frac{\mathbf{z} \cdot \mathbf{E}_1}{2} = & \frac{\lambda}{w} \cos \gamma_3 - \sin \alpha + \gamma_1 \gamma_2 \sin \alpha \sin 2\gamma_3 (1 + \cos \alpha) - \gamma_1 \gamma_2 \frac{\lambda}{w} \sin \gamma_3 \cos \alpha \\ & - \gamma_1^2 \gamma_2^2 \sin \alpha \cos^2 \gamma_3 (1 + \cos \alpha) + (\gamma_1^2 - \gamma_2^2) \sin \alpha \sin^2 \gamma_3 \cos \alpha \end{aligned} \quad (2.3.96)$$

As in the case of strictly planar deformation, it is assumed that the initial grating wavelength is w_0 and that the incoming laser beams are aligned so that the fringe spacing is infinite, which leads to the relation $\sin \alpha = \lambda/w_0$. Making these substitutions (and dividing out the dependence on λ) gives

$$\delta^* = \frac{w w_0}{2[w_0^2 + w^2 - 2w w_0 \cos \gamma_3 + \gamma_1^2 \gamma_2^2 w^2 \cos^2 \gamma_3 - \gamma_1 \gamma_2 w^2 \sin 2\gamma_3]^{\frac{1}{2}}} \quad (2.3.97)$$

$$\begin{aligned} \tan \phi^* = & \frac{w_0 \sin \gamma_3 - \gamma_1 \gamma_2 w \cos 2\gamma_3 (1 + \cos \alpha) + \gamma_1 \gamma_2 w_0 \cos \gamma_3 \cos \alpha}{[w_0 \cos \gamma_3 - w + \gamma_1 \gamma_2 w \sin 2\gamma_3 (1 + \cos \alpha) - \gamma_1 \gamma_2 w_0 \sin \gamma_3 \cos \alpha \\ & - \gamma_1^2 \gamma_2^2 w \cos^2 \gamma_3 (1 + \cos \alpha) + (\gamma_1^2 - \gamma_2^2) w \sin^2 \gamma_3 \cos \alpha]} \end{aligned} \quad (2.3.98)$$

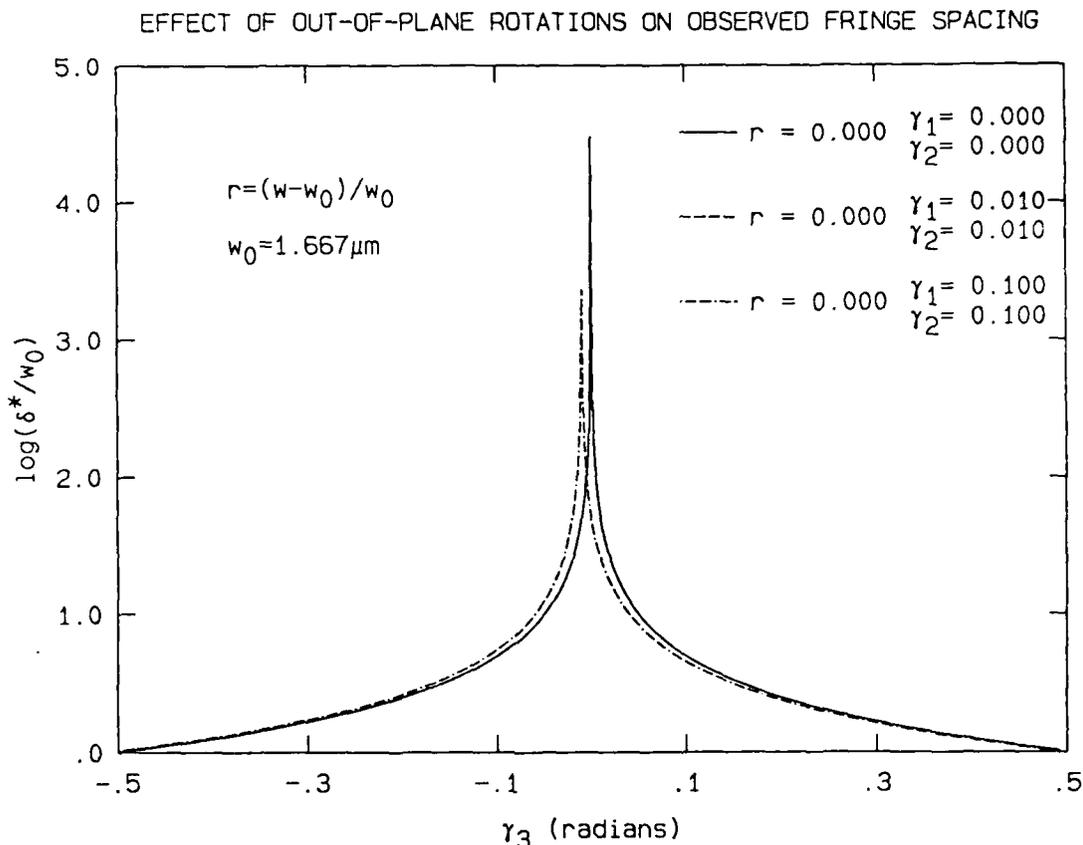


FIGURE 7. The variation for δ^* with out-of-plane rotations for the case where the stretch ratio is zero. In this case the effect of the out-of-plane rotations, is most severe. The offset of the curve for larger γ_1 and γ_2 would be in the opposite direction if γ_1 or γ_2 were negative.

2.3.10 Discussion

For planar deformation, it is clear that the equations governing the fringe spacing and inclination in moire interferometry are identical with the equations governing the fringe spacing and inclination in geometric moire techniques, which means that the fringes can be construed as contours of constant relative in-plane displacement. The increment of displacement between fringes and the direction of the displacement is determined by the angle at which the incoming laser beams intersect and the plane in which they lie. (This in turn is determined by the initial

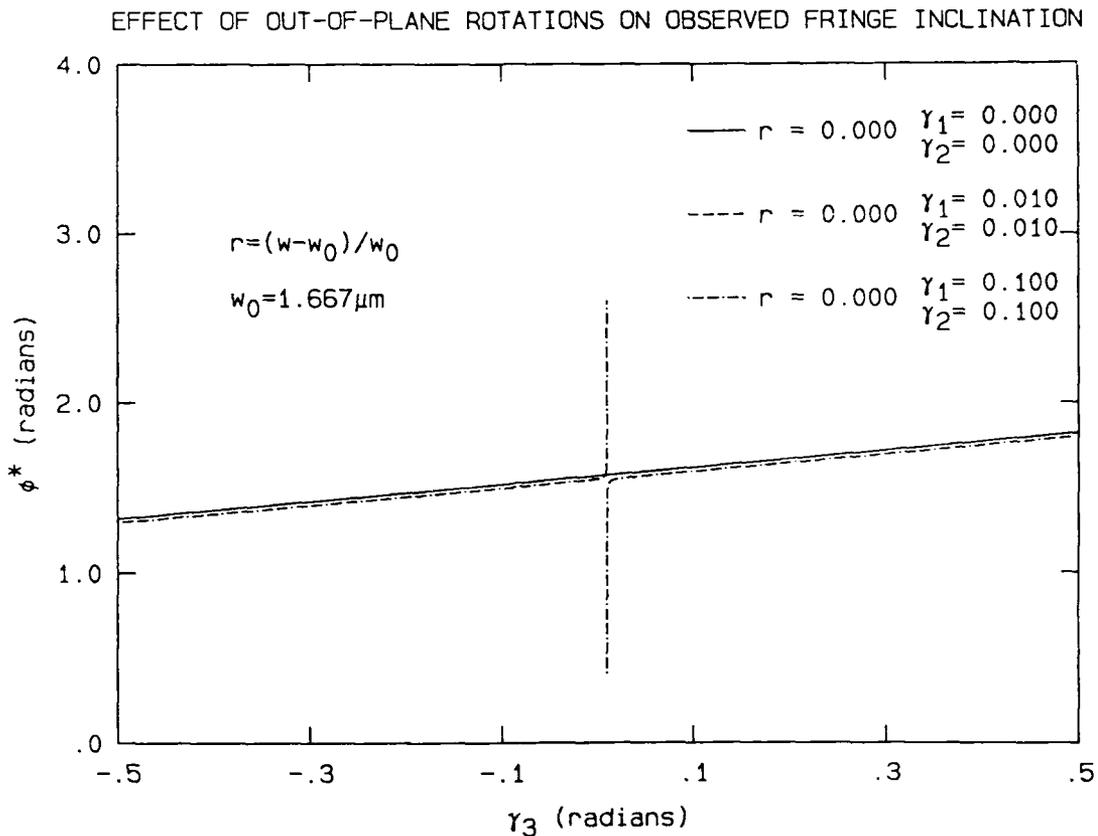


FIGURE 8. The variation for ϕ^* with out-of-plane rotations for the case where the stretch ratio is zero. Again, this is the case where the effect of the out-of-plane rotations is most severe.

state of the grating: its wavelength w_0 and orientation. The incoming laser beams are aligned using the initial grating.) Further discussion of the case of planar deformation and the relation to general moire methods can be found in papers by Livnat and Post [9] and the author [21].

These equations have been derived assuming that the deformation is homogeneous. In order to extend these equations to nonhomogeneous deformations, it must be assumed that the equations can be applied locally and that the displacement field is continuous and smooth enough for the fringes to remain continuous

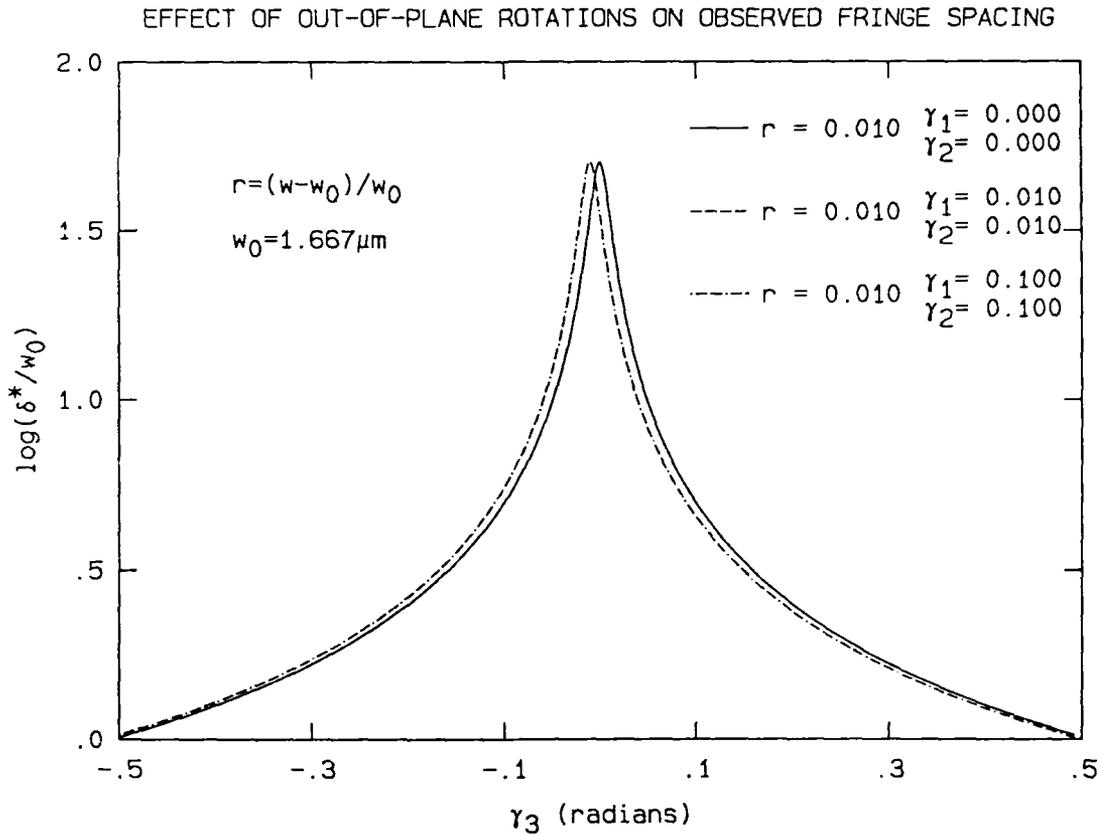


FIGURE 9. The variation for δ^* with out-of-plane rotations for the case where the stretch ratio is an intermediate value. (About 1% strain.) In this case, the error due to the out-of-plane rotations is less severe than for the case where the stretch ratio is zero.

and smooth. Of course, this raises the question of the size scale at which the localization and continuity arguments are no longer valid. Unfortunately, solutions of diffraction problems are hard to come by. Most of the work analyzing diffraction gratings is aimed at predicting the diffraction efficiency of a given grating profile. These problems are attacked by assuming a solution of plane waves for each of the diffraction orders and matching the amplitudes of those plane waves to the boundary conditions imposed by the grating profile. In the case of the interferometer, the assumption that the diffracted light is a plane wave would have to be discarded,

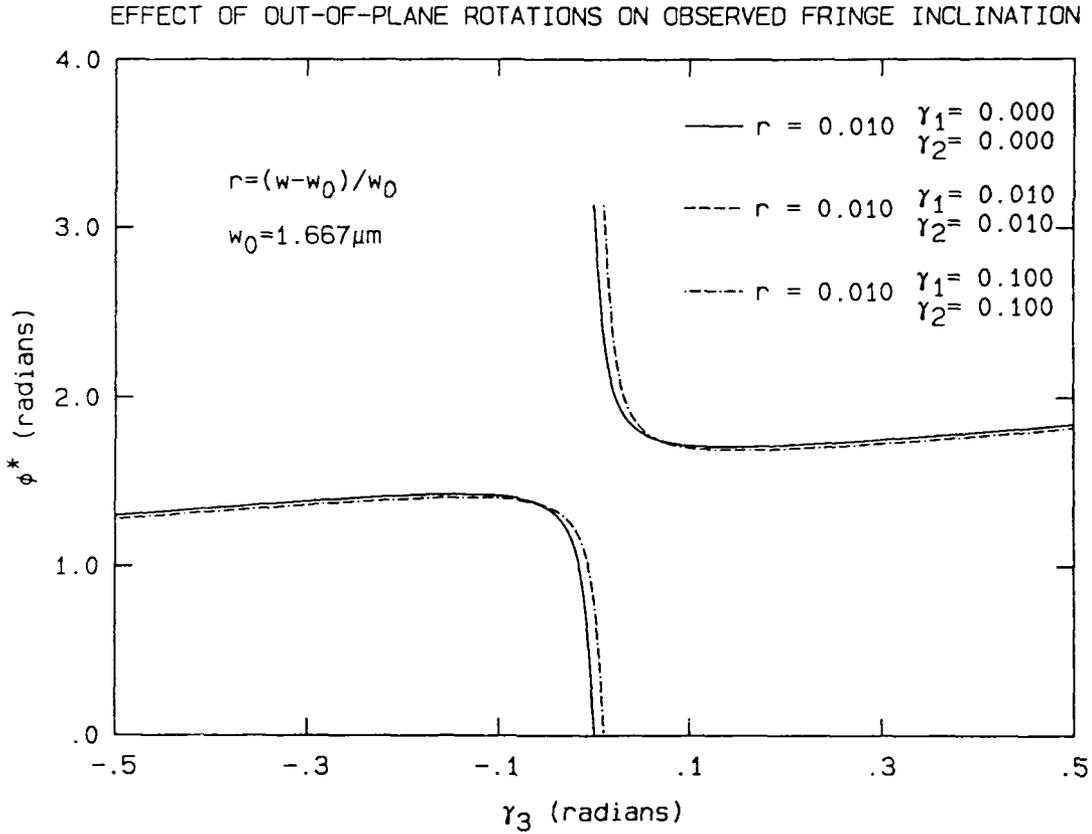


FIGURE 10. The variation for ϕ^* with out-of-plane rotations for the case where the stretch ratio is an intermediate value. (About 1% strain.) In this case the error due to the out-of-plane rotations is less severe than for the case where the stretch ratio is zero, except near the region where $\gamma_3 = 0$.

and a solution for a grating of varying spacing would need to be constructed to determine if the assumption of locally plane waves at the angle predicted by equation (2.3.1) is valid.

In undertaking the derivation of the equations describing the fringes for a general (not necessarily planar) deformation, the goal was to determine the error introduced by deformations which rotate the plane of the grating surface. This was motivated by the fact that the surface near the tip of a crack dimples inward. This

dimple is confined to a small region, and even within this region, the surface slopes are approximated well by small angle theory over most of the dimple.

The two term approximation for the out-of-plane component of the unit vectors representing the diffracted beams u_3 and v_3 is quite good for the experimental moire interferometer. The error in this approximation is less than 1 percent for deviations from the surface normal up to 30 degrees, while the theoretical limit which can be captured using the current configuration of the interferometer is about 20 degrees. The small angle approximation to the general equations simplifies the equations somewhat and helps to determine the leading sources of discrepancy in the fringe formation equations, but the equations retaining the sin and cos terms could be used as well. To a large extent the surface rotation will be in the range where small angle theory is valid. Experimentally, large surface rotations mean large deviation from the original path of the laser beams to the camera, which means that the light will be lost. Additionally, when the surface rotations are beyond small angles the out-of-plane displacement measurement made by the Twyman-Green interferometer must have a correction for its measurement.

It can be seen from Figures 7 and 8 that the out-of-plane rotations have the most effect when the stretch ratio and the in-plane rotation γ_3 are zero. As the in-plane deformation increases, the errors introduced by the out-of-plane rotations are reduced, as shown in Figures 9 and 10.

2.4 Description of the Apparatus and Experimental Technique

Moire interferometry is a method of measuring in-plane displacements at the surface of an object. The displacement information is contained in interference fringes which have been shown to represent contours of constant in-plane displacements. In order to measure these displacements, a diffraction grating must first be attached or etched onto the surface of the object to be measured. This diffraction grating is a periodic reflective surface height variation; its period is on the order of the wavelength of the light used in the measurements.

The increment of displacement between adjacent contours is proportional to the period of the diffraction grating and so is itself on the order of the wavelength of light used; the smallest theoretically possible displacement increment between contours is one-half of that wavelength. This means that moire interferometry is very sensitive to small displacement gradients, and that the fringe density becomes very high when the strains become large. This measurement technique displays the displacement contours over a field which is limited by the size of the diffraction grating and the diameter of the laser beams used to create the interference pattern. However, since the light which travels to the observer is nearly collimated, the collection device must have an aperture as large as the measurement area in order to collect all of the information from the object. As the specimen deforms, the light diverges from its initial path and at some point the aperture will not be large enough to capture all of the information from the specimen. Ordinarily, this would happen first at the edge of the field of view, but in the case of fracture the deformation may be so large at the center of the field of view near the crack tip that this region also diverts the light outside the collection apparatus.

2.4.1 Fringe Quality

The moire interferometric measurement technique relies on the interference between two laser beams at the surface of the specimen. In order to produce interference patterns of high contrast, the two laser beams must originate from the same laser and travel paths of similar length to the object. Although lasers are nearly monochromatic, the wavelength of the laser beam does vary slightly with time (and therefore with distance from the laser). This variation is often approximated by considering the laser light to be the sum of two waves with wavelengths slightly above and below the nominal laser wavelength. These two waves interfere with each other and cause the wave pattern to vary slightly in time and in space as the waves travel while still retaining the nominal laser wavelength. If the difference in wavelength between these two waves is $\Delta\lambda$ and the nominal wavelength of the laser is λ_0 , the path length difference between the two beams ΔL should be such that

$$\Delta L < \frac{1}{8} \frac{\lambda_0^2}{\Delta\lambda} \quad (2.4.1)$$

for high contrast fringes. Although the laser output is more complex than the sum of two waves, this relation is still a good estimate of the allowable path length difference, if $\Delta\lambda$ represents the bandwidth of the light waves from the laser as described by Vest [25]. This allowable path length difference is expressed as the coherence length of the laser, and for a Helium-Neon laser it is on the order of 20 cm, so that for interferometric purposes the difference in path lengths between the two legs of the interferometer should be less than 20 cm and as nearly equal to zero as possible. This notion of the coherence length is actually a description of the temporal coherence of the laser because it describes the variation of the laser output with time and is related to a length because of the propagation of the light waves and because that length is of a scale which is more easily comprehended (the

light waves travel over the coherence length in 6.67×10^{-10} seconds).

A second important coherence measurement is the spatial coherence of the light waves, which describes the similarity of the waves across the width of the beam. A laser source has excellent spatial coherence, in contrast to a diffuse monochromatic source, for instance, such as a sodium vapor tube, where the light emitted from two separate points on the source may have different phases, amplitudes and polarizations, even though they are presumed to have the same wavelength.

The coherence of the interfering light waves is the reason that the two interfering laser beams must come from a single source, since matching the outputs from two different lasers is effectively not possible. Besides matching the path lengths to maintain the temporal coherence of the laser beams, the amplitude and polarization of the two laser beams should be similar to provide high contrast fringes.

2.4.2 The Moire Interferometer

The moire interferometer used in these experiments uses two Spectra-Physics *Model 120 S* 15 mW Helium-Neon lasers, one for each of the two in-plane displacement components to be measured. It also depends strongly on precisely adjustable mirrors and beam splitters for the initial alignment with the diffraction grating. The mirrors, beam splitters and alignment hardware is all from the Newport Corporation, Fountain Valley, Ca. For each displacement component, the laser beam passes through a beam expander and lens to produce a collimated laser beam of 50 mm in diameter. This beam is then reflected from a mirror through a beam splitter where half of the beam is reflected and half transmitted. Each of the two beams from the beam splitter travels to a mirror which directs the beam to the

specimen surface. See Figures 11 through 14. This arrangement consisting of a laser, a beam splitter and three mirrors is used twice, once to provide laser beams which cross in a vertical plane to measure the vertical component of displacement, and once to provide laser beams which cross in a horizontal plane, measuring the horizontal component of displacement. All of the mirrors and beam splitters are angularly adjustable by means of micrometers which rotate the optical components about fixed pivots in their mounts. This feature of the interferometer makes it quite flexible and provides ease of alignment of the laser beams so that they intersect at the specimen surface at the proper angle and at a specific location.

All of the components are mounted on a Newport Corporation breadboard measuring three feet by four feet, and the breadboard is held upright in a vertical plane on a Newport Corporation vibration isolated optical table which is ten feet by four feet by 18 inches thick.

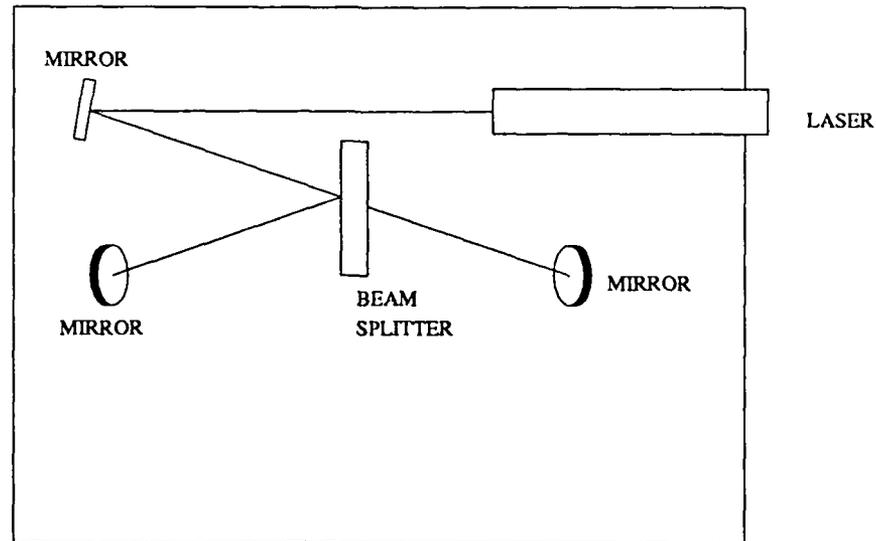


FIGURE 11. Diagram of the interferometer configuration to measure the horizontal component of displacement.

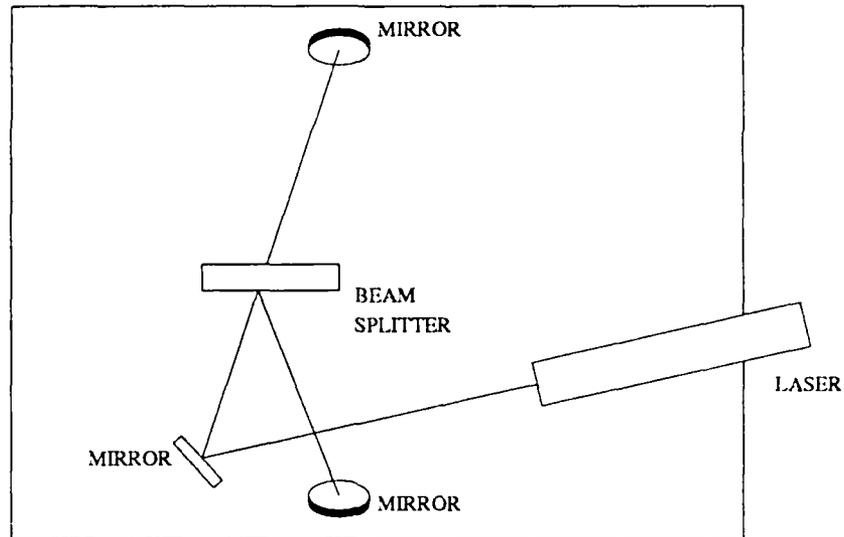


FIGURE 12. Diagram of the interferometer configuration to measure the vertical component of displacement.

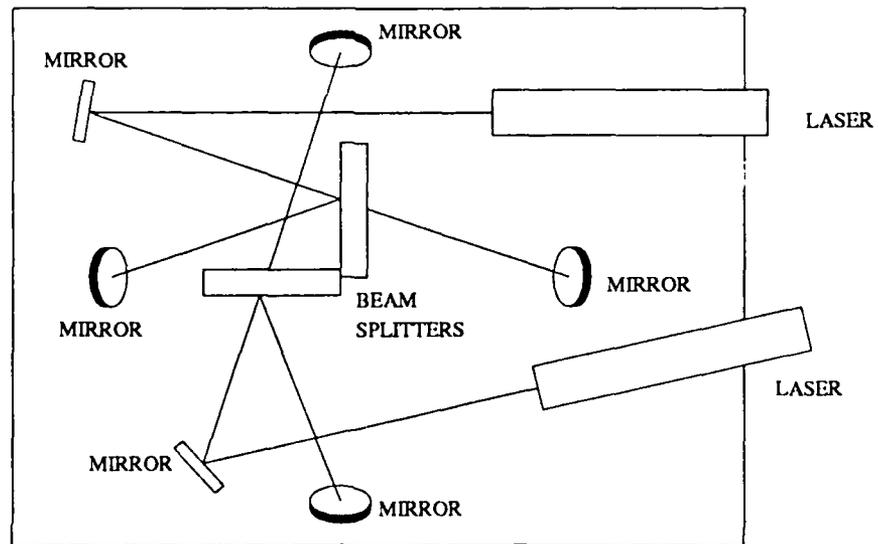


FIGURE 13. Diagram of the interferometer configuration to measure both the horizontal and vertical components of displacement.

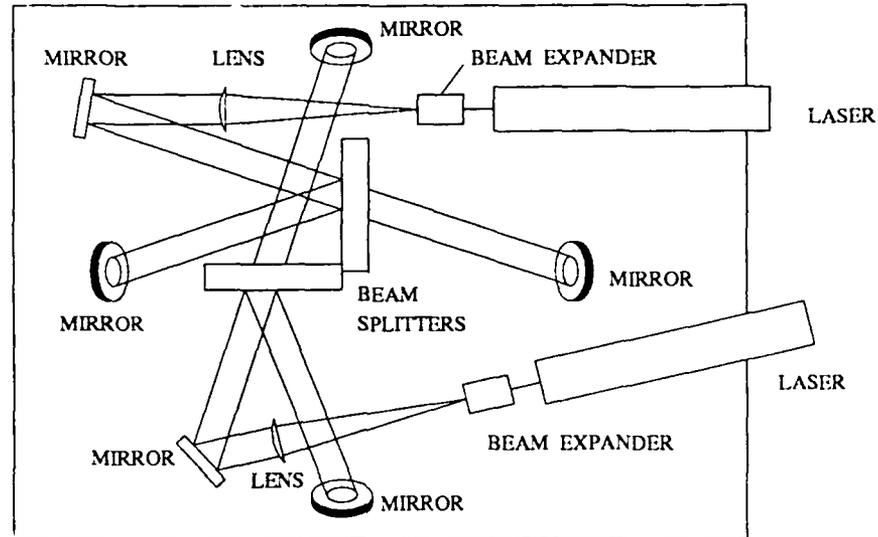


FIGURE 14. Diagram of the interferometer configuration to measure both of the displacement components with optics for beam expansion and collimation. All optics are mounted on a Newport Corporation breadboard which measures three feet by four feet. The breadboard is held in a vertical plane. All four laser beams intersect approximately one meter from the breadboard surface.

2.4.3 Interferometer Alignment

The interferometer is aligned by first aligning the unexpanded laser beams. The beam which passes through the beam splitter is aligned first by adjusting the angles of the mirror which lies between the laser and the beam splitter (Mirror A in Figure 15) and the the mirror which the laser beam reaches after passing through the beam splitter (Mirror B in Figure 15). By adjusting these mirrors alternately, the laser beam can be brought to the specimen surface at the desired location and at the proper angle. The angle can be checked by tracing the path of the second diffraction order beam from the specimen. Since the first diffraction order beam should exit along the normal to the specimen surface, the second diffraction order beam should return along the path of the incoming beam. Although this order

carries considerably less of the energy than the zero order or first order beams, it is visible when using the unexpanded laser beam. After the beam which passes through the beam splitter is aligned, the beam which is reflected from the beam splitter is aligned in an analogous manner by alternately adjusting the angles of the beam splitter and the mirror the beam reaches after reflecting from the beam splitter (Mirror C in Figure 15). Slight angular changes in the orientation of the beam splitter do not significantly alter the path of the laser beam passing through it.

Once the unexpanded laser beam has been aligned, the collimating lens is inserted into the path and adjusted so that the beam passing through it is not deflected from its original route and so that the reflection from the lens travels back to the laser. This ensures that the lens is centered on the laser beam and that it is placed perpendicularly with respect to the path of the beam. The beam expander, a microscope objective and pinhole arrangement, is then aligned. The microscope objective lens focuses the laser beam to a small diameter and the pinhole filters out noise. In essence the microscope objective acts as a Fourier transform on the laser light in spatial coordinates and the pinhole filters all but the central peak of the beam in the spatial frequency domain, leaving a smooth beam that consists only of components of low spatial frequency. The distance from the collimating lens to the pinhole is adjusted so that the pinhole lies at a focal point of the lens; the light that leaves the lens then consists of plane waves. At this stage the overlapping beams diffracted from the specimen can be adjusted to produce a null field, that is, the number of fringes visible is minimized, and zero if possible. (If there is an initial pattern, it is hoped that it is negligible compared to the pattern obtained when the specimen is loaded. For steel, the initial pattern is in general only one or two random fringes. For polymers, however, the initial pattern can be several fringes.)

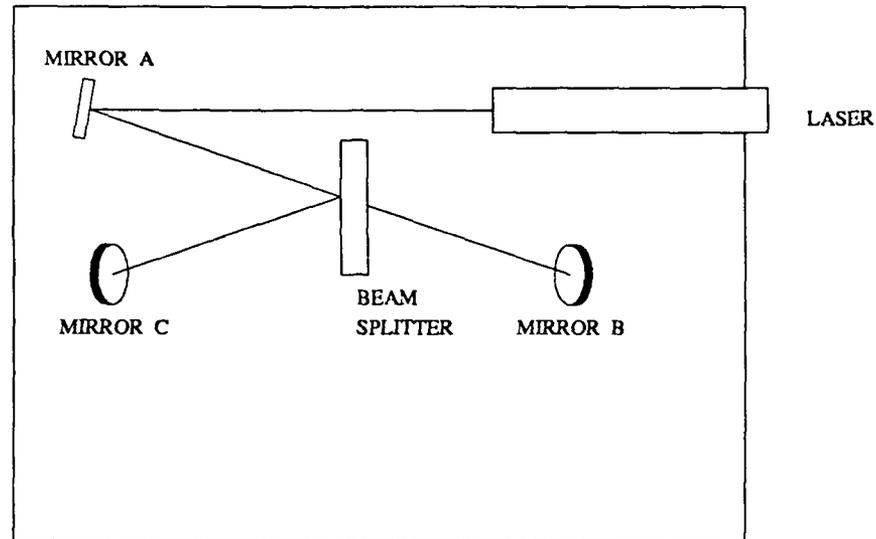


FIGURE 15. Diagram of interferometer configuration to measure the horizontal displacement components. Mirrors A, B and C and the beam splitter can be angularly adjusted about two axes with micrometers. Mirrors A and B are adjusted to bring the the beam which passes through the beam splitter to the specimen at a precise location and at proper angle so that the first diffraction order beam exits along the normal to the specimen surface. Then the beam splitter and mirror C is adjusted to bring the beam which reflects from the beam splitter to the correct location at the proper angle. (Changing the angle of the beam splitter slightly does not cause the beam which passes through the beam splitter to deviate significantly from its path.) The beam expander is then inserted into the beam, and mirror B or C adjusted to minimize the number of fringes. The interferometer configuration for measuring the vertical displacement component is adjusted in same way.

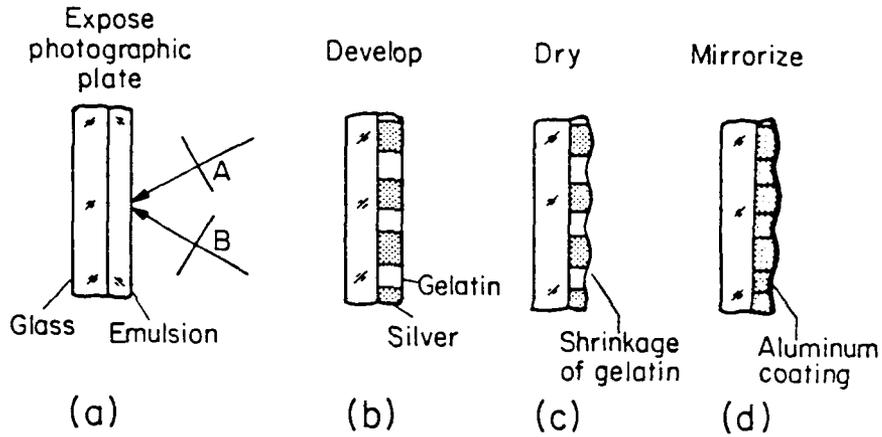
This completes the alignment of the interferometer.

2.4.4 The Grating

The specimen gratings are applied after the method of Post [18,19]. Figure 16 was taken from Post [18], and describes the process. A holographic plate is exposed to two crossed laser beams. The crossing of the beams produces an interference pattern of standing waves which creates a periodic light intensity variation across the surface of the holographic plate as shown in Figure 2. This period can be changed by changing the angle at which the laser beams cross. Again, the laser beams must originate from the same source and travel paths of similar length to create an interference pattern of high contrast. During the exposure, the structure of the emulsion on the holographic plate changes. Where the emulsion is exposed to light, the silver halide crystals suspended in the emulsion absorb energy and form metallic silver; this happens with more frequency in regions where the light is bright than where it is dark [7]. During development, the conversion to metallic silver is completed in those silver halide crystals where it has begun. When the plate is fixed, the unchanged silver halide crystals are removed, causing the emulsion to shrink in the regions where there is less metallic silver. This shrinkage therefore follows the irradiance pattern of the interfering laser beams, creating a periodic surface height variation which acts as a reflective diffraction grating. By coating this rippled surface with optical quality aluminum in a thin film, the reflectivity is enhanced and so is the diffraction efficiency. This rippled surface variation and the aluminum coating can be transferred to a specimen if the adherence between the aluminum film and the holographic emulsion is weak.

The details of the preparation of the specimen grating follow. Kodak Type 649 F spectroscopic holographic plates are exposed twice to two crossed laser beams emanating from a 15 mW Spectra-physics Helium-Neon laser, with the plates being

PREPARATION OF MOLD



REPLICATION ON SPECIMEN

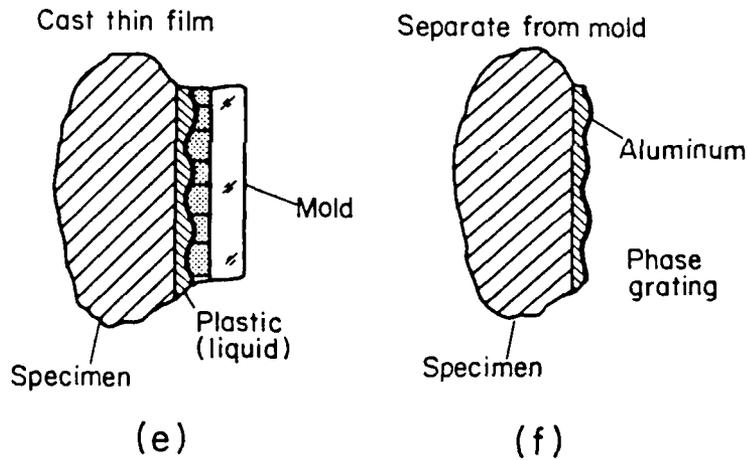


FIGURE 16. The process for preparing the specimen grating and transferring it to the specimen. (After Post [18].)

rotated by 90 degrees between exposures to produce orthogonal sets of lines making two orthogonal specimen gratings. The exposure time in each case is two seconds. The Kodak 649 F plates are sensitive to the 632.8 nm Helium-Neon laser wavelength and have an emulsion of 17 μms thick with a nominal sensitivity of 70 $\mu\text{J}/\text{cm}^2$ and a resolution of 2000 lines/mm. The plates are anti-halation backed to reduce reflection from the rear surface of the plates. The plates are developed as recommended by

Kodak, 6 minutes in Kodak D-19 developer, 30 seconds in Kodak indicator stop bath, 3 minutes in Kodak rapid fixer, rinse in water for 25 minutes and dip for 30 seconds in Kodak Photo-flo solution. After the plates have dried, they are dipped a second time in the Photo-flo solution for 30 seconds. When the plates are completely dry, they are coated with a thin coating of reflective aluminum in a vacuum coating chamber. This aluminum coating is on the order of 200 nm in thickness. The grating is transferred to the specimen using Photolastic PC-1 epoxy adhesive. (Photolastic Division, Measurements Group, Inc., P.O. Box 27777, Raleigh, North Carolina, 27611, (919) 365-3800) The double dip in the Photo-flo solution provides a very weak interface between the aluminum film and the emulsion on the holographic plate.

The gratings used in this study that have been made this way have a wavelength of $1.667 \mu\text{m}$ or 600 lines/mm. The gratings are produced to a known wavelength by aligning the crossed laser beam angle to a commercially produced grating of wavelength $3.333 \mu\text{m}$ or 300 lines/mm in exactly the same way that the moire interferometer is aligned to the specimen grating. If there are no fringes visible over a 25.4 cm field of view, the wavelength of the laser interference is twice that of the alignment grating to within 0.01 %. The grating was obtained from Diffraction Products, Inc., P.O. Box 645, Woodstock, Illinois, 60098. (815)338-6768.

2.4.5 Requirements and Limitations for the Collection Optics

The maximum strain which can be measured corresponds to the minimum fringe spacing which can be captured by the optical system and recorded on the film. Since the diffracted laser beams are still approximately collimated, the collection optics must have an aperture at least as large as the field to be viewed in order to record information from the entire field. The fringes observed in the experiment are the result of angular deviations of the diffracted beams from the original paths. Higher fringe densities, corresponding to higher displacement gradients (strains), are the result of relatively large angular deviations from the original specimen normal. Hence, in order to capture large strain behavior, the angular aperture of the collection optics must be large.

The angular aperture is a measure of the largest angle a ray may attain without being lost by the optical system; it is inversely related to the f number of the optical system, so a large angular aperture corresponds to a small f number. The f number of a lens is the ratio of the focal length of the lens divided by its aperture diameter.

The ability to capture fringes representing large strains is also limited by the resolution of the film used to record the image. For the Kodak Technical Pan film used in the experiments, the minimum fringe wavelength which can be recorded on the film is about 5 microns [8]. However, the modulation transfer function of the film does drop off quickly as the fringe spacing decreases.

The modulation transfer function is a measure of how well the contrast of the fringes is transferred to the recording film. Each optical component has a modulation transfer function, they assemble in a complicated fashion resulting in a modulation transfer function for the entire system. All modulation transfer functions tend toward zero as the fringe wavelength becomes smaller. Even a simple

aperture without a lens shows this feature because diffraction from the edges of the aperture will cause the fringes to blur when the fringe wavelength becomes small enough.

The optical system can also be limited by aberrations in the lens, as well as diffraction from the edges of any aperture the light must pass through. These aberrations cause a blurring of the image which will smear out the light intensity distribution, destroying fringes of short wavelength. The Nikon camera lenses used in the experiment are optimized to be diffraction limited, that is, the most significant aberration in the lens is diffraction due to the aperture, and not due to defects in the lens. As a first estimate of the resolution limits of the optical recording system, only the effects of the angular aperture limit on the maximum observable diffraction angle will be considered. Diffraction from the edges of the aperture and the effect of aberrations will be neglected. (These effects are all included in the modulation transfer function for the system.) The angular aperture will be assumed to be dominated by the lens aperture closest to the specimen (any ray which passes through the lens aperture closest to the specimen will be assumed to pass through the rest of the lenses in the system and fall on the film plane). In this case, the angular aperture is only a function of the distance between the specimen surface and the initial lens aperture and is assumed not to depend on the internal workings of the lens system.

If the camera aperture has a diameter A and the distance from the specimen to the aperture is L , as in Figure 17, the angular aperture, ϕ , is found from

$$\tan \phi = \frac{A}{2L} \quad (2.4.2)$$

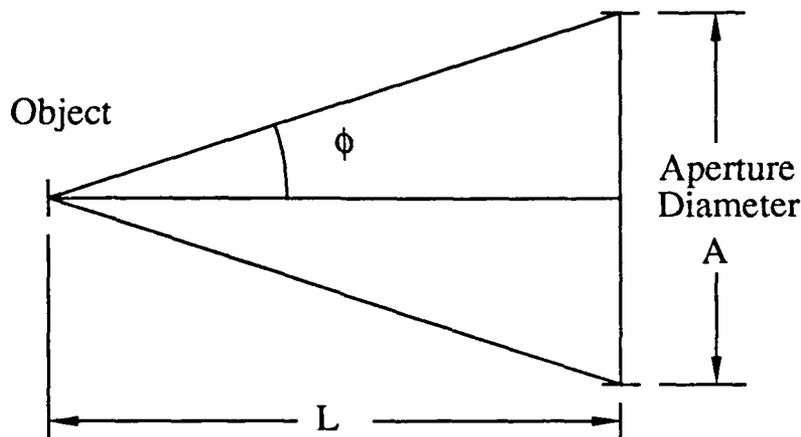


FIGURE 17. The angular aperture, ϕ , is a measure of the largest angle a ray from the center of the field may achieve and still pass through the aperture of the lens closest to the specimen. L is the distance from the specimen to the aperture, A is the aperture diameter, and $\tan \phi = A/(2L)$.

If two diffracted light rays (representing the normals of locally plane light waves) leave the specimen surface from the center of the field of view and just pass through the aperture of the lens, they will give rise to fringes with a wavelength δ , which depends on the angular aperture ϕ by

$$\delta = \frac{\lambda}{2 \sin \phi} \quad (2.4.3)$$

where λ is the wavelength of light used. In planar deformation, the light waves are diffracted symmetrically about the normal to the specimen surface, so the two rays intersect the aperture at either end of a diameter of the aperture, which leads to the relation in equation (2.4.3) above. For a small angular aperture ϕ , $\tan \phi \approx \sin \phi$. This is a good approximation for the camera lens used in this study, which had an angular aperture of about 0.08 radians (about 5°). For small angular aperture ϕ ,

then, δ be approximated by

$$\delta = \frac{\lambda L}{A} \quad (2.4.4)$$

For an optical system which has a magnification M , the fringes which arrive at the film plane have a wavelength $M\delta$. The film is able to record fringes which have wavelengths as small as $5\mu\text{m}$. Using film, the fringes must be magnified so that

$$5\mu\text{m} < M\delta \quad (2.4.5)$$

in order to capture fringes of wavelength δ at the specimen surface. From these equations, the required magnification can be related to the angular aperture, and thus to the camera aperture and specimen to aperture distance) as

$$\frac{5\mu\text{m}}{M} < \frac{\lambda L}{A} \quad (2.4.6)$$

For a Helium-Neon laser with $\lambda = 0.6328\mu\text{m}$ this gives

$$\frac{7.90}{M} < \frac{L}{A} \quad (2.4.7)$$

The required magnification M for a specimen to aperture distance L is shown in Figure 18 with an aperture A equal to 50 mm.

The Nikon Micro-Nikkor 200mm f/4 Internal Focus lens to be used in the experiments has an aperture of 50mm and at its closest focus it has a magnification of 1/2 with a specimen to aperture distance of 500mm. With the addition of the Nikon TC-200 Teleconverter, the magnification is doubled to 1 without changing the specimen to aperture distance. The lens alone does not provide enough magnification for the film to record the smallest fringes which the lens can capture. However, with the addition of the teleconverter, the magnification is sufficient for the film to resolve the smallest fringes, but at the expense of imaging only one quarter of the

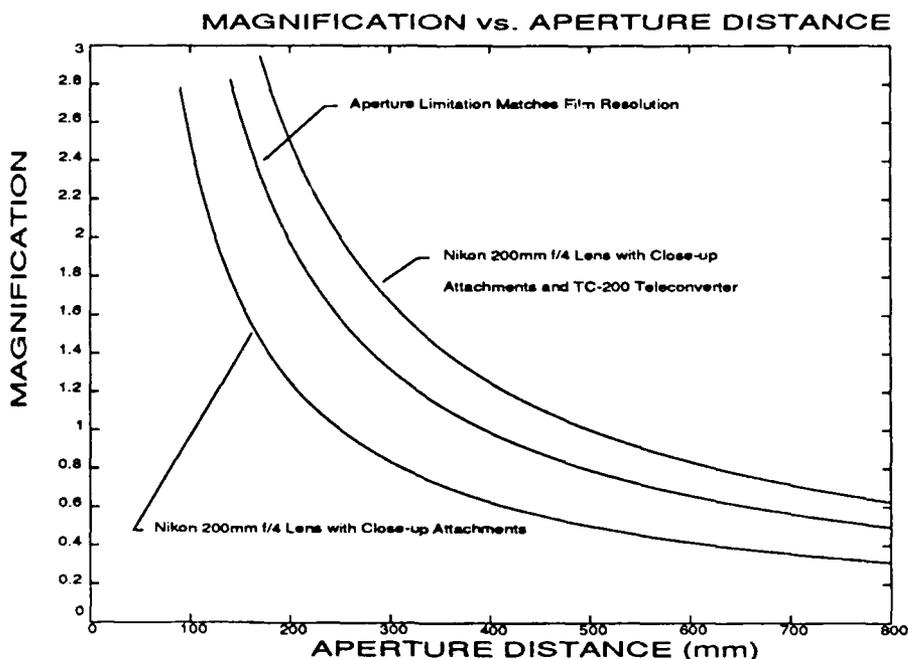


FIGURE 18. Required magnification to match film resolution to aperture limitation. Performance of Nikon 200 mm f/4 with close-up attachments.

area seen without the teleconverter. The minimum resolvable fringe wavelength can be improved by the use of close-up attachment lenses, which decrease the specimen to aperture distance and improve the magnification by creating a virtual object, which is farther away than the actual object, but is proportionately larger than the actual object. These attachment lenses can be approximated by thin lens theory, creating a virtual object which is 500mm away as seen by the primary 200mm f/4 lens. Addition of such a lens gives a relation between the magnification and the specimen to camera distance of

$$M = \frac{500\text{mm}}{2L} \tag{2.4.8}$$

The magnification is doubled with the addition of the teleconverter. The required focal length of the added close-up lens is given by

$$f = \frac{500\text{mm}}{2M - 1} = \frac{(L)500\text{mm}}{500\text{mm} - L} \tag{2.4.9}$$

The necessary magnification to match the film resolution with a 50mm aperture is shown in Figure 18, along with the performance of the Nikon 200mm f/4 lens with close-up attachments and teleconverter. It is necessary to be above the curve where the film resolution matches the aperture limitation to record all of the fringes which travel through the optical system.

The magnification can also be increased by use of an extender which moves the film plane farther away from the lens; this also reduces the specimen to aperture distance. While this is not normally recommended since the lens is optimized to focus on the film plane at a certain distance, the Nikon 200mm f/4 lens is designed to give better performance at this off-design usage than other lenses. For this means of improving the magnification and angular aperture, the relation between the magnification and the specimen to aperture distance can be described by

$$M = \frac{167\text{mm}}{L - 167\text{mm}} \quad (2.4.10)$$

The performance of the Nikon 200mm f/4 with extensions is shown in Figure 19. The addition of the teleconverter would double the magnification without changing the specimen to aperture distance.

2.4.6 Physical Limitation

The maximum angular aperture of the recording device is limited by the fact that it cannot block the incoming laser beams, as illustrated in Figure 20. Additionally, if the aperture were to get in the way of the incoming beams, it would also be in a position to capture light from the zeroth order reflected beams. For a given specimen grating diameter of D , an aperture diameter of A , separation L between

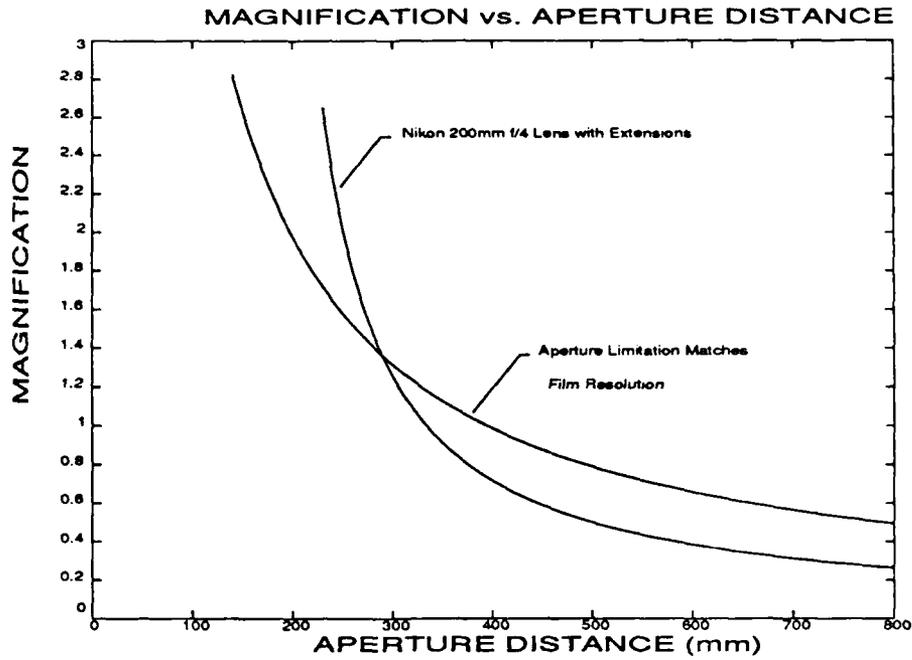


FIGURE 19. Required magnification to match film resolution to aperture limitation. Performance of Nikon 200 mm f/4 with extensions.

the two and angle α between the normal to the grating surface and incoming laser beams, it can be seen from Figure 20 and Figure 17 that

$$\tan \alpha = \frac{A + D}{2L} \quad (2.4.11)$$

$$\tan \phi = \frac{A}{2L} = \frac{A \tan \alpha}{A + D} \quad (2.4.12)$$

where ϕ is the maximum allowable angular aperture without blocking the incoming laser beams. The largest possible ϕ occurs when A becomes very large compared with D , which gives $\phi = \alpha$. Of course, as A becomes large, so must L , unless α is very large.

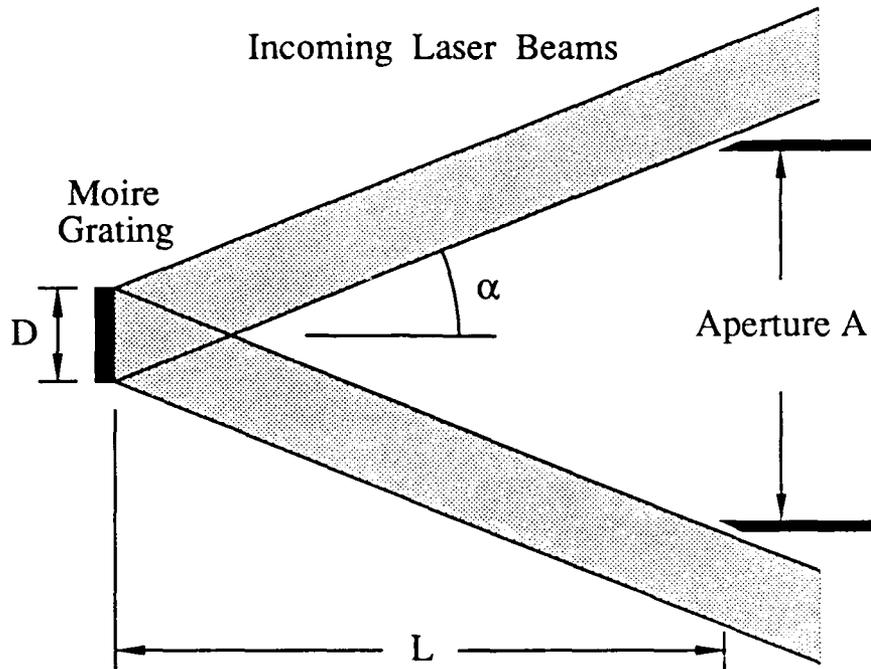


FIGURE 20. The aperture of the optical recording device may not be so large that it blocks the incoming laser beams. This limits the angular aperture to be less than the angle α of the incoming laser beams. If the aperture lies just inside the incoming beams, $\tan \alpha = (D + A)/(2L)$, while the angular aperture ϕ is determined from the relation $\tan \phi = A/(2L)$.

2.4.7 Sensitivity

The sensitivity of the measurement technique depends on the wavelength of the undeformed diffraction grating, w_0 . The displacement increment from one black fringe to an adjacent black fringe is $w_0/2$. The incoming laser beams are arranged in such a way that $\sin \alpha = \lambda/w_0$. From this relation it is seen that the minimum possible grating wavelength is $w_0 = \lambda$, and the absolute limit on the sensitivity is $\lambda/2$, identical to the sensitivity of the Twyman-Green interferometer. This requires that $\alpha = \pi/2$, which means the incoming laser beams are parallel to the surface. However, because the sensitivity is inversely proportional to the sin of α , it is

possible to make a reasonable approach to the limit with a significantly smaller value for α .

2.4.8 Strain Measurement

The minimum fringe spacing can be converted to a measure of the maximum strain if the deformation is assumed to have a simple form. If it is assumed that the fringes are due entirely to stretching (no shear or rotation), the fringe spacing δ is found to be

$$\delta = \frac{ww_0}{2|w - w_0|} \quad (2.4.13)$$

where w_0 is the initial grating wavelength, and w is the current grating wavelength. The engineering strain, e , is defined by the stretch ratio $(w - w_0)/w_0$, which leads to the equation

$$\delta = \frac{w_0(1 + e)}{2|e|} \quad (2.4.14)$$

$$\frac{|e|}{1 + e} = \frac{w_0}{2\delta} \quad (2.4.15)$$

The largest stretch ratio which can be measured corresponds to the minimum δ , which depends on ϕ , the angular aperture of the camera. The parameter $w_0/(2\delta)$ is the ratio of the sensitivity of the measurement to the minimum fringe spacing, and it is also proportional to the angular aperture of the system, since

$$\frac{w_0}{2\delta} = \frac{\sin \phi}{\sin \alpha} \quad (2.4.16)$$

Note that $\sin \phi / \sin \alpha$ varies between zero and unity. The variation of e with $w_0/(2\delta)$ is shown in Figure 21. Figure 22 shows a plot of the same data with the engineering strain e replaced by the 'true strain', ϵ . The engineering strain, e , and the 'true strain,' ϵ , are related by the equation

$$\epsilon = \ln(1 + e) \quad (2.4.17)$$

which is the same equation used in the reduction of the uniaxial stress-strain data (See the section on uniaxial data in the chapter on plastic fracture). Of course, for large stretch ratios the nonlinearity of the strain-displacement relation must be accounted for, but the observable stretch ratio is a parameter which will quickly show how well the interferometer can perform, and how close it is to its limits.

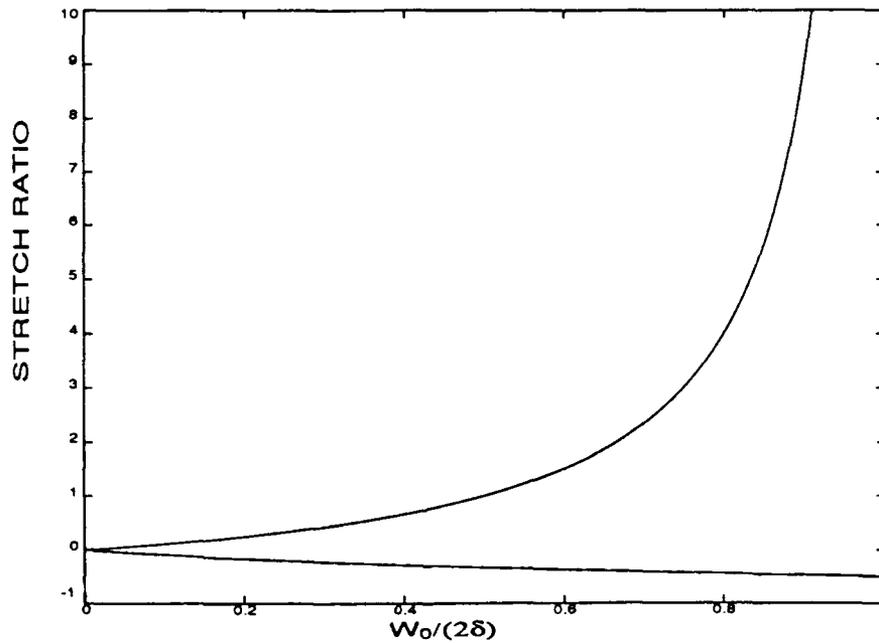


FIGURE 21. The stretch ratio $(w - w_0)/w_0$ as a function of fringe spacing δ . The parameter $w_0/(2\delta)$ is proportional to the angular aperture of the optical system.

It is interesting that the observable stretch ratios do not behave symmetrically in extension and contraction, due to the absolute value of the change in the grating wavelength, $w - w_0$, in equation (2.4.13). If the grating wavelength increases, in the limit the grating will eventually appear like a flat surface, and the diffracted beam will approach the zeroth order reflected beam. In this case, the strain can approach infinity while the angle of the diffracted ray is bounded. If the wavelength

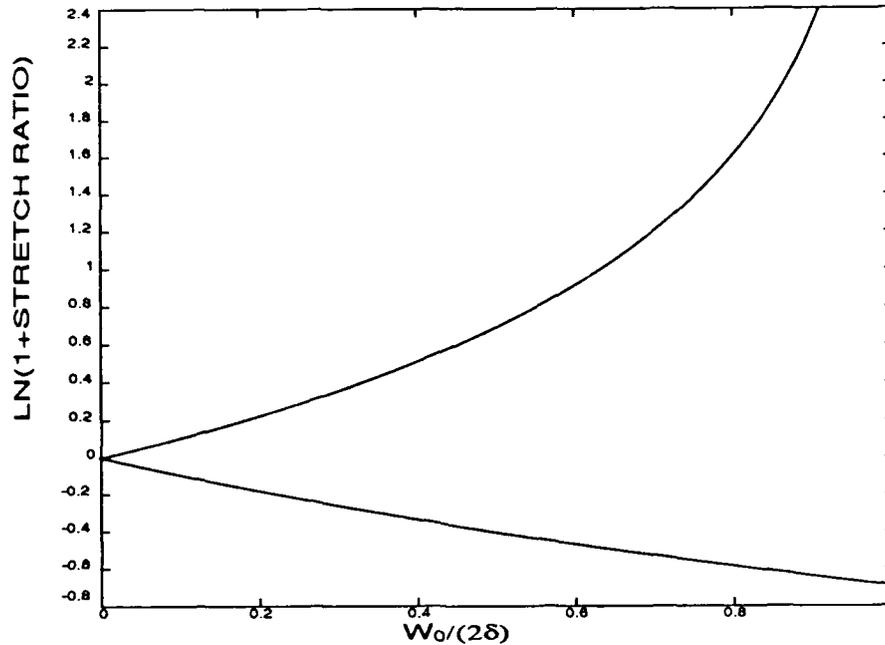


FIGURE 22. The function $\ln(w/w_0)$ as a function of fringe spacing δ . In a uniaxial test, the log of the 1 plus the stretch ratio is called the 'true strain.'

decreases, the diffracted ray has a different kind of constraint. As the wavelength of the grating decreases, the diffracted ray may keep rotating until it is parallel to the grating surface. At that point, no solution to the diffraction equation exists (it implies that $\sin \phi > 1$), but the grating wavelength may continue to decrease.

2.4.9 Compromises

A compromise must be struck between the desire for high sensitivity in the measurement technique and the desire to measure large displacement gradients. There are a number of parameters which affect these quantities that must be determined. The wavelength of light is fixed at $0.6328\mu\text{m}$ by the use of a Helium-Neon laser source. It is assumed that the aperture diameter A and grating field diameter D are equal, for this experiment it was chosen to be 50 mm. If the camera

is capable of a magnification of 2, and the film can record fringes of $5\mu\text{m}$, the smallest fringe spacing on the specimen is $2.5\mu\text{m}$. This indicates a need for an angular aperture of $\phi = 0.127$ radians $= 7.3^\circ$. Since A and D are equal, equations (2.4.11) and (2.4.12) indicate that $\tan \alpha = 2 \tan \phi$, so α must be greater than or equal to 0.250 radians $= 14.3^\circ$. This value for alpha yields a grating wavelength $w_0 = 2.56\mu\text{m}$. The parameter $w_0/(2\delta)$ is then equal to 0.51, and the corresponding stretch ratios observable range from -0.34 to 1.04, or 'true strains' between -0.36 and 0.72.

On the other hand, if the sensitivity is chosen to be about equal to the wavelength of light, so that the moire interferometer has half of the sensitivity of the Twyman-Green interferometer, $w_0 = 1.266\mu\text{m}$. Using this value for w_0 and $2.5\mu\text{m}$ for δ yields measurable stretch ratios between -0.20 and 0.34, or 'true strains' between -0.22 and 0.29.

The specimen gratings are made with a precise wavelength by aligning the intersection angle of the laser beams at the holographic plate with a commercially produced grating. A compromise was made by choosing a grating with a wavelength of $3.333\mu\text{m}$ for alignment. The specimen grating produced has a wavelength half that of the alignment grating, so $w_0 \approx 1.667\mu\text{m}$. With $\delta = 2.5\mu\text{m}$, the measurable stretch ratios lie between -0.25 and 0.50, and the 'true strains' lie between -0.29 and 0.41.

The camera magnification of 2 is chosen somewhat arbitrarily, although with a 50mm field, a magnification of 2 would fill a 4 inch by 5 inch plate for a framing camera. If larger magnifications can be made, the lower limit on the fringe spacing grows smaller. Still assuming that $A = D = 50\text{mm}$ and $\lambda = 0.6328\mu\text{m}$, and with $w_0 = 1.667\mu\text{m}$, the smallest δ that can be captured without blocking the incoming

laser beams is $1.57\mu\text{m}$, assuming the magnification is high enough to allow the film to record the fringes. This is the smallest fringe spacing which can be observed with the limitations imposed on A , D , λ and w_0 . The stretch ratios are then between -0.35 and 1.11, and the 'true strains' between -0.43 and 0.75.

In the experiment for this study, the highest magnification used was about 1.4, $A=D=50\text{mm}$, $L=320\text{ mm}$. This gives $\tan \phi = 0.08$ so $\delta = 3.8\mu\text{m}$. The measurable stretch ratios are then predicted to lie between -0.18 and 0.28, and the 'true strains' between -0.20 and 0.25.

2.4.10 Performance

With the case of $w_0 = 1.667\mu\text{m}$, the minimum possible fringe spacing is $\delta = w_0/2 = 0.8333\mu\text{m}$ for an infinitely large aperture. With the limitation to a finite aperture of 50 mm, the minimum fringe spacing increases to $1.57\mu\text{m}$. Thus the introduction of an aperture has increased the lower bound on the fringe spacing by only a factor of 1.9, assuming that aperture is just inside the limit for blocking the beams. In the experiment for this study, with a predicted minimum fringe spacing of $3.8\mu\text{m}$, the minimum δ is 4.5 the lower bound for an infinite aperture, and 2.4 times the lower bound for a 50 mm aperture. The next experiment will be done as close to the 50 mm aperture limit as possible.

2.5 Orthogonal Displacements with a Three Beam Moire Interferometer

The moire interferometer used in this study was originally designed to use three laser beams to measure orthogonal in-plane displacements instead of using four laser beams. This type of interferometer is described by McDonach, McKelvie and Walker [11]. The configuration of optical components to make such measurements is shown in Figure 23 and can be compared with the configuration of the four beam interferometer in Figure 24.

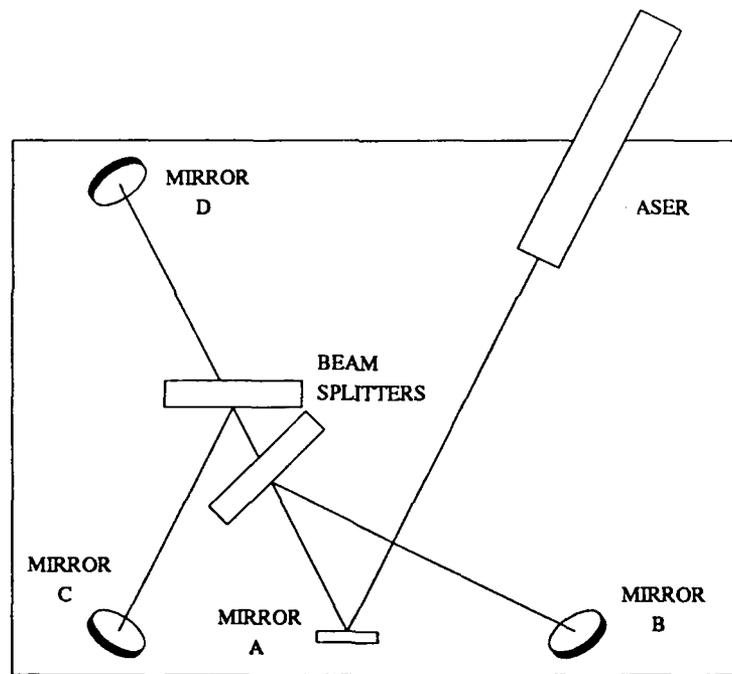


FIGURE 23. Diagram of the three beam interferometer configuration to measure both horizontal and vertical displacement components. Beams from mirrors B and C yield the horizontal component, and beams from mirrors C and D yield the vertical.

The three beam interferometer has several advantages over the four beam interferometer; it uses two fewer mirrors and only one laser source. However, there are problems due to unequal beam intensity and polarization mismatch, and there is

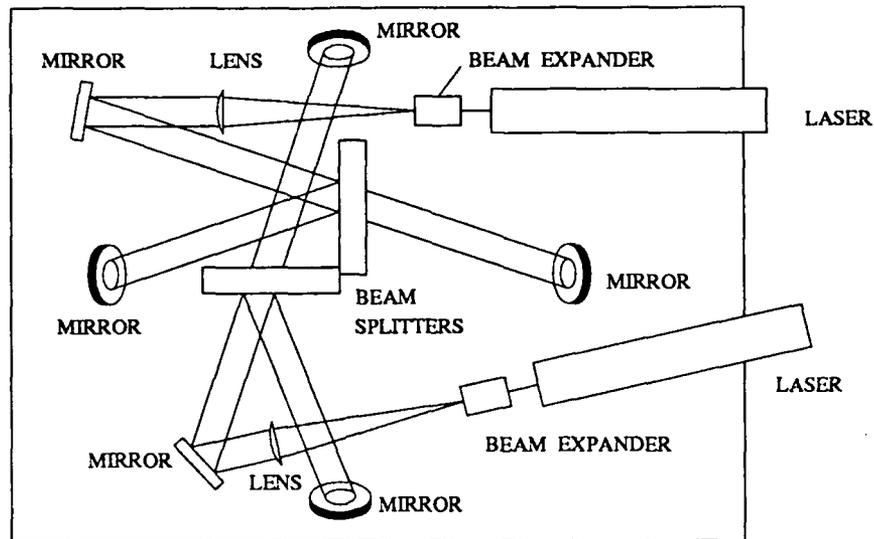


FIGURE 24. Diagram of the four beam interferometer configuration to measure both horizontal and vertical displacement components.

considerably less light as the single laser beam must be divided twice, necessitating longer exposure times and making vibrations more of a problem. Figure 25 shows the three beam interferometer with the addition of the beam expander to produce collimated beams of 50 mm diameter.

The three beam interferometer uses two crossed gratings at $\pm 45^\circ$ to the globally fixed X_1 and X_2 axes to measure components of displacement along those axes. As in the four beam interferometer, the three incoming laser beams are aligned so that the first diffraction order exists along the normal to the specimen surface (the X_3 axis). Interference between any pair of diffracted laser beams will produce fringe patterns which are related to the deformation of the grating.

Referring to Figure 23, interference between the beams from mirrors B and C produces a fringe pattern of contours of horizontal displacement; interference between the beams from mirrors C and D produces a fringe pattern of contours of

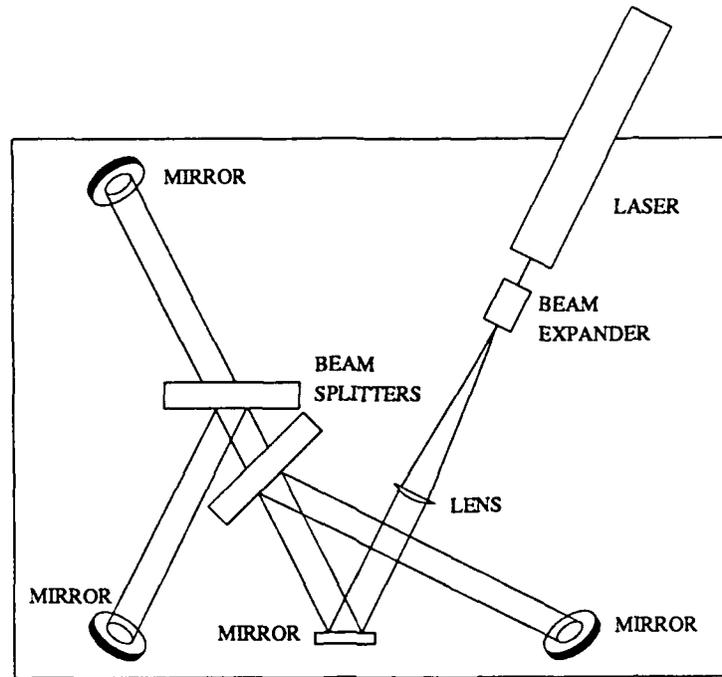


FIGURE 25. Diagram of the three beam interferometer configuration to measure both horizontal and vertical displacement components, including components to expand the beam.

vertical displacement; and interference between the beams from mirrors B and D produces a fringe pattern of contours of displacement parallel to the line between mirrors B and D.

Analysis of the governing equations of fringe formation for the three beam interferometer will follow the lines of the derivation for the four beam interferometer. However, only the case of planar deformations will be analyzed. The effects of rotations about axes in the plane can be determined in a manner analogous to that in the section describing the four beam interferometer with arbitrary changes in grating, if required.

2.5.1 Grating Geometry

The interference patterns observed will come from two light beams, each diffracted from one of the crossed gratings. The mean grating surface coincides with the $X_1 - X_2$ plane in a globally fixed right handed coordinate frame. Initially, the gratings have identical wavelengths of w_0 and are perpendicular to each other, but at angles of $\pm 45^\circ$ to the X_1 axis.

Let \mathbf{E}_1 , \mathbf{E}_2 and \mathbf{E}_3 be unit vectors aligned with the X_1 , X_2 and X_3 axes, respectively. Let the unit normal to the lines of grating #1 be \mathbf{n}_1 , and let \mathbf{t}_1 be a unit vector parallel to the lines of grating #1 such that \mathbf{n}_1 , \mathbf{t}_1 and \mathbf{E}_3 form a right handed orthogonal coordinate system. Let the unit normal to the lines of grating #2 be \mathbf{n}_2 , and let \mathbf{t}_2 be a unit vector parallel to the lines of grating #2 such that \mathbf{n}_2 , \mathbf{t}_2 and \mathbf{E}_3 form another right handed coordinate system. Initially, \mathbf{n}_1 is at an angle of $\frac{\pi}{4}$ radians to the X_1 axis, and \mathbf{n}_2 is at an angle of $\frac{3\pi}{4}$ radians to the X_1 axis. The gratings are allowed to deform independently. Grating #1 stretches or contracts to a new wavelength w_1 , and the grating lines are rotated so that the normal \mathbf{n}_1 is at an angle of $(\frac{\pi}{4} + \beta_1)$ radians to the X_1 axis. Grating #2 stretches or contracts to a new wavelength w_2 , and these grating lines rotate so that the normal \mathbf{n}_2 is at an angle of $(\frac{3\pi}{4} + \beta_2)$ radians to the X_1 axis. See Figures 26 and 26. Figure 26 shows the overlap of the crossed gratings, and Figure 27 shows the geometry of the two gratings separately.

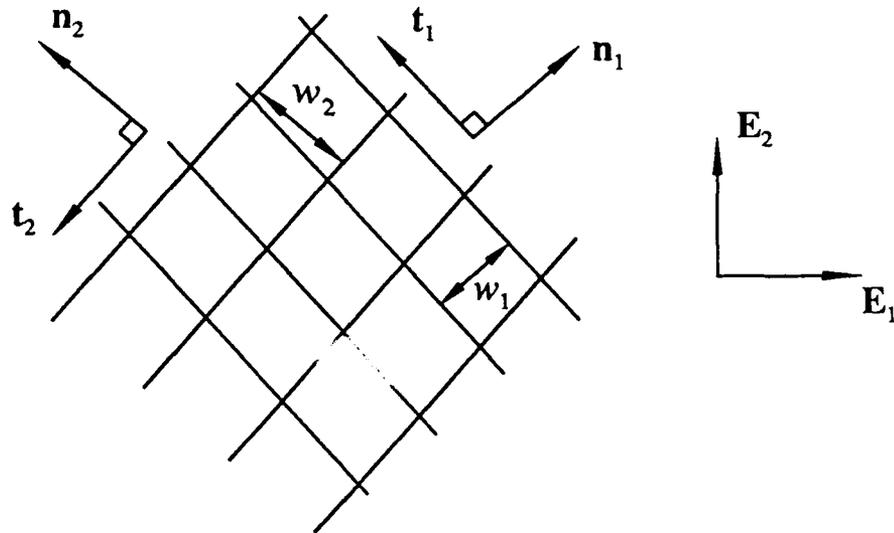


FIGURE 26. Diagram of crossed lines of grating used in three beam interferometer. Vector n_1 is normal to ridges of grating #1, which are spaced w_1 apart. Vector n_2 is normal to ridges of grating #2, which are spaced w_2 apart. Unit vectors E_1 and E_2 lie along the globally fixed coordinate axes.

2.5.2 Diffraction Equations

To calculate the direction of a diffracted light ray (of any order) from an incident light ray which is directed at a grating from an arbitrary angle, the incident light ray must be resolved into components with respect to a coordinate system which is aligned with the lines of the grating as it deforms. For instance, if unit vector a represents the incident ray and unit vector $b^{(n)}$ represents the n^{th} order ray diffracted from the incident ray, the equations relating the components of a to the components of $b^{(n)}$ are:

$$a_1 + b_1^{(n)} = \frac{n\lambda}{w} \quad (2.5.1)$$

$$a_2 + b_2^{(n)} = 0 \quad (2.5.2)$$

where a_1 and $b_1^{(n)}$ are the components of a and $b^{(n)}$ perpendicular to the lines of the grating and parallel to the mean grating surface; a_2 and $b_2^{(n)}$ are the components of

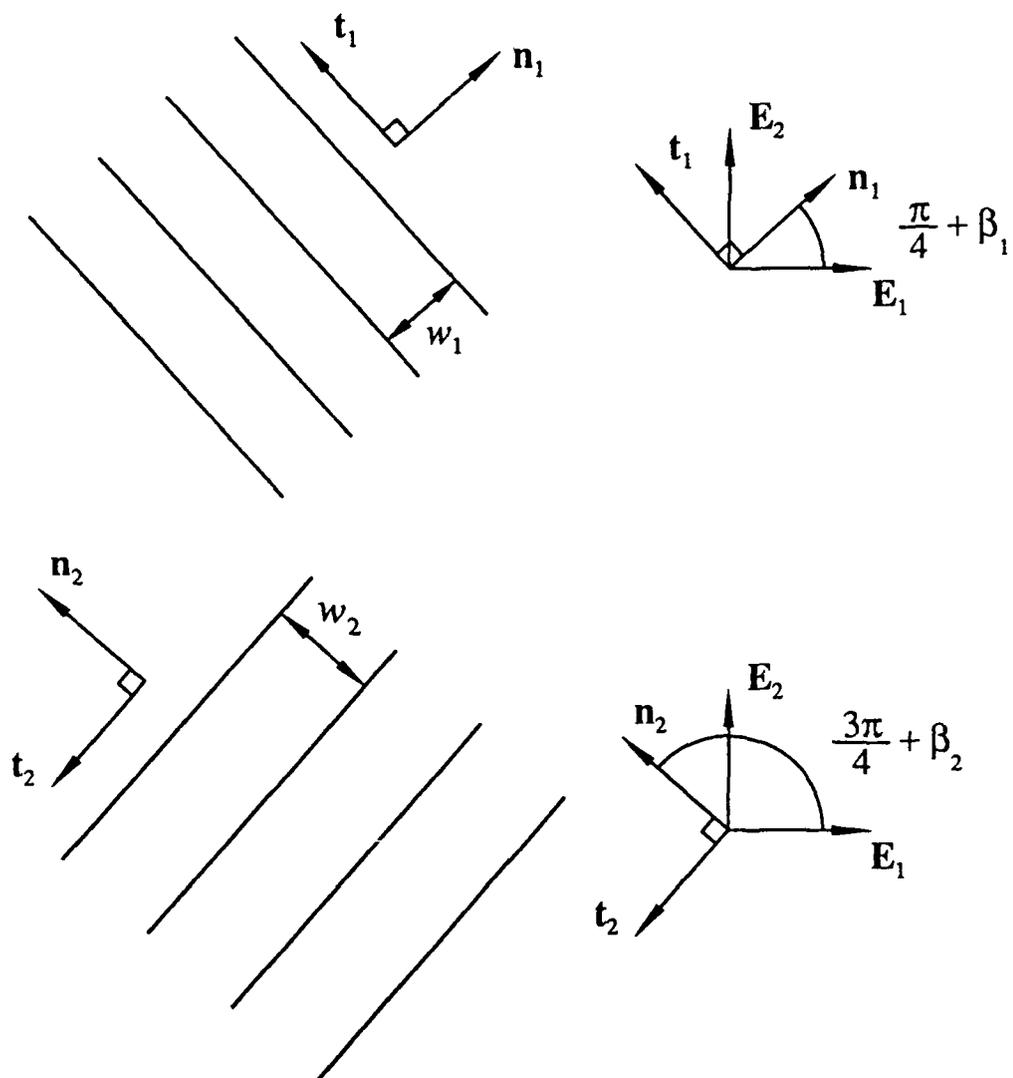


FIGURE 27. Diagram showing the geometry of gratings #1 and #2 separately. Grating #1 has wavelength w_1 . Unit vector n_1 is normal to the ridges of grating #1 and unit vector t_1 is parallel to the ridges of grating #1. The angle between the globally fixed unit vector E_1 and the unit vector n_1 is $\frac{\pi}{4} + \beta_1$. Grating #2 has wavelength w_2 . Unit vector n_2 is normal to the ridges of grating #2 and unit vector t_2 is parallel to the ridges of grating #2. The angle between the globally fixed unit vector E_1 and the unit vector n_2 is $\frac{3\pi}{4} + \beta_2$.

a and $b^{(n)}$ parallel to the both the lines of the grating and the mean grating surface.

Again, λ is the wavelength of light used, and w is the current wavelength of the

grating. a_3 and b_3 lie along the normal to the mean grating surface, and both are positive, which means that while the diffracted light rays are directed along $\mathbf{b}^{(n)}$, the incident light rays are actually directed along $-\mathbf{a}$. Since \mathbf{a} and $\mathbf{b}^{(n)}$ are unit vectors, equations (2.5.1) and (2.5.2) uniquely determine $\mathbf{b}^{(n)}$ from any given \mathbf{a} .

Let the diffracted light ray of interest from grating #1 be represented by the unit vector \mathbf{u} , and let the diffracted light ray of interest from grating #2 be represented by unit vector \mathbf{v} . Let \mathbf{u} have components u_1 , u_2 and U_3 in the $\mathbf{n}_1\mathbf{t}_1\mathbf{E}_3$ coordinate system and components U_1 , U_2 and U_3 in the $\mathbf{E}_1\mathbf{E}_2\mathbf{E}_3$ coordinate system. Let \mathbf{v} have components v_1 , v_2 and V_3 in the $\mathbf{n}_2\mathbf{t}_2\mathbf{E}_3$ coordinate system and components V_1 , V_2 and V_3 in the $\mathbf{E}_1\mathbf{E}_2\mathbf{E}_3$ coordinate system. The fixed unit vector representing the incident light ray to grating #1 is at an angle α to the X_3 axis and at an angle of $-\frac{3\pi}{4}$ radians to the X_1 axis. The fixed unit vector representing the incident light ray to grating #2 is also at an angle α to the X_3 axis and at an angle of $-\frac{\pi}{4}$ radians to the X_1 axis. The components of the diffracted rays, u_1 , u_2 , v_1 and v_2 are found using equations (2.5.1) and (2.5.2),

$$-\sin \alpha \cos \beta_1 + u_1 = -\frac{\lambda}{w_1} \quad (2.5.3)$$

$$\sin \alpha \sin \beta_1 + u_2 = 0 \quad (2.5.4)$$

$$-\sin \alpha \cos \beta_2 + v_1 = -\frac{\lambda}{w_2} \quad (2.5.5)$$

$$\sin \alpha \sin \beta_2 + v_2 = 0 \quad (2.5.6)$$

Again, since \mathbf{u} and \mathbf{v} are unit vectors with $u_3 > 0$ and $v_3 > 0$, \mathbf{u} and \mathbf{v} are completely determined from equations (2.5.3) through (2.5.6).

$$U_1 = u_1 \cos\left(\frac{\pi}{4} + \beta_1\right) - u_2 \sin\left(\frac{\pi}{4} + \beta_1\right) \quad (2.5.7)$$

$$U_2 = u_1 \sin\left(\frac{\pi}{4} + \beta_1\right) + u_2 \cos\left(\frac{\pi}{4} + \beta_1\right) \quad (2.5.8)$$

$$V_1 = v_1 \cos\left(\frac{3\pi}{4} + \beta_2\right) - v_2 \sin\left(\frac{3\pi}{4} + \beta_2\right) \quad (2.5.9)$$

$$V_2 = v_1 \sin\left(\frac{3\pi}{4} + \beta_2\right) + v_2 \cos\left(\frac{3\pi}{4} + \beta_2\right) \quad (2.5.10)$$

Using trigonometric identities for addition of angles,

$$\cos\left(\frac{\pi}{4} + \beta_1\right) = \frac{1}{\sqrt{2}}(\cos \beta_1 - \sin \beta_1) \quad (2.5.11)$$

$$\sin\left(\frac{\pi}{4} + \beta_1\right) = \frac{1}{\sqrt{2}}(\cos \beta_1 + \sin \beta_1) \quad (2.5.12)$$

$$\cos\left(\frac{3\pi}{4} + \beta_2\right) = \frac{-1}{\sqrt{2}}(\cos \beta_2 + \sin \beta_2) \quad (2.5.13)$$

$$\sin\left(\frac{3\pi}{4} + \beta_2\right) = \frac{1}{\sqrt{2}}(\cos \beta_2 - \sin \beta_2) \quad (2.5.14)$$

which leads to the result,

$$U_1 = \frac{1}{\sqrt{2}} \left[u_1(\cos \beta_1 - \sin \beta_1) - u_2(\cos \beta_1 + \sin \beta_1) \right] \quad (2.5.15)$$

$$U_2 = \frac{1}{\sqrt{2}} \left[u_1(\cos \beta_1 + \sin \beta_1) - u_2(\cos \beta_1 - \sin \beta_1) \right] \quad (2.5.16)$$

$$V_1 = \frac{-1}{\sqrt{2}} \left[v_1(\cos \beta_2 + \sin \beta_2) + v_2(\cos \beta_2 - \sin \beta_2) \right] \quad (2.5.17)$$

$$V_2 = \frac{1}{\sqrt{2}} \left[v_1(\cos \beta_2 - \sin \beta_2) - v_2(\cos \beta_2 + \sin \beta_2) \right] \quad (2.5.18)$$

From equations (2.5.4), (2.5.5), (2.5.6) and (2.5.7)

$$u_1 = -\frac{\lambda}{w_1} + \sin \alpha \cos \beta_1 \quad (2.5.19)$$

$$u_2 = -\sin \alpha \sin \beta_1 \quad (2.5.20)$$

$$v_1 = -\frac{\lambda}{w_2} + \sin \alpha \cos \beta_2 \quad (2.5.21)$$

$$v_2 = -\sin \alpha \sin \beta_2 \quad (2.5.22)$$

Therefore,

$$U_1 = \frac{1}{\sqrt{2}} \left[-\frac{\lambda}{w_1} (\cos \beta_1 - \sin \beta_1) + \sin \alpha \right] \quad (2.5.23)$$

$$U_2 = \frac{1}{\sqrt{2}} \left[-\frac{\lambda}{w_1} (\cos \beta_1 + \sin \beta_1) + \sin \alpha \right] \quad (2.5.24)$$

$$V_1 = \frac{1}{\sqrt{2}} \left[\frac{\lambda}{w_2} (\cos \beta_2 + \sin \beta_2) - \sin \alpha \right] \quad (2.5.25)$$

$$V_2 = \frac{1}{\sqrt{2}} \left[-\frac{\lambda}{w_2} (\cos \beta_2 - \sin \beta_2) + \sin \alpha \right] \quad (2.5.26)$$

2.5.3 Initial Alignment

Initially, before any deformation occurs, the incoming laser beams are aligned so that the diffracted beams are parallel to the X_3 axis. The initial conditions are that $\beta_1 = \beta_2 = 0$, $w_1 = w_2 = w_0$ and $u_1 = u_2 = v_1 = v_3 = 0$. This leads to the conclusion that

$$\sin \alpha = \frac{\lambda}{w_0} \quad (2.5.27)$$

This does not change during the deformation.

Making the substitution from equation (2.5.27) into equations (2.5.23) through (2.5.26),

$$U_1 = \frac{1}{\sqrt{2}} \left[-\frac{\lambda}{w_1} (\cos \beta_1 - \sin \beta_1) + \frac{\lambda}{w_0} \right] \quad (2.5.28)$$

$$U_2 = \frac{1}{\sqrt{2}} \left[-\frac{\lambda}{w_1} (\cos \beta_1 + \sin \beta_1) + \frac{\lambda}{w_0} \right] \quad (2.5.29)$$

$$V_1 = \frac{1}{\sqrt{2}} \left[\frac{\lambda}{w_2} (\cos \beta_2 + \sin \beta_2) - \frac{\lambda}{w_0} \right] \quad (2.5.30)$$

$$V_2 = \frac{1}{\sqrt{2}} \left[-\frac{\lambda}{w_2} (\cos \beta_2 - \sin \beta_2) + \frac{\lambda}{w_0} \right] \quad (2.5.31)$$

2.5.4 Observed Fringe Spacing and Inclination

As described in Section 2.3, the observed fringe spacing, δ^* , and the angle between the fringes and the X_2 axis, ϕ^* , can be determined from the vector difference $\mathbf{z} = \mathbf{u} - \mathbf{v}$. These relations are found to be

$$\delta^* = \frac{\lambda}{[\mathbf{z} \cdot \mathbf{z} - (\mathbf{z} \cdot \mathbf{E}_3)^2]^{\frac{1}{2}}} \quad (2.5.32)$$

$$\tan \phi^* = \frac{\mathbf{z} \cdot \mathbf{E}_2}{\mathbf{z} \cdot \mathbf{E}_1} \quad (2.5.33)$$

Note that if the vector \mathbf{z} is written in terms of components in the globally fixed X_1, X_2, X_3 coordinate system, only the components Z_1 and Z_2 are needed since

$$\begin{aligned} \mathbf{z} \cdot \mathbf{z} - (\mathbf{z} \cdot \mathbf{E}_3)^2 &= (Z_1)^2 + (Z_2)^2 + (Z_3)^2 - (Z_3)^2 \\ &= (Z_1)^2 + (Z_2)^2 \end{aligned} \quad (2.5.34)$$

and

$$\mathbf{z} \cdot \mathbf{E}_2 = Z_2 \quad \mathbf{z} \cdot \mathbf{E}_1 = Z_1 \quad (2.5.35)$$

From equations (2.5.28) through (2.5.31) it is found

$$\begin{aligned} Z_1 &= U_1 - V_1 \\ &= \frac{1}{\sqrt{2}} \left[-\frac{\lambda}{w_1} (\cos \beta_1 - \sin \beta_1) - \frac{\lambda}{w_2} (\cos \beta_2 + \sin \beta_2) + 2 \frac{\lambda}{w_0} \right] \end{aligned} \quad (2.5.36)$$

$$\begin{aligned} Z_2 &= U_2 - V_2 \\ &= \frac{1}{\sqrt{2}} \left[-\frac{\lambda}{w_1} (\cos \beta_1 + \sin \beta_1) + \frac{\lambda}{w_2} (\cos \beta_2 - \sin \beta_2) \right] \end{aligned} \quad (2.5.37)$$

2.5.5 The Third Grating

It is proposed that the interference pattern produced by the light diffracted from gratings #1 and #2, originally with lines at $\pm 45^\circ$ to the X_1 or X_2 axes as described above, is equivalent to an interference pattern due to the deformation of a grating (grating #3) which originally had lines parallel to the X_2 axis. From Figure 28, it can be seen that the peaks and valleys of a crossed grating at $\pm 45^\circ$ to the X_1 axis with wavelength w_0 make peaks and valleys for a grating at 90° with a wavelength of $w_0/\sqrt{2}$.

As seen in Figure 29, and from Durelli and Parks [2], or Theocaris [23], it can be shown that after deformation, grating #3 will have a wavelength w which is related to the wavelengths of the deformed gratings #1 and #2 (originally at $\pm 45^\circ$) by the equation:

$$w = \frac{w_1 w_2}{[w_1^2 + w_2^2 - 2w_1 w_2 \cos \theta]^{\frac{1}{2}}} \quad (2.5.38)$$

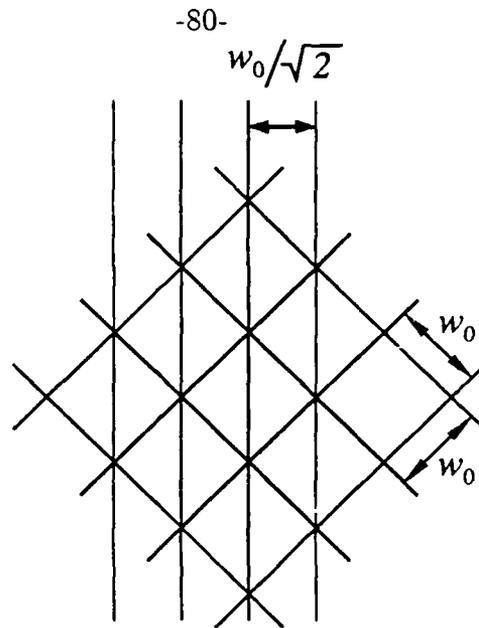


FIGURE 28. The ridges and valleys of gratings which cross at $\pm 45^\circ$ and which have identical wavelengths of w_0 form a third grating which has a wavelength of $w_0/\sqrt{2}$.

where w_1 is the wavelength of grating #1, w_2 is the wavelength of grating #2, and θ is the angle between the unit normals of gratings #1 and #2, \mathbf{n}_1 and \mathbf{n}_2 (or, equivalently, the angle between the lines of grating #1 and the lines of grating #2). The angle between the lines of grating #1 and the lines of grating #3 is denoted ψ , and it is related to gratings #1 and #2 by the equations:

$$\sin \psi = \frac{w_1 \sin \theta}{[w_1^2 + w_2^2 - 2w_1 w_2 \cos \theta]^{\frac{1}{2}}} = \left(\frac{w_1 \sin \theta}{w_2} \right) \quad (2.5.39)$$

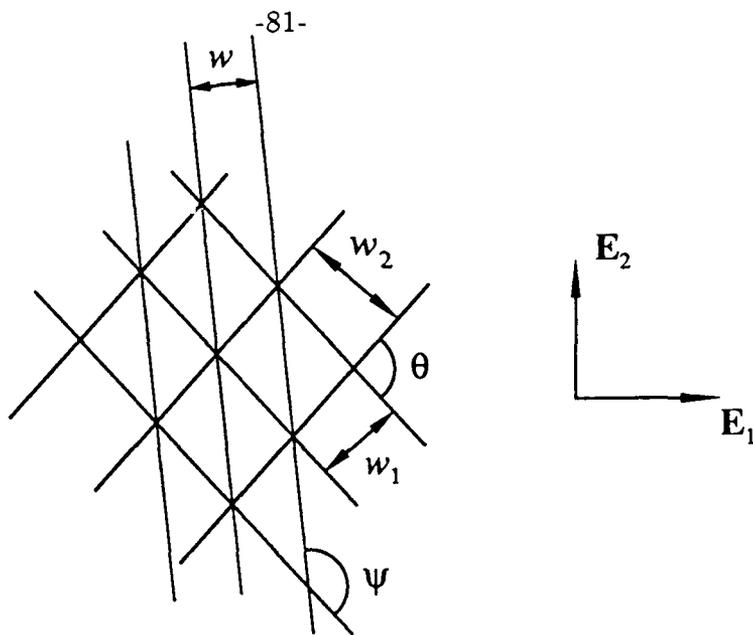


FIGURE 29. The geometry of the third grating after deformation. Grating #1 has wavelength w_1 , grating #2 has wavelength w_2 , and the angle between the ridges of grating #1 and grating #2 is θ . The third grating has wavelength w and the angle between the ridges of grating #1 and grating #3 is ψ .

$$\cos \psi = \frac{w_1 \cos \theta - w_2}{[w_1^2 + w_2^2 - 2w_1w_2 \cos \theta]^{\frac{1}{2}}} = \left(\frac{w \cos \theta}{w_2} - \frac{w}{w_1} \right) \quad (2.5.40)$$

Vector \mathbf{n}_1 is at an angle of $(\frac{\pi}{4} + \beta_1)$ radians to the X_1 axis, and vector \mathbf{n}_2 is at an angle of $(\frac{3\pi}{4} + \beta_2)$ radians to the X_1 axis. The angle θ can be calculated from the angles β_1 and β_2

$$\begin{aligned} \theta &= \left(\frac{3\pi}{4} + \beta_2 \right) - \left(\frac{\pi}{4} + \beta_1 \right) \\ &= \left(\frac{\pi}{2} + \beta_2 - \beta_1 \right) \end{aligned} \quad (2.5.41)$$

2.5.6 Equations in Terms of the Third Grating

In order to show the correspondence between the equations describing the fringes observed with the three beam interferometer and the equations describing the fringes observed using geometric moire methods, it is necessary to write the equations governing the fringe formation, equations (2.5.32) and (2.5.33), in terms of variables associated with grating #3. These variables are the wavelength of grating #3, w , and the angle β , which is the angle between the lines of grating #3 and the X_2 axis (the same as the angle between the normal to the grating lines and the X_1 axis).

$$\beta = \psi - \frac{\pi}{2} - \left(\frac{\pi}{4} - \beta_1 \right) = \psi + \beta_1 - \frac{3\pi}{4} \quad (2.5.42)$$

From equation (2.5.41),

$$\begin{aligned} \cos \theta &= \cos \left(\frac{\pi}{2} + \beta_2 - \beta_1 \right) \\ &= -\sin(\beta_2 - \beta_1) \\ &= \sin \beta_1 \cos \beta_2 - \sin \beta_2 \cos \beta_1 \end{aligned} \quad (2.5.43)$$

$$\begin{aligned} \sin \theta &= \sin \left(\frac{\pi}{2} + \beta_2 - \beta_1 \right) \\ &= \cos(\beta_2 - \beta_1) \\ &= \cos \beta_2 \cos \beta_1 + \sin \beta_2 \sin \beta_1 \end{aligned} \quad (2.5.44)$$

From equations (2.5.42), (2.5.43) and (2.5.44),

$$\begin{aligned}
 \cos \beta &= \cos\left(\psi + \beta_1 - \frac{3\pi}{4}\right) \\
 &= \frac{1}{\sqrt{2}}\left(\sin(\psi + \beta_1) - \cos(\psi + \beta_1)\right) \\
 &= \frac{1}{\sqrt{2}}\left(\sin \psi(\cos \beta_1 + \sin \beta_1) + \cos \psi(\sin \beta_1 - \cos \beta_1)\right) \\
 &= \frac{1}{\sqrt{2}}\left(\frac{w \sin \theta}{w_2}(\cos \beta_1 + \sin \beta_1) + \frac{w \cos \theta}{w_2}(\sin \beta_1 - \cos \beta_1) \right. \\
 &\quad \left. - \frac{w}{w_1}(\sin \beta_1 - \cos \beta_1)\right) \\
 &= \frac{1}{\sqrt{2}}\left(\frac{w}{w_2} \sin(\theta + \beta_1) - \frac{w}{w_2} \cos(\theta + \beta_1) - \frac{w}{w_1}(\sin \beta_1 - \cos \beta_1)\right)
 \end{aligned} \tag{2.5.45}$$

$$\begin{aligned}
 \sin \beta &= \sin\left(\psi + \beta_1 - \frac{3\pi}{4}\right) \\
 &= \frac{1}{\sqrt{2}}\left(-\sin(\psi + \beta_1) - \cos(\psi + \beta_1)\right) \\
 &= \frac{1}{\sqrt{2}}\left(-\sin \psi(\cos \beta_1 - \sin \beta_1) - \cos \psi(\sin \beta_1 + \cos \beta_1)\right) \\
 &= \frac{1}{\sqrt{2}}\left(-\frac{w \sin \theta}{w_2}(\cos \beta_1 - \sin \beta_1) - \frac{w \cos \theta}{w_2}(\sin \beta_1 + \cos \beta_1) \right. \\
 &\quad \left. + \frac{w}{w_1}(\sin \beta_1 + \cos \beta_1)\right) \\
 &= \frac{1}{\sqrt{2}}\left(-\frac{w}{w_2} \sin(\theta + \beta_1) + \frac{w}{w_2} \cos(\theta + \beta_1) + \frac{w}{w_1}(\sin \beta_1 + \cos \beta_1)\right)
 \end{aligned} \tag{2.5.46}$$

From the formula for θ , equation (2.5.41),

$$\begin{aligned}
 \sin(\theta + \beta_1) &= \sin\left(\frac{\pi}{2} + \beta_2 - \beta_1 + \beta_1\right) \\
 &= \cos \beta_2
 \end{aligned} \tag{2.5.47}$$

$$\begin{aligned}\cos(\theta + \beta_1) &= \cos\left(\frac{\pi}{2} + \beta_2 - \beta_1 + \beta_1\right) \\ &= -\sin \beta_2\end{aligned}\tag{2.5.48}$$

From equations (2.5.45) through (2.5.48),

$$\cos \beta = \frac{1}{\sqrt{2}} \left(\frac{w}{w_2} (\cos \beta_2 + \sin \beta_2) + \frac{w}{w_1} (\cos \beta_1 - \sin \beta_1) \right)\tag{2.5.49}$$

$$\sin \beta = \frac{1}{\sqrt{2}} \left(\frac{w}{w_2} (\sin \beta_2 - \cos \beta_2) + \frac{w}{w_1} (\sin \beta_1 + \cos \beta_1) \right)\tag{2.5.50}$$

2.5.7 Initial Condition of the Third Grating

Let \tilde{w}_0 be the initial wavelength grating #3, before any deformation. From geometry (or equation (2.5.38) with $w_1 = w_2 = w_0$ and $\theta = \pi$, the initial conditions for gratings #1 and #2),

$$\tilde{w}_0 = \frac{w_0}{\sqrt{2}}\tag{2.5.51}$$

2.5.8 Result

The equations for the fringe spacing δ^* and inclination ϕ^* , equations (2.5.32) and (2.5.33), can now be rewritten in terms of variables associated with grating #3: w , \tilde{w}_0 and β . Substituting equations (2.5.49) and (2.5.50) into equations (2.5.36) and (2.5.37), it is found that

$$\begin{aligned}Z_1 &= U_1 - V_1 \\ &= \frac{\lambda}{\tilde{w}_0} - \frac{\lambda}{w} \cos \beta\end{aligned}\tag{2.5.52}$$

$$\begin{aligned} Z_2 &= U_2 - V_2 \\ &= -\frac{\lambda}{w} \sin \beta \end{aligned} \quad (2.5.53)$$

Therefore,

$$Z_1^2 + Z_2^2 = \frac{\lambda^2}{\tilde{w}_0^2} + \frac{\lambda^2}{w^2} - \frac{2\lambda^2}{w\tilde{w}_0} \cos \beta \quad (2.5.54)$$

and finally,

$$\delta^* = \frac{\lambda}{[Z_1^2 + Z_2^2]^{\frac{1}{2}}} = \frac{w\tilde{w}_0}{[w^2 + \tilde{w}_0^2 - 2w\tilde{w}_0 \cos \beta]^{\frac{1}{2}}} \quad (2.5.55)$$

$$\tan \phi^* = \frac{Z_2}{Z_1} = \frac{\tilde{w}_0 \sin \beta}{\tilde{w}_0 \cos \beta - w} \quad (2.5.56)$$

2.5.9 Discussion

Equations (2.5.55) and (2.5.56) are exactly the equations for the fringes to represent contours of constant displacement parallel to the X_1 axis with an increment of displacement between fringes of \tilde{w}_0 , as shown in Durelli and Parks [2] or Theocaris [23]. They are identical to the equations for the four beam interferometer under planar deformation except for a multiplying factor of 2 in the fringe spacing δ^* . This factor can be thought of as due to the spacing of the virtual reference grating formed by the interference between the two incoming laser beams. In the four beam interferometer, the incoming laser beams intersect at an angle of 2α , creating an interference pattern spaced at $w_0/2$, which is half the initial wavelength of the grating used of measurement. In the three beam interferometer, the incoming laser beams intersect at an angle such that the interference pattern is spaced at \tilde{w}_0 ,

which is exactly the same as the initial wavelength of the grating being measured. See the paper by Post [17], or the author [21], for further discussion of this fringe multiplication factor.

Though the three beam interferometer will measure the same quantities as the four beam interferometer, the disadvantages of the three beam interferometer as shown in Figure 23 outweigh its advantages compared to the four beam interferometer shown in Figure 24. In order to produce fringe patterns of high contrast, one requirement is that the laser beams follow paths of similar length to maintain the coherence between the two interfering light wavefronts. Both the three and four beam interferometers meet this requirement. However, high contrast fringe patterns also require that the two interfering beams have similar amplitudes and are polarized similarly. Here the three beam interferometer falls short (at least in the configuration shown in the figure). At the first beam splitter in the path of the laser, the light is divided approximately in half, with half of the light going to one of the mirrors and then to the specimen, and half of the light going to the next beam splitter. At the second beam splitter, the remaining light is again divided approximately in half, with each beam passing to a mirror and then to the specimen. Thus, one of the beams is approximately half of the strength of the original laser beam, while two of the beams are one quarter of the strength of the original laser beam. In order to match the amplitude of the beams, an additional beam splitter would have to be inserted into the path of the beam which only encounters one beam splitter. (Or the mirror directing the beam to the specimen could be replaced with a half reflecting mirror.) The intensity of each of the beams would then be one-quarter the original laser beam, in contrast to the four beam interferometer, where each of the beams is one-half the intensity of the original laser beam. The lower light levels in the three beam case would lead to longer exposure times for the fringe

recording system, which makes the experiment more sensitive to vibrations. Also, if an additional beam splitter is used, it reduces the cost and material advantage of the three beam interferometer over the four beam interferometer.

The polarization of the diffracted beams must be similar for the interference patterns to have high contrast. The three beam interferometer does not preserve the polarization of the three beams as well as the four beam interferometer. The polarization mismatch can be improved with a rotating polarizing filter on the recording camera, which forces the diffracted beams to be similarly polarized, and can be adjusted to find the best polarization angle at which this occurs. However, the addition of a polarizing filter also reduces the available light, increasing exposure time and the sensitivity to vibrations.

Another difficulty with the three beam interferometer when compared with the four beam interferometer is that it will lose information at lower displacement gradient levels. For the same fringe spacing, the diffracted light deviates from the normal to the specimen surface at a larger angle for the three beam interferometer than for the four beam interferometer. This is because the three beam interferometer uses light diffracted from two different gratings to infer displacement information about a third grating. Stretching of the two diffracting gratings can occur with only part of the resulting angular deflection of the diffracted light being related to the stretch of the third grating.

SECTION 3

Image Processing

3.1 Introduction

The fringes recorded using the moire interferometer or the Twyman-Green interferometer represent displacement contours over a field. The amount of information is somewhat staggering. Consider a 35 mm photographic negative which has a smallest spot size of about $2.5 \mu\text{m}$. The 35 mm film frame is 24 mm by 36 mm, which means there are approximately 10,000 by 15,000 pixels in the frame. If each pixel can be assigned an intensity value between 0 and 255, this is 150 megabytes of data. In contrast, a normal television screen is about 512 by 512 pixels with 0.25 megabytes of data. In order to attack this problem, the data reduction is automated to some extent, allowing a computer to trace the fringes and determine the displacement at any point in a video field. In order to do this, the computer must be able to translate from the light intensity input to the displacement data output. The computer operator controls the assignment of the data value to each fringe and oversees the fringe tracing to correct mistakes.

3.2 Hardware and Software

The system hardware consists of a stand alone Digital Equipment Corporation 11-73 microprocessor linked to 4 MB of RAM, 180 MB of hard disk, a pair of 1 MB floppy disks and a 70 MB tape drive for long term storage. The operating system is Digital Equipment Corporation RT-11 (RT stands for real time), which supports Macro-11 machine language and Fortran-77. A quad serial I/O port allows communication with a Visual 550 terminal, a C-Itoh dot matrix printer, a terminal emulator connection with the CADRE VAX cluster and the Microtek MSF 300 G digitizing scanner. Faster communication with the CADRE VAX cluster is accomplished through a Digital Equipment Corporation DEQNA parallel interface card and Ethernet network connection line controlled by Interfield Research Associates IRANET software.

Images to be analyzed are acquired through the Microtek digitizing scanner. The scanner is capable of digitizing an eight and one-half inch by fourteen inch field at 300 points per inch with eight-bit resolution (256 intensity levels) at each point. Digitization of the entire field produces 10.21 megabytes of data which can be stored on memory locally or transmitted over the Ethernet network connection to the CADRE VAX cluster memory. The operation of the scanner is controlled through software written by Guillermo Pulos.

Images may also be acquired with a black and white television camera through an Imaging Technology Incorporated Analog Processor, which contains an eight-bit flash analog-to-digital (A/D) converter. The Analog Processor can digitize the television signal at 512 by 480 pixels (picture elements) in 1/30 second, which is real time for television transmission of a frame. This results in data transfer rates of 7.5 megabytes/second, which must be accomplished over a parallel communication

line to an Imaging Technology Incorporated Frame Buffer, which stores the frame. Frames can be acquired continuously, with the Frame Buffer being overwritten every 1/30 second, or a single frame can be acquired and stored in the Frame Buffer indefinitely. Once a frame has been stored in the Frame Buffer, it can be written to the system memory, but at a significantly slower speed. Each frame contains 0.23 megabytes. There are four Frame Buffers in the system, each of which can be individually addressed by the Analog Processor, so four frames can be acquired before they must be written to the system memory. The data from the Analog Processor A/D converter passes through a programmable look-up-table before reaching the Frame Buffer; the look-up-table takes an input data value as an address and passes the value at that address along in place of the input data value. This provides the ability to adjust for contrast or brightness before the image reaches the Frame Buffer, although the information content of the image cannot be increased in this way. The Analog Processor also makes it possible for images stored in a Frame Buffer to be displayed on a color television monitor. The Analog Processor contains three eight-bit digital-to-analog (D/A) converters, one each for the red, blue and green color guns of the monitor. Again, the Analog Processor can read the Frame Buffer memory and send it to the television screen in 1/30 second. Each of the three D/A converters is preceded by a programmable look-up-table, which allows the generation of false colors on the screen; again there is no change in the information content, the color just makes it easier for humans to discriminate data levels. Although the four Frame Buffers can only be accessed singly by the Analog Processor, to the computer system they are configured as two sixteen-bit memory arrays. This is also how the Frame Buffers are accessed by the Imaging Technology Incorporated Arithmetic Logic Unit, which is a fast processor for manipulation of images stored in a Frame Buffer. These manipulations can

include additions, multiplications and convolutions by limited size kernels. Control of the image display and manipulation is through software written by Dr. Jon D. Melvin, Guillermo Pulos, Peter Washabaugh and the author.

Portions of images digitized by the Microtek scanner can be written from memory storage into a Frame Buffer for display on the monitor. Images can also be translated into a language called Postscript (Adobe Systems Incorporated) for hardcopy (black and white) output on an Apple Laserwriter. Images can also be displayed on the CADRE Vaxstation 3000 through a program written by Peter Washabaugh and then an eight color dotmatrix printer can make a color approximation of the image with a screen dump.

3.3 Fringe Tracing Approach

The fringe analysis programs used with the Moire Interferometric patterns obtained from the fracture experiments begin with a smoothing of the data to reduce noise in the pictures and equalize the light intensity variation of the fringes about a common value. The image is first low-pass filtered to extract any large scale intensity variations, such as those due to the approximately Gaussian intensity variation across the width of the laser beams. The low-pass filter simply replaces the intensity value of each pixel with the average of the intensity values over a square array of pixels centered about the pixel to be replaced. The smoothness of the result is controlled by the choice of the size of the square array. By using a very large array size, on the order of one-quarter of the the size of the image, intensity variations of very low spatial frequency can be extracted. By subtracting this low frequency variation from the original image, the fringes can be seen to oscillate

about a common mean. This subtracted image is again low-pass filtered with an array size of about three to five pixels to remove high spatial frequency noise.

Once the fringe pattern intensity has been equalized and smoothed, the fringes can be traced and assigned a displacement value. In this case, the crossings of the mean intensity value are traced rather than trying to trace the peaks or valleys of the intensity variation. These crossings correspond to the zero crossings of the phase of the of the fringe intensity and also fall on curves representing contours of constant relative displacement. The computer begins at a point and moves in an operator chosen direction, looking for a crossing of the mean intensity value (or any chosen intensity value) between two pixels. Once a crossing has been found, the pixel of higher intensity is saved and the computer sweeps the nine pixels adjacent to the saved pixel in a clockwise manner, searching for another crossing of the chosen value. The next crossing found now becomes the center for a new clockwise sweep. This continues until a boundary is reached or no crossings of the chosen data value can be found. The computer then returns to the initial saved pixel and sweeps in a counter-clockwise direction looking for crossings. In this way the entire edge of a fringe is traced and assigned an integer corresponding to the displacement. As the fringe edge is traced, the pixels discovered as crossings are stored in a buffer. This buffer can be used to undo accidental wanderings away from the correct fringe edge. Once a fringe edge has been completely traced, the computer can continue the search for edges in the direction originally chosen with the fringe counter incremented by a chosen amount.

The final result of the edge tracing is a skeleton of the original image, with each fringe edge reduced to a curve of pixels having a data value corresponding to the relative displacement of that edge with respect to the other fringe edges.

In order to extract the data for comparison with a finite element analysis, the locations of the finite element nodes are converted into screen coordinates. At each of those screen coordinates, the surrounding pixels are searched for a fringe edge. Once an edge is discovered, an approximate normal to that edge is constructed and the search for fringe edges is continued along that normal for a second fringe edge. The search is then directed backwards along the constructed normal to find a third fringe, at which another normal is constructed and used to find a fourth fringe. If all works properly, the computer has found the four fringes surrounding the desired data point, if the routine should find a boundary or screen edge, it will adapt to find any set of four (or three or two) fringes in the neighborhood of the chosen location. Having found a set of fringes, the computer will space five points along each fringe to fit a surface through. The surface fit is a polynomial with six parameters, $a_1x^2 + a_2xy + a_3y^2 + a_4x + a_5y + a_6$. Determination of the coefficients requires solution of a six by six matrix equation which is accomplished through Gaussian reduction followed by an iteration procedure to reduce any error.

For the three point bend experiment, which has a numerical model for comparison, the locations of the nodes in the finite element model are used as the locations for the determination of the displacements and slopes.

SECTION 4

Plastic Fracture

4.1 Elastic and Plastic Constitutive Relations

4.1.1 Elastic Strain

For a given current stress state σ , and strain state ϵ , it is assumed that the strain can be decomposed into an elastic part ϵ^e and a plastic part ϵ^p .

$$\epsilon = \epsilon^e + \epsilon^p \quad (4.1.1)$$

The elastic strain is related to the stress state by the usual linearly elastic constitutive law, which will be written in this case as a compliance relation between the stress and strain tensors. This relation is valid in both the elastic regime and the plastic regime, and includes the possibility of elastic unloading from a plastically strained state. In terms of components, the elastic constitutive law is written as

$$\sigma_{ij} = C_{ijkl} \epsilon_{kl}^e \quad (4.1.2)$$

where C_{ijkl} is the constant, positive definite elastic compliance four tensor for homogeneous, isotropic material behavior.

4.1.2 Plastic Strain

The plastic strain is assumed to be independent of the hydrostatic pressure, or plastically incompressible. This means that the plastic strain ϵ^p is only a function of the deviatoric stress \mathbf{S} , where \mathbf{S} is the current stress state minus the hydrostatic pressure. In components,

$$S_{ij} = \sigma_{ij} - \frac{1}{3}\sigma_{kk}\delta_{ij} \quad (4.1.3)$$

In particular, the small strain, incremental form of the plastic constitutive relation used in the numerical simulation is assumed to depend only on the second invariant (J_2) of the deviatoric stress tensor, \mathbf{S} , and the history of the loading, which determines the magnitude and direction of the plastic strain increment from the current state. The history will be contained in a parameter $\bar{\epsilon}^p$, called the accumulated equivalent plastic strain. The invariant J_2 is found to be

$$J_2 = \frac{1}{2}S_{ij}S_{ij} \quad (4.1.4)$$

The incremental form of the plastic constitutive relation will be written in rate form, with time as a parameter. The numerical code uses a specified load history and time steps to produce load steps. Thus, the elastic and plastic strain increments will be converted to strain rates, $\dot{\epsilon}^e$ and $\dot{\epsilon}^p$.

The rate of change of the history parameter, the accumulated equivalent plastic strain $\bar{\epsilon}^p$, will be defined by

$$\frac{d\bar{\epsilon}^p}{dt} = \left(\frac{2}{3}\dot{\epsilon}_{ij}^p\dot{\epsilon}_{ij}^p\right)^{\frac{1}{2}} \quad (4.1.5)$$

from which it follows that the current accumulated equivalent plastic strain, $\bar{\epsilon}^p$ is described by an integral over the loading history:

$$\bar{\epsilon}^p = \int \left(\frac{2}{3}\dot{\epsilon}_{ij}^p\dot{\epsilon}_{ij}^p\right)^{\frac{1}{2}} dt \quad (4.1.6)$$

4.1.3 Yield Surface — Condition for Plastic Deformation

The yield surface is a function of the current stress and strain states, $f(\boldsymbol{\sigma}, \boldsymbol{\epsilon})$, such that the condition for plastic deformation to occur is that $f(\boldsymbol{\sigma}, \boldsymbol{\epsilon}) = 0$. If the stress and strain states are such that $f(\boldsymbol{\sigma}, \boldsymbol{\epsilon}) < 0$, the deformation is purely elastic. The yield surface moves during plastic deformation so that it is impossible for the stress and strain to reach a state where $f(\boldsymbol{\sigma}, \boldsymbol{\epsilon}) > 0$. The yield surface follows the Huber-Von Mises criterion, making it a sphere in deviatoric stress space, specifically,

$$f(\boldsymbol{\sigma}, \boldsymbol{\epsilon}) = 3J_2 - \bar{\sigma}^2(\bar{\epsilon}^p) \quad (4.1.7)$$

The function $\bar{\sigma}$, which is a function of the accumulated equivalent plastic strain $\bar{\epsilon}^p$, ($\bar{\sigma} = \bar{\sigma}(\bar{\epsilon}^p)$) is assumed to follow a power law of the form:

$$\frac{\bar{\epsilon}^p}{\epsilon_0} = \left(\frac{\bar{\sigma}}{\sigma_0} \right)^n - \frac{\bar{\sigma}}{\sigma_0} \quad (4.1.8)$$

Here σ_0 is the yield stress in uniaxial tension and ϵ_0 is the corresponding strain at yield in uniaxial tension.

Again, for elastic loading or unloading, the stress state must be such that $f(\boldsymbol{\sigma}, \boldsymbol{\epsilon}) < 0$, with plastic deformation taking place when $f(\boldsymbol{\sigma}, \boldsymbol{\epsilon}) = 0$. The yield surface expands during plastic loading so that the current stress state must always lie on or inside it.

4.1.4 Normality

The J_2 incremental theory assumes that any plastic strain increment is normal to the yield surface. This assumption yields the result that the plastic strain rate is proportional to the rate of change of the deviatoric stress invariant, J_2 , and is parallel to the deviatoric stress tensor itself, since the yield surface is a sphere in deviatoric space. The magnitude of the plastic strain rate is contained in a parameter H , which is a measure of the hardening behavior of the material. Therefore, while plastic strain is occurring, the plastic strain rate can be written as:

$$\dot{\epsilon}_{ij}^p = \frac{3}{4H J_2} S_{ij} \frac{dJ_2}{dt} = \dot{\lambda} S_{ij} \quad (4.1.9)$$

where $\dot{\lambda}$ is a parameter containing the rate and hardening terms.

4.1.5 Hardening

Using equations (4.1.5) and (4.1.9), the rate of change of the accumulated, equivalent plastic strain is found to be

$$\frac{d\bar{\epsilon}^p}{dt} = \frac{3}{2H} \frac{dJ_2}{dt} \left(\frac{1}{3J_2} \right)^{\frac{1}{2}} \quad (4.1.10)$$

However, during plastic deformation, the function describing the yield surface, $f(\sigma, \epsilon)$, is equal to zero. Therefore,

$$\bar{\sigma}^2 = 3J_2 \quad (4.1.11)$$

$$\bar{\sigma} = (3J_2)^{\frac{1}{2}} \quad (4.1.12)$$

$$\frac{d\bar{\sigma}}{dt} = \frac{3}{2} \frac{dJ_2}{dt} \left(\frac{1}{3J_2} \right)^{\frac{1}{2}} \quad (4.1.13)$$

Substituting this relation into equation (4.1.10),

$$\frac{d\bar{\epsilon}^p}{dt} = \frac{1}{H} \frac{d\bar{\sigma}}{dt} \quad (4.1.14)$$

from which it can be seen that H is the rate of change of function $\bar{\sigma}$ with respect to the accumulated equivalent plastic strain, $\bar{\epsilon}^p$,

$$H = \frac{d\bar{\sigma}}{d\bar{\epsilon}^p} \quad (4.1.15)$$

The function H can be derived explicitly from equation (4.1.8), and is found to be

$$H = \frac{d\bar{\sigma}}{d\bar{\epsilon}^p} = \frac{\sigma_0}{\epsilon_0 \left[n \left(\frac{\bar{\sigma}}{\sigma_0} \right)^{n-1} - 1 \right]} \quad (4.1.16)$$

4.1.6 Constitutive Relation for the Total Strain Rate

The total strain rate $\dot{\epsilon}$ can be related to the stress rate $\dot{\sigma}$ by combining the plastic and elastic parts. Writing the elastic stress-strain relation in rate form,

$$C_{ijkl} \dot{\epsilon}_{kl}^e = \dot{\sigma}_{ij} \quad (4.1.17)$$

Using the rate form of the plastic strain relation, equation (4.1.9), and operating on both sides with the constant elastic compliance tensor C_{ijkl} ,

$$C_{ijkl} \dot{\epsilon}_{kl}^p = \frac{3}{4HJ_2} C_{ijkl} S_{kl} \frac{dJ_2}{dt} \quad (4.1.18)$$

Adding equations (4.1.17) and (4.1.18), noting that $\dot{\epsilon} = \dot{\epsilon}^e + \dot{\epsilon}^p$,

$$C_{ijkl} \dot{\epsilon}_{kl} = \dot{\sigma}_{ij} + \frac{3}{4HJ_2} C_{ijkl} S_{kl} \frac{dJ_2}{dt} \quad (4.1.19)$$

From the definition of J_2 , equation (4.1.4),

$$\begin{aligned}
 \frac{dJ_2}{dt} &= \frac{1}{2} \frac{d(S_{mn}S_{mn})}{dt} = S_{mn} \frac{dS_{mn}}{dt} \\
 &= S_{mn} \frac{d}{dt} (\sigma_{mn} - \frac{1}{3} \delta_{mn} d\sigma_{kk}) \\
 &= S_{mn} \dot{\sigma}_{mn} - \frac{1}{3} S_{mn} \delta_{mn} \frac{d\sigma_{kk}}{dt} \\
 &= S_{mn} \dot{\sigma}_{mn}
 \end{aligned} \tag{4.1.20}$$

since

$$S_{mn} \delta_{mn} = 0 \tag{4.1.21}$$

Replacing dJ_2/dt with $S_{mn} \dot{\sigma}_{mn}$ in equation (4.1.19),

$$C_{ijkl} \dot{\epsilon}_{kl} = \dot{\sigma}_{ij} + \frac{3}{4HJ_2} C_{ijkl} S_{kl} S_{mn} \dot{\sigma}_{mn} \tag{4.1.22}$$

Using the rate form for the elastic strain, equation (4.1.17),

$$C_{ijkl} \dot{\epsilon}_{kl}^e = \dot{\sigma}_{ij} + \frac{3}{4HJ_2} C_{ijkl} S_{kl} S_{mn} C_{mnpq} \dot{\epsilon}_{pq}^e \tag{4.1.23}$$

Equation (4.1.18), the rate form of the plastic strain can also be rewritten using the substitution for dJ_2/dt of $S_{mn} \dot{\sigma}_{mn}$.

$$C_{ijkl} \dot{\epsilon}_{kl}^p = \frac{3}{4HJ_2} C_{ijkl} S_{kl} S_{mn} C_{mnpq} \dot{\epsilon}_{pq}^e \tag{4.1.24}$$

Operating on both sides of equation (4.1.24) with the deviatoric stress S_{ij} yields a form which can be used to transform the dependence of equation (4.1.23) on the elastic strain rate $\dot{\epsilon}^e$ into a dependence on the total strain rate $\dot{\epsilon}$.

$$S_{ij} C_{ijkl} \dot{\epsilon}_{kl}^p = \frac{3}{4HJ_2} S_{ij} C_{ijkl} S_{kl} S_{mn} C_{mnpq} \dot{\epsilon}_{pq}^e \tag{4.1.25}$$

Since all indices are summed out, the indices $ijkl$ may be replaced on the left by the indices $mnpq$, respectively.

$$S_{mn} C_{mnpq} \dot{\epsilon}_{pq}^p = \frac{3}{4HJ_2} S_{ij} C_{ijkl} S_{kl} S_{mn} C_{mnpq} \dot{\epsilon}_{pq}^e \tag{4.1.26}$$

Adding $S_{mn}C_{mnpq}\dot{\epsilon}_{pq}^e$ to both sides of equation (4.1.26) gives the total strain rate on the left-hand side:

$$S_{mn}C_{mnpq}\dot{\epsilon}_{pq} = \left(\frac{3}{4HJ_2} S_{ij}C_{ijkl}S_{kl} + 1 \right) S_{mn}C_{mnpq}\dot{\epsilon}_{pq}^e \quad (4.1.27)$$

$$\frac{3S_{mn}C_{mnpq}\dot{\epsilon}_{pq}^e}{4HJ_2} = \frac{S_{mn}C_{mnpq}\dot{\epsilon}_{pq}}{\frac{4}{3}HJ_2 + S_{ij}C_{ijkl}S_{kl}} \quad (4.1.28)$$

The relation between the strain rate and the stress rate, equation (4.1.22), can now be written as

$$C_{ijkl}\dot{\epsilon}_{kl} = \dot{\sigma}_{ij} + \frac{1}{\frac{4}{3}HJ_2 + S_{rt}C_{rtuv}S_{uv}} C_{ijpq}S_{pq}S_{mn}C_{mnkl}\dot{\epsilon}_{kl} \quad (4.1.29)$$

Grouping the strain terms on the right, the total strain rate can be related to the stress rate by

$$\dot{\sigma}_{ij} = \left[C_{ijkl} - \frac{C_{ijpq}S_{pq}S_{mn}C_{mnkl}}{S_{rt}C_{rtuv}S_{uv} + \frac{4}{3}HJ_2} \right] \dot{\epsilon}_{kl} \quad (4.1.30)$$

During plastic deformation, the function describing the yield surface, $f(\sigma, \epsilon)$ is identically zero. From equation (4.1.7), it is found that

$$J_2 = \frac{1}{3}\bar{\sigma}^2(\bar{\epsilon}^p) \quad (4.1.31)$$

Replacing the quantity J_2 in equation (4.1.30) with the relation in equation (4.1.31),

$$\dot{\sigma}_{ij} = \left[C_{ijkl} - \frac{C_{ijpq}S_{pq}S_{mn}C_{mnkl}}{S_{rt}C_{rtuv}S_{uv} + \frac{4}{9}\bar{\sigma}^2 H} \right] \dot{\epsilon}_{kl} = C_{ijkl}^* \dot{\epsilon}_{kl} \quad (4.1.32)$$

The quantity C_{ijkl}^* in equation (4.1.32) is the history dependent compliance tensor used to relate the stress and strain rates during plastic deformation. During elastic deformation the compliance C_{ijkl}^* reduces to the constant elastic compliance C_{ijkl} .

4.1.7 Uniaxial Test to Determine Hardening Behavior

The yield stress, strain at yield and hardening exponent n must all be determined through a uniaxial test. The procedure for determining the hardening exponent will be described here. In the uniaxial test, the stress and strain are measured along one axis, in this case the x_1 axis. It is assumed that all stress components are zero except the tensile stress in the x_1 direction, so that $\sigma_{11} = \sigma_{11}(t)$. The stress tensor has the form

$$\boldsymbol{\sigma} = \sigma_{11}(t) \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \quad (4.1.33)$$

The deviatoric stress is also a constant tensor multiplied by $\sigma_{11}(t)$.

$$\mathbf{S} = \sigma_{11}(t) \begin{pmatrix} 2/3 & 0 & 0 \\ 0 & -1/3 & 0 \\ 0 & 0 & -1/3 \end{pmatrix} \quad (4.1.34)$$

The second invariant of the deviatoric stress, J_2 , is thus $\frac{1}{3}\sigma_{11}^2(t)$. The condition for plastic yielding, $f(\boldsymbol{\sigma}, \boldsymbol{\epsilon}) = 0$ implies that

$$\bar{\sigma}^2(\bar{\epsilon}^p) = \sigma_{11}^2(t) \quad (4.1.35)$$

Since both $\sigma_{11}(t)$ and $\bar{\sigma}$ are positive for this case, this is equivalent to

$$\sigma_{11}(t) = \bar{\sigma}(\bar{\epsilon}^p) \quad (4.1.36)$$

In proportional loading conditions such as the uniaxial test, where the deviatoric stress can be described by a scalar function of load multiplying a constant tensor, the plastic strain is also proportional to that constant tensor, as can be seen from equation (4.1.9). From that equation, it is found that

$$\dot{\boldsymbol{\epsilon}}^p = \dot{\lambda} \sigma_{11}(t) \begin{pmatrix} 2/3 & 0 & 0 \\ 0 & -1/3 & 0 \\ 0 & 0 & -1/3 \end{pmatrix} \quad (4.1.37)$$

The measured strain is $\epsilon_{11}(t)$. The plastic part of the measured strain is $\epsilon_{11}^p(t)$ and in terms of strain rate, $\dot{\epsilon}_{11}^p(t) = \frac{2}{3}\dot{\lambda}\sigma_{11}(t)$. Rewriting equation (4.1.37) in terms of $\dot{\epsilon}_{11}^p(t)$ leads to the result

$$\dot{\epsilon}^p = \dot{\epsilon}_{11}^p(t) \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1/2 & 0 \\ 0 & 0 & -1/2 \end{pmatrix} \quad (4.1.38)$$

Using equation (4.1.5), the rate of change of the accumulated equivalent plastic strain, $\bar{\epsilon}$, is given by

$$\frac{d\bar{\epsilon}^p}{dt} = \left(\frac{2}{3} \dot{\epsilon}_{ij}^p \dot{\epsilon}_{ij}^p \right)^{1/2} = \dot{\epsilon}_{11}^p(t) \quad (4.1.39)$$

Integrating over the loading history,

$$\bar{\epsilon}^p = \epsilon_{11}^p(t) \quad (4.1.40)$$

Substituting equations (4.1.36) and (4.1.40) relating the equivalent accumulated plastic strain $\bar{\epsilon}^p$ to the plastic strain $\epsilon_{11}^p(t)$ and the function $\bar{\sigma}(\bar{\epsilon}^p)$ to the measured stress $\sigma_{11}(t)$ into equation (4.1.8), the power law relating $\bar{\epsilon}^p$ and $\bar{\sigma}(\bar{\epsilon}^p)$,

$$\frac{\epsilon_{11}^p(t)}{\epsilon_0} = \left(\frac{\sigma_{11}(t)}{\sigma_0} \right)^n - \frac{\sigma_{11}(t)}{\sigma_0} \quad (4.1.41)$$

However, $\sigma_{11}(t)/\sigma_0$ is the elastic part of $\epsilon_{11}(t)$ divided by ϵ_0 , the strain at yield. Therefore, the measured strain $\epsilon_{11}(t)$ is related to the measured stress $\sigma_{11}(t)$ by the power law

$$\frac{\epsilon_{11}}{\epsilon_0} = \left(\frac{\sigma_{11}}{\sigma_0} \right)^n \quad (4.1.42)$$

This equation is fit to the measured data from a uniaxial test to provide the parameter n for the numerical simulation. The yield stress and strain at yield are also determined from the measured data.

4.2 Uniaxial Test Results

The uniaxial tests were done to ensure that a meaningful comparison could be made between the measurements made for this study, the measurements by Zehnder [27,28] and the numerical simulation of Zehnder's experiment by Narasimhan [15]. Although the specifications for the steel and heat treatment of the specimens matched that of Zehnder, and the steel was purchased from the same source, it would be unrealistic to expect that the constitutive relations would be identical. Consequently, it was hoped that the uniaxial tests would indicate that the different studies matched closely enough for the various studies to be compared directly, which was the original reason for choosing the specimen geometry, material and heat treatment. As an alternative, if the material properties failed to match closely, it was hoped that the numerical simulation might be redone with material properties generated from the current study.

4.2.1 Testing Procedure

The uniaxial tests were done in the Baldwin hydraulic loading machine (capacity 1,000,000 N) which does not allow precise control of either the force or displacement history. The tests were done with the loading valve slightly open to give a slowly increasing load to failure of the specimen. The overall strain rate was less than 0.5%/sec in each case. The load was read from a SENSOTEC 400,000 N capacity load cell and the strain was read from two Instron strain gage extensometers, a 10% extensometer with a 2 inch gauge length and a 100% extensometer with a 1 inch gauge length. All three signals were recorded on a Masscomp computer data acquisition and control system using Laboratory Workbench software under a Unix operating system. The data are shown in Figures 30 through 33. Two specimens

were tested with identical geometry and heat treatment to check the repeatability of the measurement. Figures 30 and 31 show the measurements done with the 10 % extensometer, which was released from each specimen before it reached the end of its travel.

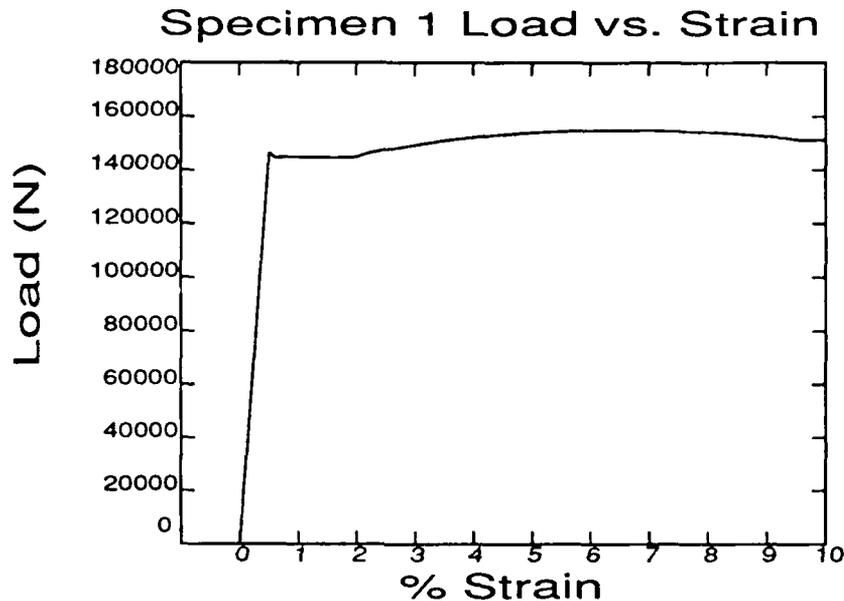


FIGURE 30. Load versus Strain history for uniaxial test specimen number 1. Load in Newtons. The 10 % extensometer was freed from the specimen at 9.5 % Strain.

Figures 32 and 33 show the data from the 100% extensometer, which shows the effect of necking beyond about 8 %. Obviously, when the specimen has begun to neck, the stress and strain states are no longer homogeneous throughout the specimen, and the extensometer measurement no longer represents the actual lengthening of the specimen material.

The Laboratory Workbench software allows for the inclusion of multiplying factors and offset voltages during the recording of the data, so the calibrations of the load cell and extensometers could be used to convert the voltages to load and dis-

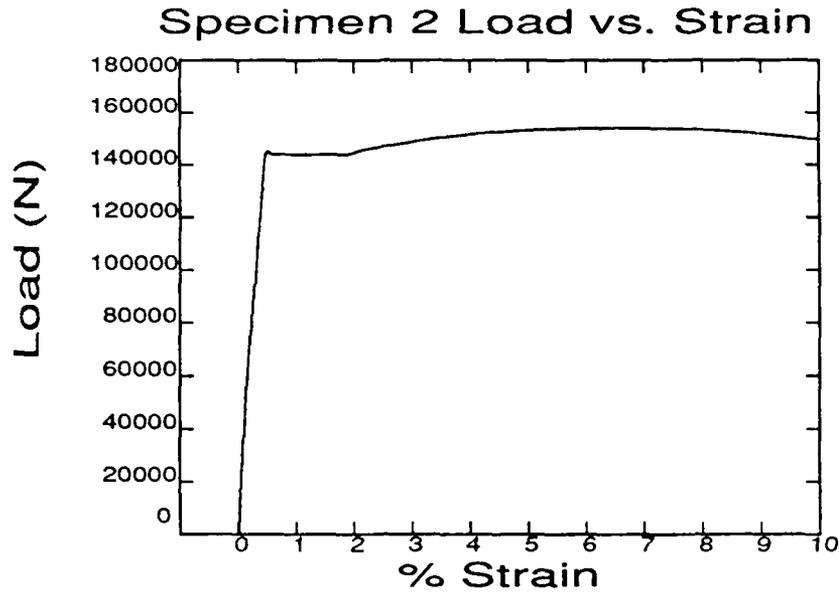


FIGURE 31. Load versus Strain history for uniaxial test specimen number 2. Load is in Newtons. Specimen number 2 was identical to specimen number 1. The 10 % extensometer was freed from the specimen at 9.5 % Strain.

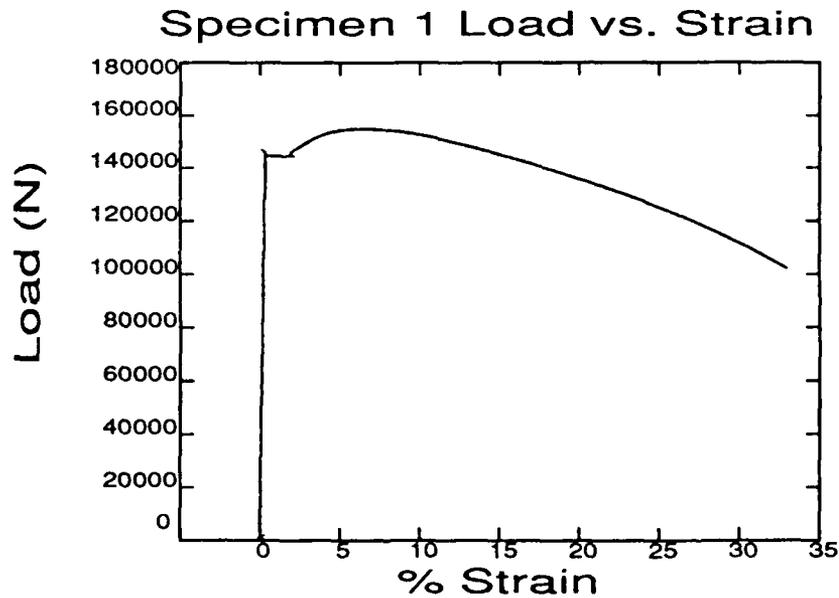


FIGURE 32. Load versus Strain history for uniaxial test specimen number 1. Load in Newtons. The 100 % extensometer shows the effect of necking. The specimen failed at about 34 % Strain.

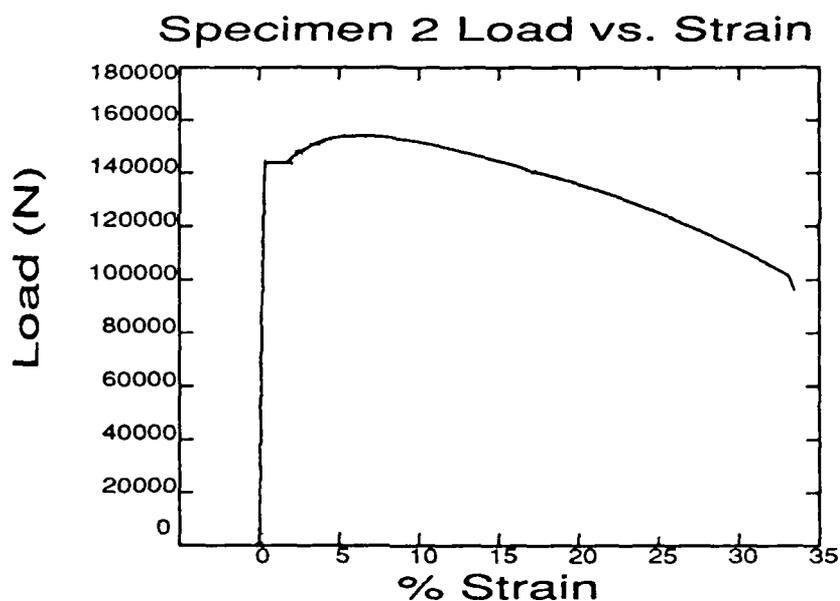


FIGURE 33. Load versus Strain history for uniaxial test specimen number 2. Load is in Newtons. Specimen number 2 was identical to specimen number 1. The 100 % extensometer shows necking. The specimen failed at about 33 % Strain.

placement units during the experiment. Also, the gauge length of each extensometer could be divided out to give the engineering strain $e = \Delta l/l_0$. The engineering stress S was then calculated from the recorded load L and the initial cross sectional area, A as $S = L/A$. When the specimen is deforming plastically, but before the onset of necking, the assumption of plastic incompressibility leads to the relation that the current area $A^* = A/(1 + e)$. Thus the true stress σ is calculated by $\sigma = S(1 + e)$. The true strain ϵ is calculated by noting that $d\epsilon = dl/l$. Integrating this quantity from the initial length l_0 to the current length l gives $\epsilon = \ln(l/l_0) = \ln(1 + e)$. The data in this form are shown in Figures 34 and 35 for specimens 1 and 2. The data from both specimens is plotted on the same graph in Figure 36, showing the repeatability of the measurement.

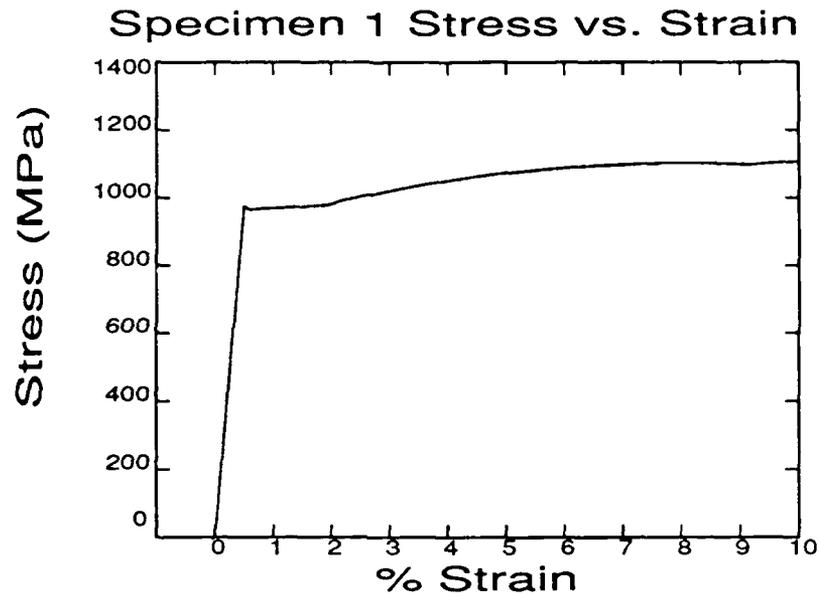


FIGURE 34. Stress versus True Strain for specimen number 1. Strain data from the 10 % extensometer. Stress includes correction for plastic incompressibility.

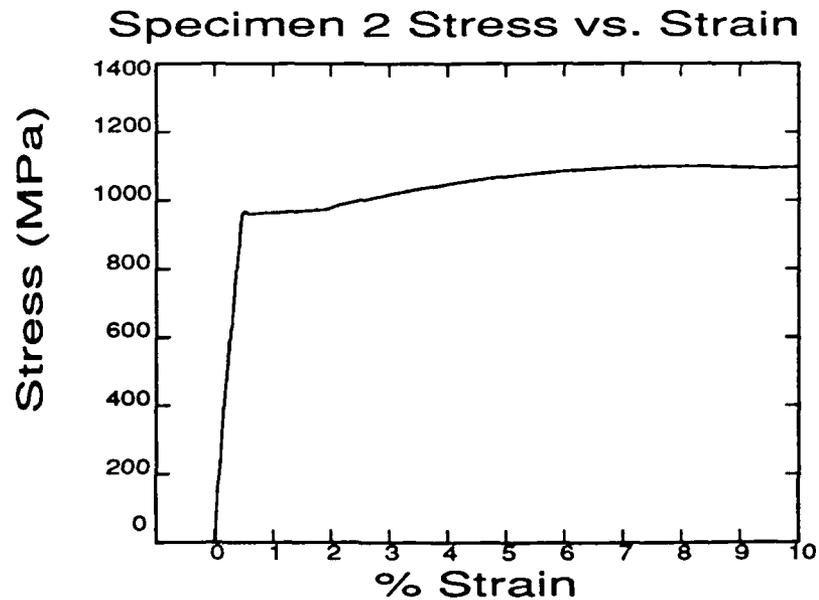


FIGURE 35. Stress versus True Strain for specimen number 2. Strain data from the 10 % extensometer. Stress includes correction for plastic incompressibility.

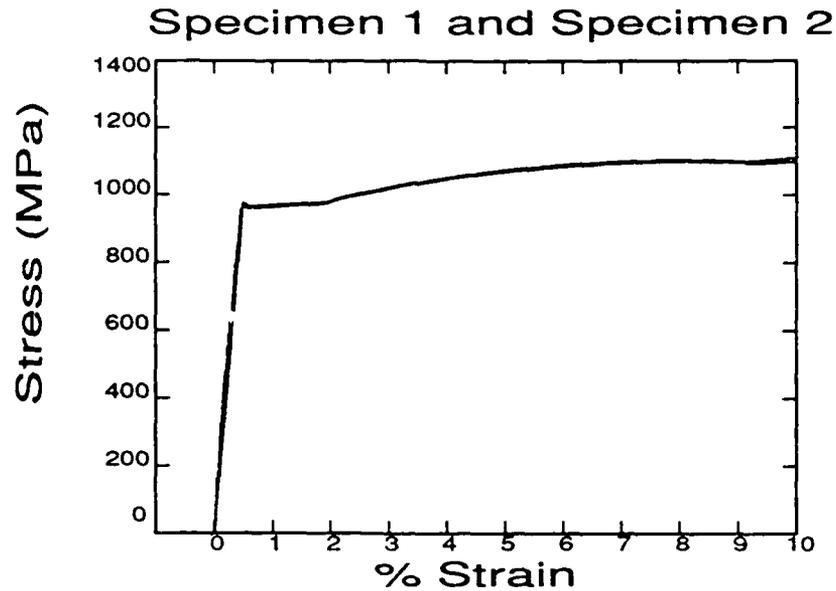


FIGURE 36. Stress versus True Strain for specimens 1 and 2 on the same figure. This figure shows the repeatability of the measurement procedure.

4.2.2 Comparison and Discussion

The uniaxial constitutive relation for input to the numerical model is expressed in the following way:

$$\frac{\epsilon}{\epsilon_0} = \begin{cases} \frac{\sigma}{\sigma_0} & \sigma < \sigma_0 \\ \left(\frac{\sigma}{\sigma_0}\right)^n & \sigma \geq \sigma_0 \end{cases}$$

where σ_0 is the yield stress, ϵ_0 is the corresponding strain at yield and n is the hardening exponent. Young's Modulus is calculated as σ_0/ϵ_0 .

The yield stress σ_0 and the strain at yield ϵ_0 were determined from the stress-strain diagrams of Figure 36, with $\sigma_0 = 960.0\text{MPa}$ and $\epsilon_0 = 0.500\%$. Young's Modulus is therefore found to be 192.0 GPa . A power law of the form above was least squares fit minimizing the error in stress as a function of strain for the stress-strain data from yield to a strain of 8% , where necking can be seen to initiate. The hardening exponent n is calculated to be $n = 23.7$. The resulting stress-strain law

is shown in Figure 37. The power law is compared to the experimentally measured stress-strain law in Figure 38.

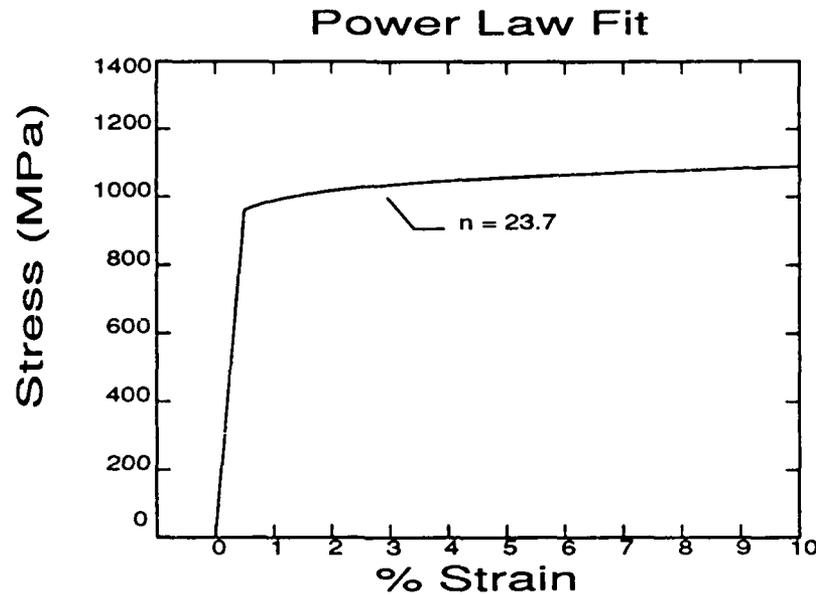


FIGURE 37. Power Law Hardening Law least squares fit to the data below 8 % Strain from Figure 36. The measured hardening exponent was 23.7.

Although a different law could fit the experiment better, the power law is useful because of its ease of implementation in the finite element code and because of the existence of the HRR analytical asymptotic solution for comparison. (A very good fit would be provided by a law consisting of two linear regions and then a power law region.)

A comparison of the uniaxial stress strain curves from Zehnder's experiment and the current experiment is shown in Figure 39. It can be seen from the figure that the yield stresses differ by about 10% between the two cases, but the plastic hardening behavior is similar. This indicates that if the two studies could be normalized with respect to the yield stress, the comparison of the plastic deformations

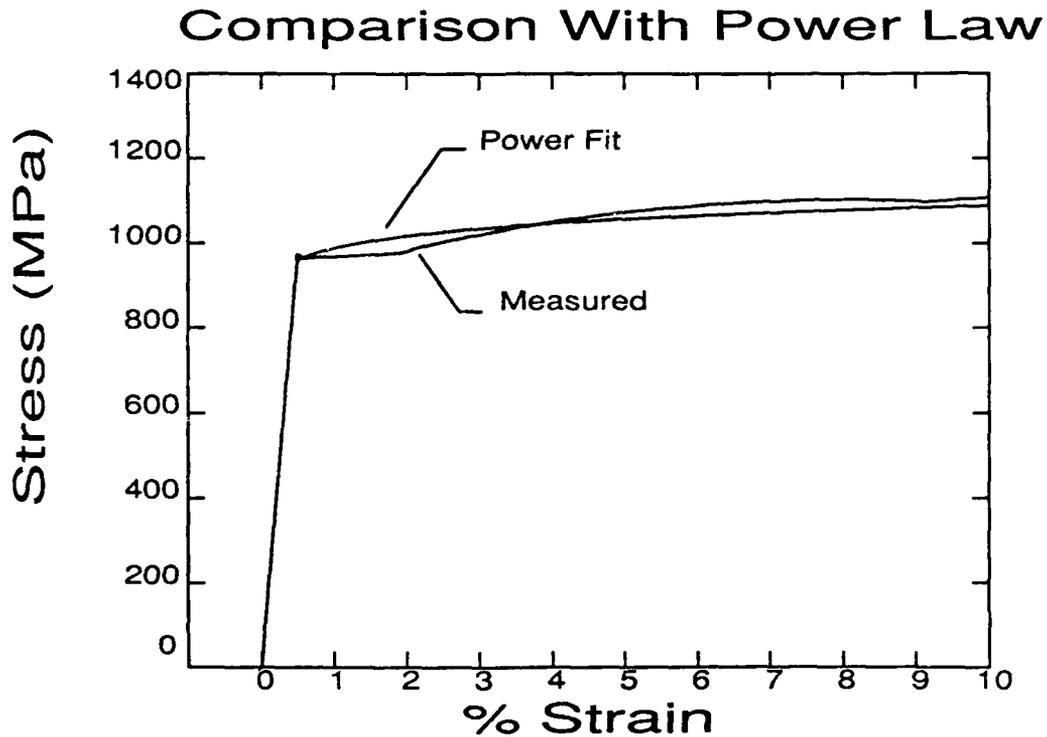


FIGURE 38. Comparison of measured Stress-Strain relation with Power Law Fit. Although a fit of a different law might be better, the power law is used because of its simplicity for input to the finite element calculation, as well as the existence of the HRR asymptotic analytical solution for comparison.

should be quite close. The data from the two cases is shown in Figure 40 with Zehnder's data normalized to the current study by the yield stresses.

In the original numerical simulation, the Young's Modulus was 199.2 GPa while the yield stress σ_0 and strain at yield ϵ_0 were determined to be $\sigma_0 = 1067\text{MPa}$ and $\epsilon_0 = 0.536\%$. The hardening exponent n was found to be $n = 22.0$.

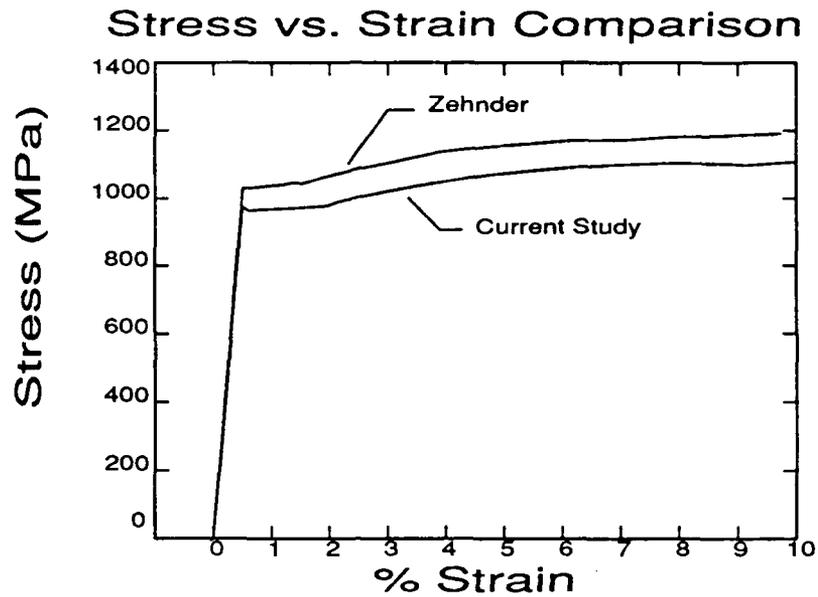


FIGURE 39. Comparison of Stress-Strain relation for the current study and the identical study by Zehnder. The yield stress differs by about 10 %.

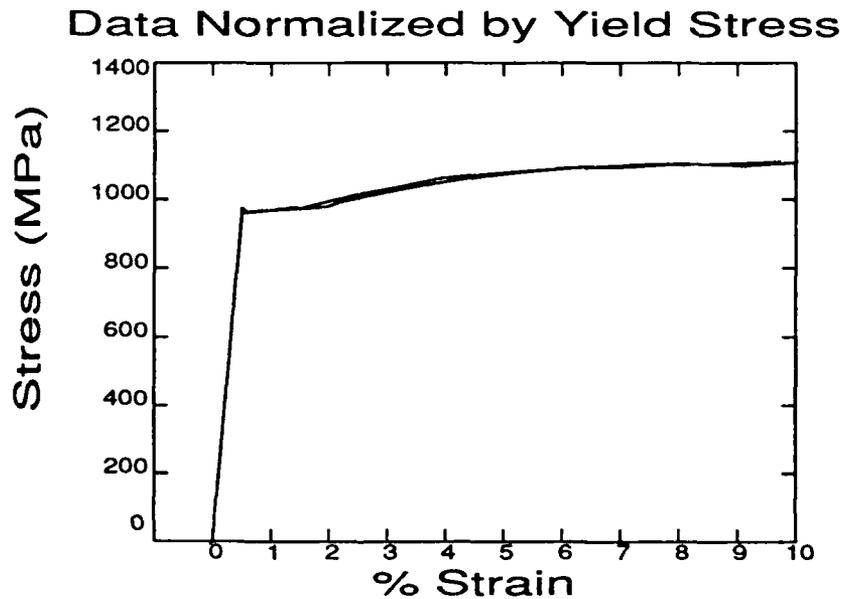


FIGURE 40. Comparison of Stress-Strain relation for the current study and the study by Zehnder, with Zehnder's data normalized by the yield stress. This suggests that comparisons between studies should be normalized by the yield stress.

4.3 Finite Element Analysis

4.3.1 Model Description

The three point bend experiment of Figure 41 was modeled using a three dimensional finite element code with a power hardening elastic-plastic constitutive relation. Due to the symmetry about the plane of the crack and about the midplane of the specimen, only one quarter of the specimen must be modeled.

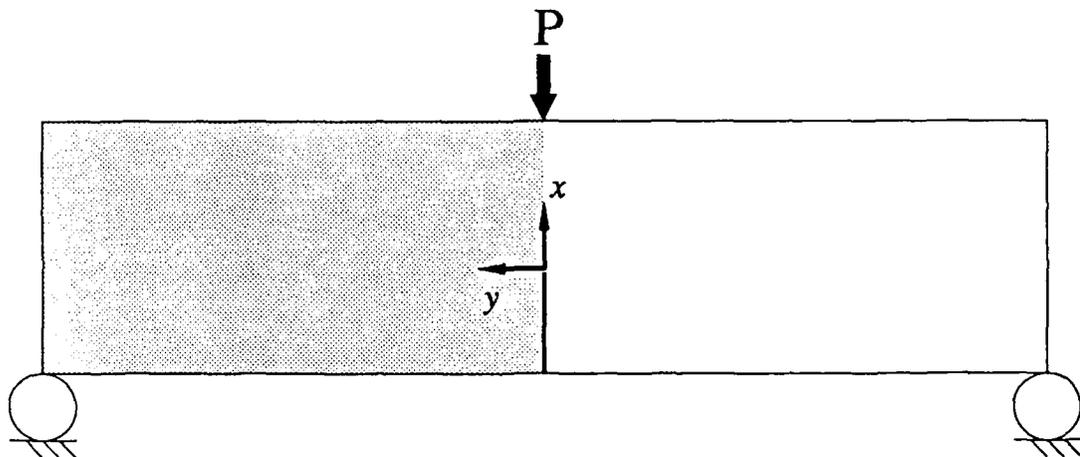


FIGURE 41. Diagram of the geometry of the three point bend experiment. Due to symmetry, only the shaded side of the specimen needs to be modeled. Also, symmetry about the midplane of the specimen means only one quarter of the entire specimen must be modeled.

Two different numerical simulations of the three point bend experiments were performed using a three dimensional code with a small strain, incremental plasticity law; both of the models were constructed by Dr. R. Narasimhan of the India Institute of Technology. The first model was designed to simulate the experiments by Zehnder and Rosakis [27,28]. The specimen modeled was a plate of 30.0 cm by

7.5 cm by 0.975 cm thick with a crack 3.0 cm in length centered in the longest side. The elastic modulus, yield stress and power hardening exponent were calculated from a uniaxial tension test. In this case, the crack was modeled as having a sharp tip. This model consisted of 2100 eight noded block elements arranged in 5 layers and having 2814 nodes and 8442 degrees of freedom. The mesh for this model is shown in Figures 42, 43 and 44. The elements become smaller near the crack tip, and around the tip the elements are arranged in eighteen wedges of 10° each. At the tip, the nodes at two adjacent corners of the elements coincide to produce the sharp crack tip. The element size varies through the thickness as well, with the elements being smaller near the free surface to capture the corner singularity.

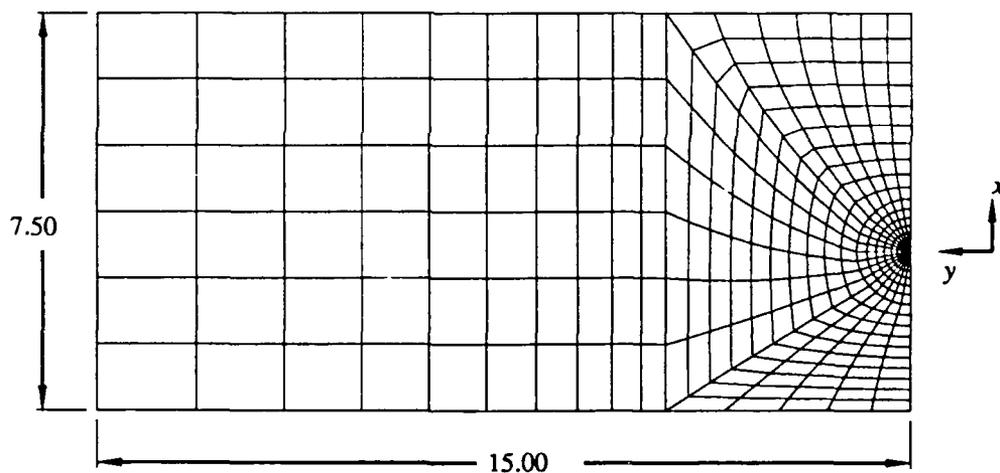


FIGURE 42. Diagram of mesh used in the original three point bend model. The crack is at the lower right, and is modeled as having a sharp tip. The elements are concentrated near the crack tip. All dimensions in centimeters.

The existence of the previous work by Zehnder and Rosakis and by Narasimhan and Rosakis was the basis for choosing the specimen geometry for the current study, so that direct comparisons could be made. However, the data available from the

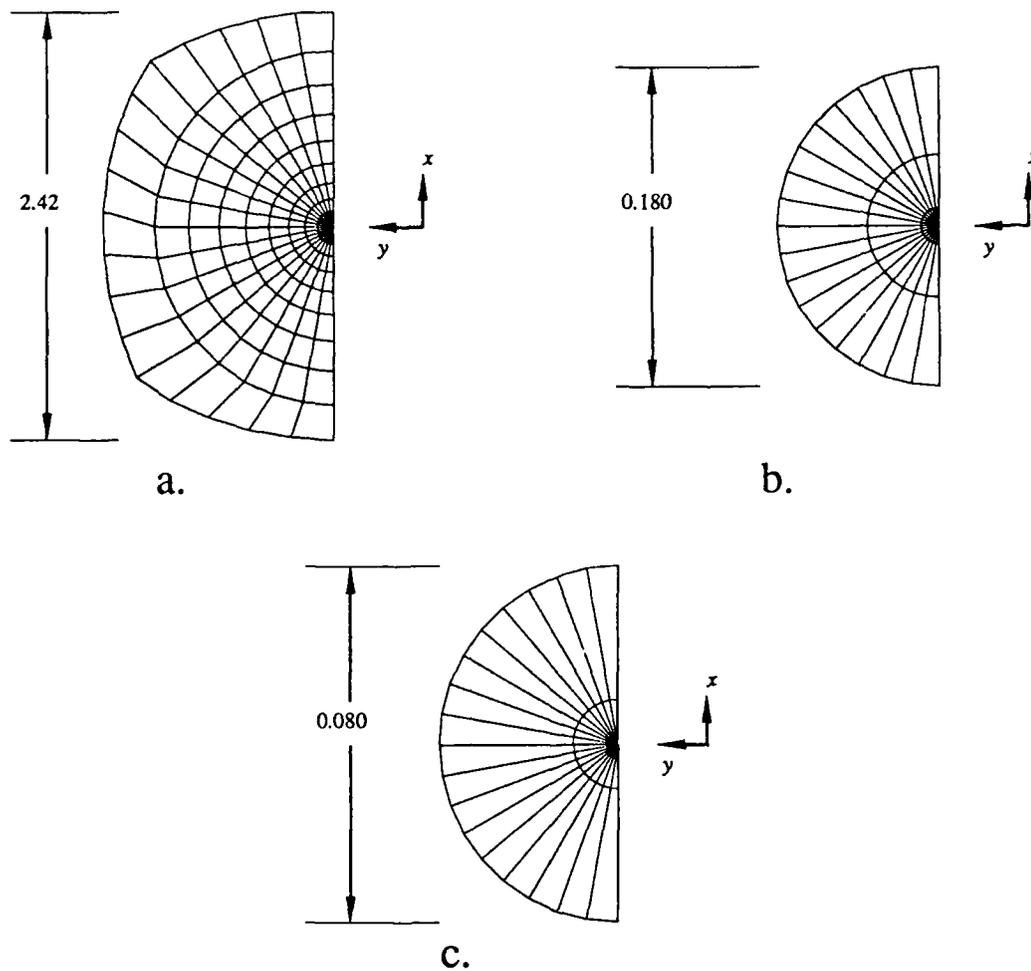


FIGURE 43. Detail of the mesh near the crack tip in the original three point bend model. The crack is modeled as having a sharp tip. All dimensions in centimeters. Diagram c shows the two smallest rings of the block elements around the crack tip, each element in the smallest ring has two coincident nodes at the crack tip.

finite element model was somewhat sparse, as displacement data was only saved for five load steps: 24 N, which serves to characterize the linearly elastic response, 35000 N, 52296 N, 68256 N, and 78863 N, which was approximately the failure load of the corresponding experiment. The lack of information available from loads between 24 N and 35000 N, which is where plastic behavior initiates and where the

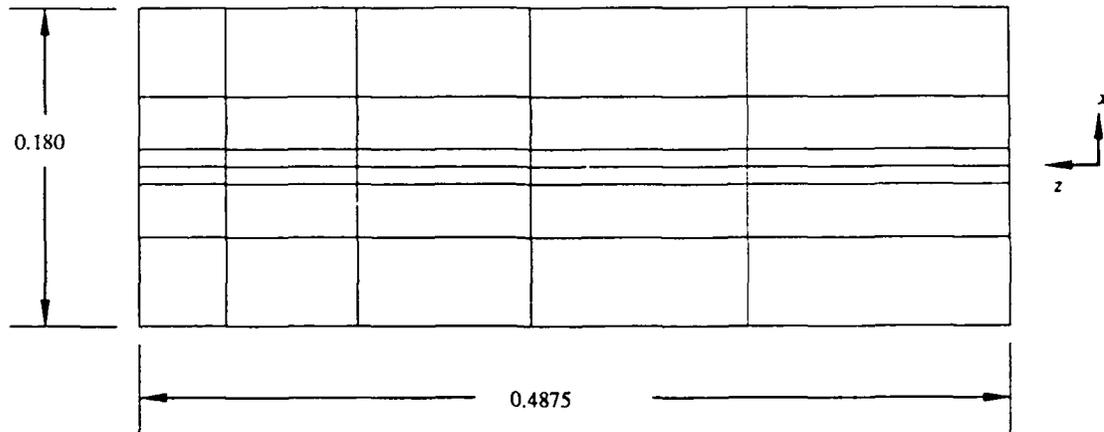


FIGURE 44. Detail of the element variation through the thickness of the original three point bend model. The crack tip is in the center of the diagram, running horizontally. The scale corresponds to diagram b in Figure 43. The free surface is at the left, the midplane at the right. All dimensions in centimeters.

use of the moire interferometer is likely to be most fruitful, made it desirable to run a second simulation of the experiment. Additionally, the material properties of the two experiments were different, most noticeable in the yield stresses, which differed by about 10 %. A second model of the three point bend specimen had been constructed by Dr. Narasimhan to study the growth of voids using a Gurson theory. This model was adapted by Dr. Narasimhan and the author to duplicate the material properties of the current study. The second model was similar to the first: a plate of 30.0 cm by 7.6 cm by 0.975 cm thick with a crack 3.0 cm in length centered in the longest side. However, the second model does represent the physical specimen better in that it contains a notch with a circular tip of radius 0.01 cm, rather than a sharp crack. (The steel plate in the experiment is cut with an electrical discharge, which results in a notch which is about 0.03 cm across.) In this case, the model consisted of 1920 eight noded block elements arranged in 6 layers and having 2527 nodes and 7581 degrees of freedom. Again the element size varied with

distance from the crack tip and with distance from the free surface. Near the tip, the elements are arranged in wedges of 15° each. The mesh geometry of the second simulation is shown in Figures 45 through 47.

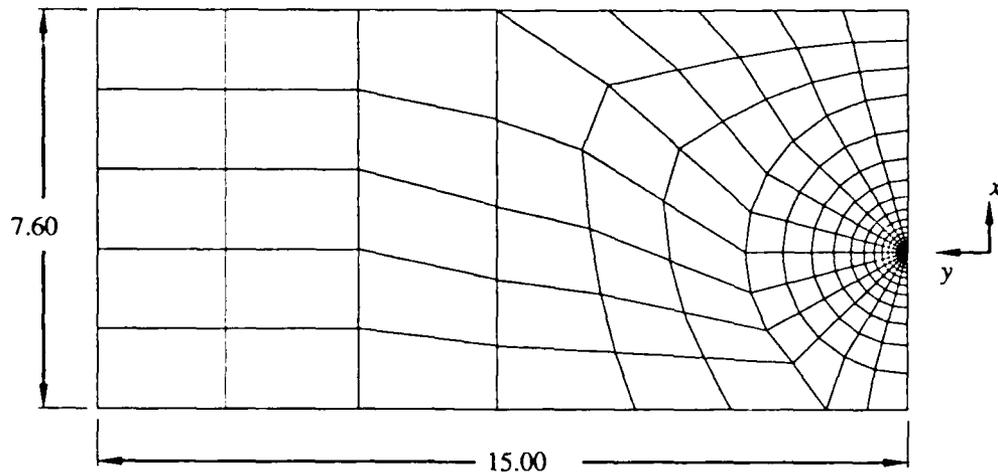


FIGURE 45. Diagram of the mesh in the second three point bend model. The crack is at the lower right, and is modeled as a notch having a circular tip of radius 0.01 centimeter. Again, the elements are concentrated near the crack tip. All dimensions in centimeters.

Displacement output from the second simulation is available for fourteen load steps: 390 N, 3950 N, 8450 N, 13450 N, 18450 N, 23450 N, 28450 N, 33450 N, 38450 N, 43450 N, 48450 N, 58450 N, 63450 N, and 73450 N.

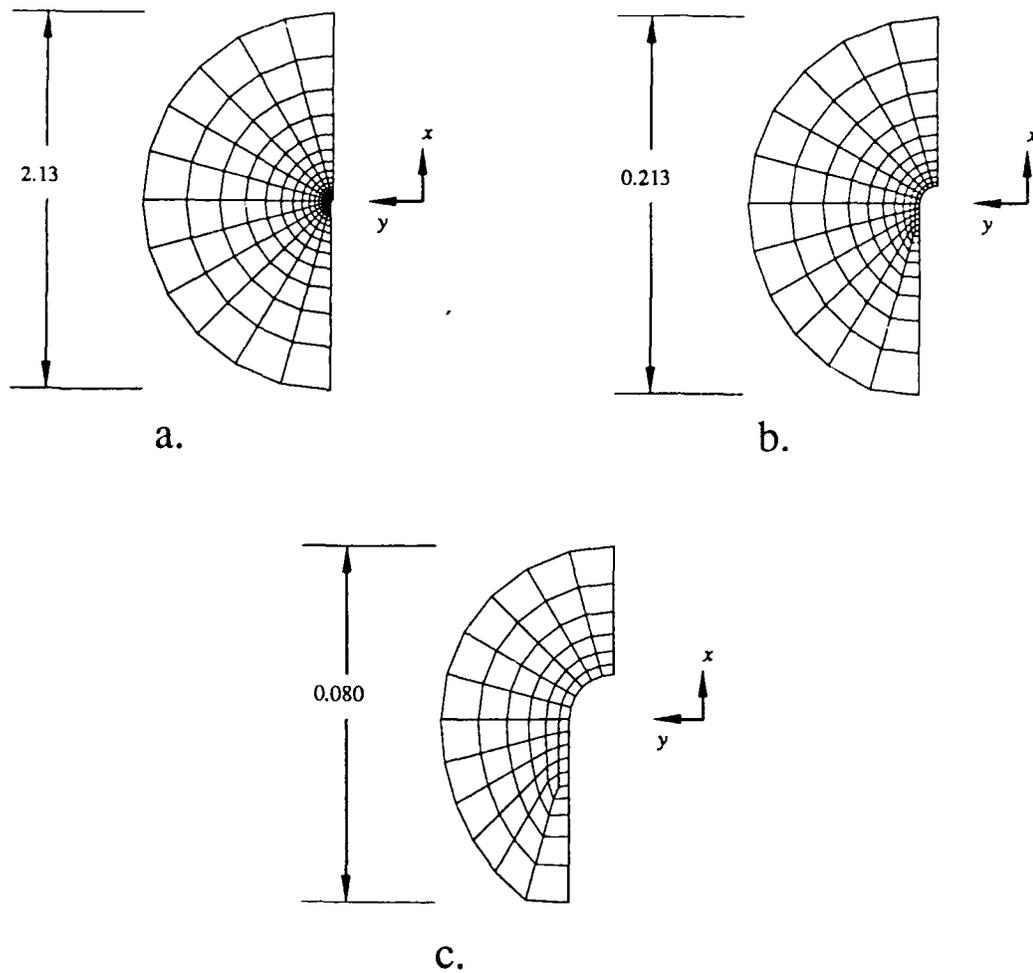


FIGURE 46. Detail of the mesh near the crack tip for the second three point bend model. The crack is modeled as a notch having a circular tip of radius 0.01 centimeter. The scale of each diagram corresponds to the diagrams in Figure 43. The elements for the second model are smaller and more nearly square than those in the first model. All dimensions in centimeters.

4.3.2 Model Constitutive Relations

The numerical models assume material homogeneity and isotropy in both the elastic and plastic constitutive relations, and the numerical simulations are carried out using a small strain, incremental plasticity theory. The stresses are computed

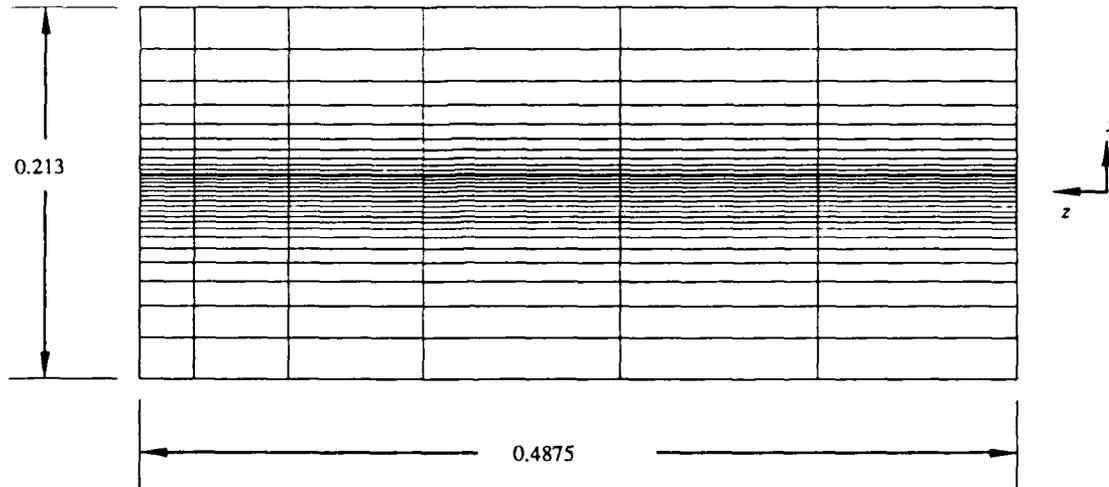


FIGURE 47. Detail of the element variation through the thickness of the second three point bend model. The crack tip is in the center of the diagram, running horizontally. The scale corresponds to diagram b in Figure 46. The free surface is at the left, the midplane at the right. All dimensions in centimeters.

and the yield surface updated using a tangential predictor - radial return method with subincrementation. The stress update procedure is described by Masimhan and Rosakis [14]. The \bar{B} method of Hughes [5] modified by Nakamura, Shih and Freund [12] was used to relieve artificial mesh locking due to plastic incompressibility. As recommended by Nakamura, Shih and Freund [12], a stabilization parameter of 0.05 was used in the calculations.

Both models used a power law hardening relation between the stress and strain in the plastic regime. The initial yield stress of the first model was about 10% higher than that of the second model. The plastic behavior of the two models was quite similar, as the hardening exponent of the first model was 22 as compared to 23.7 in the second model. An exponent of 1 corresponds to linear elasticity, while an exponent of infinity corresponds to an elastic-perfectly plastic material.

Although the constitutive relations used in the numerical models contain no

explicit time dependence (no rate dependent effects, no dynamic effects and no time dependent effects such as creep), the solution of the finite element equations is performed in a time domain to allow for the inclusion of such explicit time dependence. In this model the load history alone is explicitly a function of time, and time steps are used to implement load steps. The load history is monotonically increasing and is specified to be linear with time in this study.

4.3.3 Solution Approach

An iterative solution method is used to solve for the increments of the displacements, strains and stresses corresponding to the increment of the applied load at each time step.

It is assumed that an equilibrium solution for time t has been achieved. At time t , the displacements are held in the displacement matrix $\mathbf{U}(t)$. The strains at time t are calculated from the displacements through the $\bar{\mathbf{B}}$ matrix

$$\boldsymbol{\epsilon}(t) = \bar{\mathbf{B}}\mathbf{U}(t) \quad (4.3.1)$$

The $\bar{\mathbf{B}}$ matrix is used in place of the normal \mathbf{B} matrix to eliminate numerical error caused by mesh locking due to plastic incompressibility. The normal \mathbf{B} matrix is calculated from the derivatives of the shape functions for the element. The $\bar{\mathbf{B}}$ matrix is derived from the \mathbf{B} matrix by dividing the \mathbf{B} matrix into a deviatoric part and a volumetric part, $\mathbf{B} = \mathbf{B}_{dev} + \mathbf{B}_{vol}$. The volumetric part of the \mathbf{B} matrix is replaced by a volumetric matrix calculated at a reduced number of quadrature points using reduced order shape functions, $\bar{\mathbf{B}} = \mathbf{B}_{dev} + \bar{\mathbf{B}}_{vol}$. In elastic analysis, the $\bar{\mathbf{B}}$ method yields an identical stiffness matrix when compared to selective reduced

integration in which the compliance matrix, \mathbf{C} , is separated into deviatoric and volumetric parts, and the volumetric part is numerically integrated at the reduced quadrature points. A stabilization parameter of 0.05 was used to create a $\bar{\mathbf{B}}$ matrix which lies between the usual \mathbf{B} matrix and the full $\bar{\mathbf{B}}$ matrix, $\bar{\mathbf{B}} = \mathbf{B}_{dev} + \bar{\mathbf{B}}_{vol} + 0.05(\mathbf{B}_{vol} - \bar{\mathbf{B}}_{vol})$. This stabilization parameter is intended to eliminate spurious zero energy modes which may occur due to the reduced integration, while still providing a strain displacement matrix that does not significantly over-estimate the stiffness of the element in the event of the nearly incompressible behavior in fully plastic deformations.

Since small strains are assumed in the model, the $\bar{\mathbf{B}}$ matrix remains constant throughout the loading history. The equilibrium stresses at time t are held in the matrix $\boldsymbol{\sigma}(t)$, and the nodal force vector $\mathbf{P}(t)$ is calculated from the stresses through the $\bar{\mathbf{B}}$ matrix in a numerically calculated integral over the volume of the model

$$\mathbf{P}(t) = \int_{Volume} \bar{\mathbf{B}} \boldsymbol{\sigma}(t) \, dV \quad (4.3.2)$$

The applied nodal force vector $\mathbf{F}(t)$ is specified by the load history. Gross equilibrium of the model requires that

$$\mathbf{F}(t) = \mathbf{P}(t) \quad (4.3.3)$$

for all time t . The tangent stiffness matrix $\mathbf{K}(t)$ at time t is calculated by numerical integration over the volume

$$\mathbf{K}(t) = \int_{Volume} \bar{\mathbf{B}}^T \mathbf{C}^*(t) \bar{\mathbf{B}} \, dV \quad (4.3.4)$$

where $\mathbf{C}^*(t)$ is the history dependent compliance matrix, derived from the history dependent compliance four tensor. As described in the section on the elastic and

plastic constitutive relations, $C^*(t)$ depends on the accumulated plastic strain as well as the current stress and strain states. For purely elastic behavior, C^* reduces to the elastic compliance C .

The time is incremented to $t + \Delta t$, where Δt is a suitably chosen time increment. At time $t + \Delta t$, the externally applied forces $F(t + \Delta t)$ must still balance with the forces due to the element stresses $P(t + \Delta t)$,

$$F(t + \Delta t) = P(t + \Delta t) \quad (4.3.5)$$

The forces due to the element stresses at time $t + \Delta t$ are decomposed into the known forces at time t , $P(t)$, and the increment at time $t + \Delta t$, ΔP . The increment of the nodal forces due to the element stresses is approximated through use of the tangent stiffness matrix of time t , $K(t)$, multiplying the displacement increment at time $t + \Delta t$, ΔU .

$$\Delta P \approx K(t)\Delta U \quad (4.3.6)$$

Thus the equation that must be solved for the displacement increment ΔU at time $t + \Delta t$ is

$$\begin{aligned} K(t)\Delta U &= F(t + \Delta t) - P(t) \\ &= \Delta F \end{aligned} \quad (4.3.7)$$

The stiffness matrix is inverted and the matrix equation is solved using Gaussian elimination, leading to an approximation for the displacement increment ΔU .

Equation (4.3.7) is a first approximation to the equilibrium solution at time $t + \Delta t$. In general, the actual increment of the nodal forces corresponding to the displacement increments will not satisfy the equilibrium condition of equation (4.3.5). In order to refine the solution, an iteration procedure is used.

4.3.4 BFGS Iteration

The solution for the displacement increment at time $t + \Delta t$ in equation (4.3.7) is used as a first approximation in an iterative scheme to improve the solution. This iteration method is called the BFGS method (after Broyden, Fletcher, Goldfarb and Shanno), and was suggested for use in finite elements by Matthies and Strang [10]. This method is labelled a quasi-Newton approach because it updates the stiffness matrix (or actually its inverse) in a limited way at each iteration to improve speed of convergence. A full Newton method updates the stiffness matrix completely at each iteration, while a modified Newton method does not update the stiffness matrix at all during iteration.

The matrix equation to be solved is still as shown in equation (4.3.7), however, it is now done in an iterative procedure. At the i^{th} iteration,

$$\begin{aligned} \left(\mathbf{K}(t + \Delta t) \right)^{(i-1)} \Delta \mathbf{U}^{(i)} &= \mathbf{F}(t + \Delta t) - \left(\mathbf{P}(t + \Delta t) \right)^{(i-1)} \\ &= \Delta \mathbf{F}^{(i-1)} \end{aligned} \quad (4.3.8)$$

where $\left(\mathbf{K}(t + \Delta t) \right)^{(0)} = \mathbf{K}(t)$ and $\left(\mathbf{P}(t + \Delta t) \right)^{(0)} = \mathbf{P}(t)$ are from the equilibrium solution at time t . $\Delta \mathbf{F}^{(i-1)}$ is the out of balance force vector which becomes zero (or very small) when equilibrium is satisfied.

As mentioned, the BFGS method effectively acts to update the inverse of the stiffness matrix at each iteration. Actually, the stiffness matrix is not inverted, but is factored. The BFGS method uses vector multiplications and the factors of the original stiffness matrix to solve for the displacement increment at each iteration. The effect of the vector multiplications is identical to updating of the inverse of the stiffness matrix. Rewriting equation (4.3.8) to use the inverse form of the stiffness

matrix,

$$\begin{aligned}\Delta \mathbf{U}^{(i)} &= \left(\mathbf{K}^{-1}(t + \Delta t) \right)^{(i-1)} \left[\mathbf{F}(t + \Delta t) - \left(\mathbf{P}(t + \Delta t) \right)^{(i-1)} \right] \\ &= \left(\mathbf{K}^{-1}(t + \Delta t) \right)^{(i-1)} \Delta \mathbf{F}^{(i-1)}\end{aligned}\quad (4.3.9)$$

The total displacement is then approximated by

$$\mathbf{U}^{(i)}(t + \Delta t) = \mathbf{U}^{(i-1)}(t + \Delta t) + \beta \Delta \mathbf{U}^{(i)} \quad (4.3.10)$$

with $\mathbf{U}^{(0)}(t + \Delta t) = \mathbf{U}(t)$. The scalar parameter β is varied until the component of the new out of balance load vector parallel to the vector $\Delta \mathbf{U}^{(i)}$ is small compared to some tolerance. This is accomplished by varying β until the calculated nodal force vector, $\left(\mathbf{P}(t + \Delta t) \right)^{(i)}$, which corresponds to β is such that the inner product

$$\left(\Delta \mathbf{U}^{(i)} \right)^T \left[\mathbf{F}(t + \Delta t) - \left(\mathbf{P}(t + \Delta t) \right)^{(i)} \right] = \left(\Delta \mathbf{U}^{(i)} \right)^T \Delta \mathbf{F}^{(i)} \quad (4.3.11)$$

is less than some tolerance.

The new out of balance force vector depends on the parameter β through the constitutive law. Increasing β increases the displacement increment and therefore the strain increment. The stress increment is a nonlinear function of the strain increment, and the new nodal forces are calculated from the stresses as in equation, (4.3.2), which is also put in iterative form

$$\left(\mathbf{P}(t + \Delta t) \right)^{(i)} = \int_{Volume} \bar{\mathbf{B}} \left(\boldsymbol{\sigma}(t + \Delta t) \right)^{(i)} dV \quad (4.3.12)$$

The procedure for calculating the stresses corresponding to the displacement increment $\beta \Delta \mathbf{U}$ are described in the next section.

As described in Matthies and Strang, the BFGS method is an attempt to make an efficient iterative search for the solution $\Delta \mathbf{F} = \mathbf{0}$. The variation of β produces a search direction which is nearly orthogonal to the out of balance force $\Delta \mathbf{F}^{(i)}$. In particular, the updating of the inverse of the stiffness matrix is made inexpensive to compute.

Two difference vectors, $\delta^{(i)}$ and $\gamma^{(i)}$ are formed to represent the change in the displacements and out of balance forces between increments.

$$\delta^{(i)} = \left(\mathbf{U}(t + \Delta t) \right)^{(i)} - \left(\mathbf{U}(t + \Delta t) \right)^{(i-1)} \quad (4.3.13)$$

$$\gamma^{(i)} = \Delta \mathbf{F}^{(i-1)} - \Delta \mathbf{F}^{(i)} \quad (4.3.14)$$

The updated inverse of the stiffness matrix, $(\mathbf{K}^{-1}(t + \Delta t))^{(i)}$, should satisfy the quasi-Newton equation

$$\delta^{(i)} = (\mathbf{K}^{-1}(t + \Delta t))^{(i)} \gamma^{(i)} \quad (4.3.15)$$

In the BFGS method, the update of the inverse of the stiffness matrix is accomplished through a matrix multiplication of the form

$$(\mathbf{K}^{-1}(t + \Delta t))^{(i)} = (\mathbf{A}^T)^{(i)} \left(\mathbf{K}^{-1}(t + \Delta t) \right)^{(i-1)} \mathbf{A}^{(i)} \quad (4.3.16)$$

where $\mathbf{A}^{(i)}$ is a square matrix calculated from $\delta^{(i)}$ and $\gamma^{(i)}$. $\mathbf{A}^{(i)}$ is defined by

$$\mathbf{A}^{(i)} = \mathbf{I} + \mathbf{v}^{(i)} (\mathbf{w}^{(i)})^T \quad (4.3.17)$$

$$\mathbf{v}^{(i)} = - \left[\frac{(\delta^{(i)})^T \gamma^{(i)}}{(\delta^{(i)})^T (\mathbf{K}(t + \Delta t))^{(i-1)} \delta^{(i)}} \right]^{\frac{1}{2}} \left(\mathbf{K}(t + \Delta t) \right)^{(i-1)} \delta^{(i)} - \gamma^{(i)} \quad (4.3.18)$$

$$\mathbf{w}^{(i)} = \frac{\delta^{(i)}}{(\delta^{(i)})^T \gamma^{(i)}} \quad (4.3.19)$$

The term in brackets in equation (4.3.18) is the condition number of the update, $c^{(i)}$.

$$c^{(i)} = \left[\frac{(\delta^{(i)})^T \gamma^{(i)}}{(\delta^{(i)})^T (\mathbf{K}(t + \Delta t))^{(i-1)} \delta^{(i)}} \right]^{\frac{1}{2}} \quad (4.3.20)$$

If the condition number becomes very large, the update in equation (4.3.16) would be numerically dangerous, so a preset tolerance is used which prevents updating if the condition number exceeds the tolerance.

With the newly calculated matrices $(\mathbf{K}^{-1}(t + \Delta t))^{(i)}$ and $(\mathbf{P}(t + \Delta t))^{(i)}$, the process can begin again at equation (4.3.9) to evaluate the displacement increment at iteration $i + 1$.

The iteration for the displacement and force increments continues until the Euclidean norm of the out of balance force vector $\Delta \mathbf{F}^{(i)}$ is smaller than the out of balance force at the start of the iteration by a preset tolerance.

$$\|\Delta \mathbf{F}^{(i)}\| \leq \text{TOL} \|\Delta \mathbf{F}^{(1)}\| \quad (4.3.21)$$

The displacement at time $t + \Delta t$, $\mathbf{U}(t + \Delta t)$, is then given by

$$\mathbf{U}(t + \Delta t) = \mathbf{U}(t) + \sum_{i=1}^n \Delta \mathbf{U}^{(i)} \quad (4.3.22)$$

where n is the number of iterations to convergence.

4.3.5 Stress Update

The stresses in the elements are computed from the calculated strain increment using a tangential predictor – radial return method. In order to prevent artificial unloading during the iteration procedure, the stresses and strains at each iteration are calculated from the known equilibrium solution from the previous time step and the total displacement increment for the current time step. At the i^{th} iteration for time step $t + \Delta t$, the total displacement increment, $\Delta \mathbf{U}^{\text{total}}$, is the sum of the displacement increments for iterations 1 through i for time step $t + \Delta t$. The strain increment due to that displacement increment is given by

$$\Delta \boldsymbol{\epsilon} = \bar{\mathbf{B}} \Delta \mathbf{U}^{\text{total}} \quad (4.3.23)$$

In calculating the stress increment for the current iteration, this total strain increment is used along with the previous equilibrium solution to prevent the possibility that the stress increment at one iteration will lie outside the current yield surface, while the the next stress increment will be a correction in the opposite direction, causing artificial elastic unloading due to finite step sizes.

The procedure for updating the stress and yield surface requires only the constant elastic compliance matrix, \mathbf{C} , and the current parameter defining the yield surface, $\bar{\sigma}$. First, an elastic stress increment is determined from the strain increment through the elastic compliance tensor \mathbf{C} .

$$\Delta\sigma^e = \mathbf{C} \Delta\epsilon \quad (4.3.24)$$

A trial stress state σ^A is calculated from the stress increment and the stress state from the known equilibrium solution at time t , σ^0 .

$$\sigma^A = \sigma^0 + \Delta\sigma^e \quad (4.3.25)$$

The stress state σ^0 is assumed to lie inside the yield surface at time t . The yield surface at time t is defined by the function $f(\sigma, \epsilon) = 3J_2(\sigma) - \bar{\sigma}^2$, where $\bar{\sigma}$ is a function of the accumulated equivalent plastic strain, and is a measure of the current yield stress. At time t , $\bar{\sigma} = \bar{\sigma}^0$. If the trial stress state σ^A is such that

$$3J_2(\sigma^A) \leq (\bar{\sigma}^0)^2 \quad (4.3.26)$$

the stress path has stayed in the elastic regime, and no further calculation of the stress state is necessary. Otherwise, the yield surface has been crossed by the trial stress solution, and the yield surface and stress state must be updated together. A stress state between the equilibrium stress of time t , σ^0 and the trial stress state σ^A is found which lies on the yield stress of time t . This contact stress state, σ^C , is defined by

$$\sigma^C = \sigma^0 + q\sigma^A \quad (4.3.27)$$

and

$$3J_2(\boldsymbol{\sigma}^C) = (\bar{\sigma}^0)^2 \quad (4.3.28)$$

For the isotropic hardening material satisfying the Huber–Von Mises yield condition, this results in a quadratic equation for the scalar parameter q . Note that the path from $\boldsymbol{\sigma}^0$ to $\boldsymbol{\sigma}^C$ is completely elastic.

A corrected trial stress $\boldsymbol{\sigma}^T$ is determined from the contact stress $\boldsymbol{\sigma}^C$ using the assumption that the plastic strain rate is always normal to the yield surface. If the plastic strain rate is directed along a normal to the Huber–Von Mises yield surface, it will be given by

$$\dot{\boldsymbol{\epsilon}}^p = \dot{\lambda} \mathbf{S} \quad (4.3.29)$$

where \mathbf{S} is the deviatoric stress and the invariant $J_2 = \frac{1}{2} S_{ij} S_{ij}$. An approximate plastic strain increment is given by $\Delta\lambda \mathbf{S}$, where \mathbf{S} and $\Delta\lambda$ are calculated at the contact stress $\boldsymbol{\sigma}^C$. This plastic strain increment is subtracted from the calculated strain increment from the stress state $\boldsymbol{\sigma}^C$ to the stress state $\boldsymbol{\sigma}^A$, approximated by $(1 - q)\Delta\boldsymbol{\epsilon}$, to estimate the elastic strain increment during plastic deformation.

$$\Delta\boldsymbol{\epsilon}^e \approx (1 - q)\Delta\boldsymbol{\epsilon} - \Delta\lambda \mathbf{S} \quad (4.3.30)$$

The stress increment corresponding to that elastic strain increment is

$$\Delta\boldsymbol{\sigma} = \mathbf{C}((1 - q)\Delta\boldsymbol{\epsilon} - \Delta\lambda \mathbf{S}) \quad (4.3.31)$$

The corrected trial stress $\boldsymbol{\sigma}^T$ is the contact stress $\boldsymbol{\sigma}^C$ plus the calculated stress increment.

$$\begin{aligned} \boldsymbol{\sigma}^T &= \boldsymbol{\sigma}^C + \mathbf{C}((1 - q)\Delta\boldsymbol{\epsilon} - \Delta\lambda \mathbf{S}) \\ &= \boldsymbol{\sigma}^C + ((\boldsymbol{\sigma}^A - \boldsymbol{\sigma}^C) - \Delta\lambda \mathbf{C} \mathbf{S}) \\ &= \boldsymbol{\sigma}^A - \Delta\lambda \mathbf{C} \mathbf{S} \end{aligned} \quad (4.3.32)$$

The yield surface is updated using the hardening law H , in this case a power law. The accumulated equivalent plastic strain increment $\Delta\bar{\epsilon}^p$ is given by

$$\begin{aligned}
 \Delta\bar{\epsilon}^p &= \left(\frac{2}{3}\Delta\epsilon_{ij}^p\Delta\epsilon_{ij}^p\right)^{\frac{1}{2}} \\
 &= \left(\frac{2}{3}\Delta\lambda S_{ij}\Delta\lambda S_{ij}\right)^{\frac{1}{2}} \\
 &= \Delta\lambda\left(\frac{2}{3}S_{ij}S_{ij}\right)^{\frac{1}{2}} \\
 &= \Delta\lambda\left(\frac{2}{3}2J_2\right)^{\frac{1}{2}} \\
 &= \Delta\lambda\left(\frac{2}{3}\frac{2}{3}(\bar{\sigma}^0)^2\right)^{\frac{1}{2}} \\
 &= \frac{2}{3}\Delta\lambda\bar{\sigma}^0
 \end{aligned} \tag{4.3.33}$$

H is defined as the rate of change of the function $\bar{\sigma}$ with respect to the equivalent accumulated plastic strain $\bar{\epsilon}^p$, which is determined from the power law described in the section on the constitutive behavior. In order to update the yield surface, H will be determined at $\bar{\sigma}^0$.

$$H(\bar{\sigma}^0) = \left.\frac{d\bar{\sigma}}{d\bar{\epsilon}^p}\right|_{\bar{\sigma}=\bar{\sigma}^0} \tag{4.3.34}$$

The final value of the yield stress, $\bar{\sigma}^F$, is then calculated as

$$\bar{\sigma}^F = \bar{\sigma}^0 + H(\bar{\sigma}^0)\Delta\bar{\epsilon}^p \tag{4.3.35}$$

In general, the corrected trial stress σ^T will not lie on the updated surface. A final update of the stress state is scaled from σ^T , forcing it to lie on the updated yield surface

$$\sigma^F = \left(\frac{\bar{\sigma}^F}{\sqrt{3J_2(\sigma^T)}}\right)\sigma^T \tag{4.3.36}$$

The stress path from σ^C to σ^F is an elastic-plastic path. To minimize the effect of finite increments, the excess stress increment $\sigma^E - \sigma^C$ is divided into m subincrements and the corrected trial stress σ^T , the yield surface update, and the final correction to σ^F are calculated m times.

4.3.6 Discussion

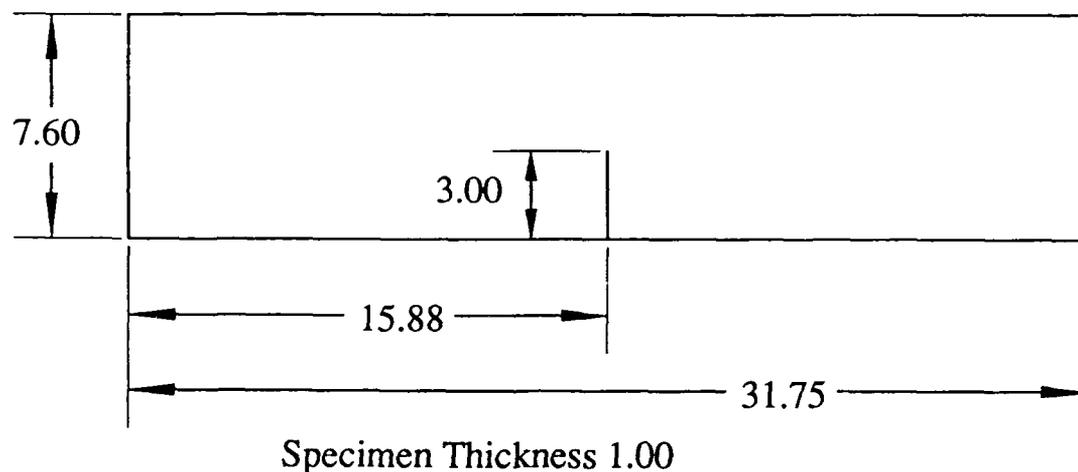
The BFGS method is an attempt to increase the speed of the convergence of the iterative solution. In general, the most time consuming portion of the solution is the factoring of the tangent stiffness matrix for solution of the equation $\mathbf{K} \Delta \mathbf{U} = \Delta \mathbf{F}$. This must be done only once per time step, however, with additional matrix updates done by simple multiplications. In contrast, the full Newton method must update and solve the entire matrix equation at each iteration. The modified Newton method must also update and solve the stiffness matrix equation at each time step, but retains that stiffness matrix throughout all iterations. The modified Newton method might seem the best choice, but the slowness of convergence in large plastic deformations offsets the gain in not updating the matrix. If the stiffness matrix were kept constant for all deformations, as in the elastic case, it would only have to be inverted once, but as the deformation moved away from its original state the convergence would be extremely slow because the search for a solution would take place along poorly chosen directions. However, if the number of iterations becomes large, the task of updating the matrix itself involves a large number of multiplications. In that case, the factors of the original stiffness matrix for that time step may be used alone, and the BFGS sequence started again.

This finite element model contains nonlinearity only in the constitutive relation, and not in the strain-displacement relation. The small strain assumption will become invalid in a region around the crack tip which grows with the applied load.

4.4 Three Point Bend Experiments

4.4.1 Specimen Material and Dimensions

The three point bend experiments were performed on Aircraft Quality Vacuum Degassed 4340 steel plates 1 cm thick and 30.5 cm long by 7.6 cm high with a single edge crack 3 cm in length as shown in Figure 48. The loading configuration is shown in Figure 49.



All Dimensions in centimeters.

FIGURE 48. Three point bend specimen geometry. Material is a ductile heat treatment of 4340 steel.

All of the test specimens were cut from the same piece of 4340 steel, a plate four inches wide by one-half inch thick by eight feet long. The steel was heat treated to give a ductile response, first being "normalized" at 871 C with air cooling to remove any effects of rolling or cutting, then heated to 815 C and oil quenched to produce a Martensitic transformation and finally tempered at a temperature of 538 C for

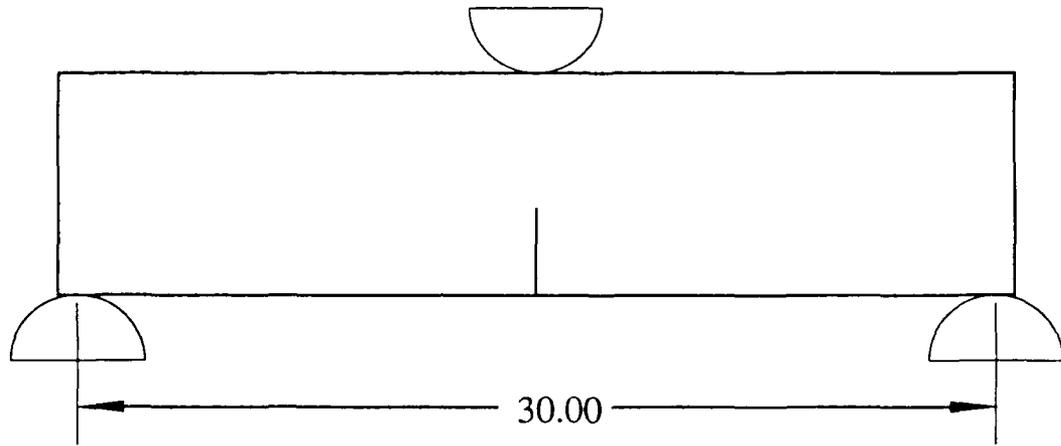


FIGURE 49. Loading configuration for three point specimen, showing the cylindrical supports and cylindrical indenter opposite the crack. (In the experiment, the specimen was inverted, with the crack in the top edge.)

one hour and air cooled. Two uniaxial test specimens were also cut from the same piece of steel and heat treated identically with the three point bend specimens. The uniaxial test geometry is shown in Figure 50. The composition of the 4340 steel is shown in Table 1.

Chemical Composition, %											
C	Mn	P	S	Si	Ni	Cr	Mo	Cu	Sn	Al	Fe
0.420	0.740	0.010	0.012	0.250	1.710	0.800	0.240	0.060	0.000	0.036	balance

Table 1. Composition of 4340 Steel. Aircraft quality, vacuum degassed, cold drawn.

One side of the fracture specimen was lapped flat and polished to a mirror finish in order to measure the out-of-plane displacements with a Twyman-Green interferometer, while a diffraction grating was epoxied to the other side of the specimen as described in the section describing the moire interferometer to measure the two

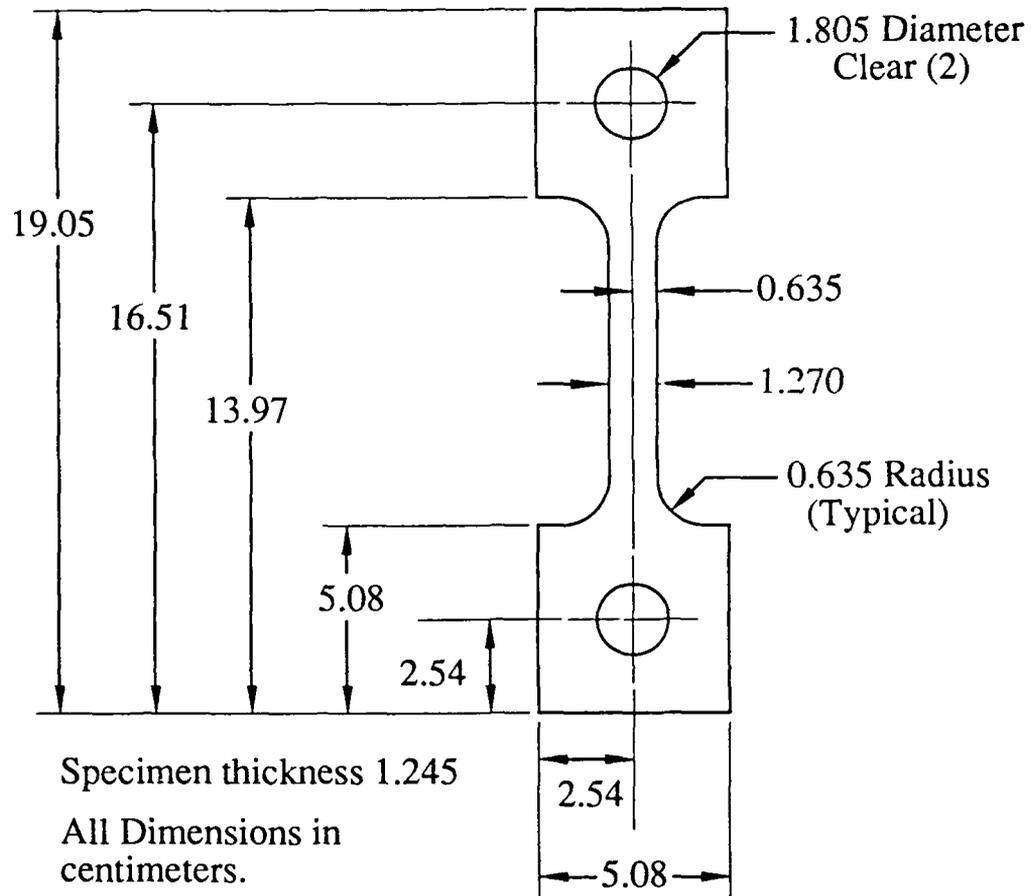


FIGURE 50. Uniaxial test specimen geometry. Material is a ductile heat treatment of 4340 steel. Test section is ground flat and parallel with a thickness of 1.245 cm.

in-plane displacement components. It is assumed that the specimen deformation is symmetric, so that the three measured components can be treated as if they were measured on the same surface.

The current material properties were found to be different from the previous work, by Zehnder and Rosakis, as seen from Figure 39. The Young's modulus was about 3.5% lower, and the yield stress was about 10% lower than that of Zehnder's

experiment. The calculated hardening exponent was 23.7 as opposed to 22 from the earlier experiments. Because of these discrepancies, and also because the numerical data available from Narasimhan's numerical comparison was somewhat sparse, it was decided to try to rerun the numerical simulation with the current material parameters (see the sections describing the elastic-plastic constitutive relation and the finite element model). Narasimhan's data was retained for only five load steps: 24 N, 35000 N, 52296 N, 68256 N and 78863 N. Data from the second numerical simulation is available at fourteen load steps: 390 N, 3950 N, 8450 N, 13450 N, 18450 N, 23450 N, 28450 N, 33450 N, 38450 N, 433450 N, 48450 N, 58450 N, 63450 N and 73450 N.

4.4.2 Experiment

The experiments were carried out on the Newport Corporation vibration isolated optical table (ten feet by four feet by eighteen inches thick) to reduce vibrations from the surroundings. The loading frame was designed and constructed at Caltech and has a test section length which is adjustable by means of threaded rods. The load itself is introduced through a hydraulic ram which is supplied from a Miller 559-Hydraulic Power Unit 66 which is rated for 11.6 gallons per minute at a working pressure of 2670 psi. The pressure is supplied through a Moog Series 62 Flow Control Servovalve which is controlled by a Shore Western Mfg., Inc. SC1100-1 DC Servo Controller.

The servo controller balances the load feedback against an input command voltage to provide control of the desired applied load. The load was measured and fed back to the servo controller with the same Sensotec 400,000 Newton capacity load cell used in the uniaxial tests. As seen from the figure, the central point of

the three point bend arrangement is held fixed, transmitting the load to the load cell, while the two outer loading points are attached to an I beam which is moved by the hydraulic ram. Although this setup makes it slightly more difficult to align the apparatus under no load, it does have the advantage that the crack tip remains nearly stationary under loading. This is important because it means that the optics do not need gross adjustments during a test.

During a test, the servo controller is supplied with a command voltage corresponding to a load where information is available from one of the finite element simulations. That load is held fixed while the optics are adjusted to try to remove any effects of rigid rotations which are most easily seen as loss of symmetry in the fringe patterns. For small deviations, this adjustment is orthogonal to the adjustment for the spacing of the virtual reference grating formed by the crossed incoming laser beams, so the virtual reference grating spacing and the displacement increment between contours remains constant. The two in-plane displacement fringe patterns are then photographed, as is the out-of-plane displacement pattern on the opposite side of the specimen. The load and load point displacement are recorded continuously by either the Masscomp Laboratory Workbench software or the A/D converter card in the Image Processing computer system.

In the first experiment, the in-plane displacements were photographed using a Nikon FA camera body with a Nikon Micro-Nikkor 200mm f/4 lens and a Nikon 3T close-up lens (focal length 665 mm) attached to the front of the 200mm lens. This camera setup produced a magnification of about 0.8 (that is, 1 mm on the specimen corresponded to 0.8 mm on the negative). After the fringes were photographed this way, a Nikon TC-200 Teleconverter was placed between the camera and the 200 mm lens. The teleconverter doubles the magnification of the image without disturbing

the focus of the lens. Six pictures of each displacement component were taken at each load step, three using the teleconverter and three without. The three pictures were each taken using a different length of exposure, in an attempt to bracket the optimum exposure and also to provide some redundancy in the event of a mishap in the development. (Since the image is made up of essentially collimated light, the exposure cannot be controlled by the aperture without losing some of the information around the edges.) The pictures taken without the teleconverter were intended to capture the fringe information away from the crack, but the magnification was not high enough to record the smallest fringe spacing which the optical system was capable of transmitting to the film. With the teleconverter in the path, the field of view is reduced, but the image is magnified enough so that the limitations due to the film resolution are less restrictive than the limits due to the finite aperture of the camera. The film used was Kodak Technical Pan set at an exposure index of 120 and developed in Kodak HC-110 (dilution D) for 6 minutes. This film has extremely fine grain size, and it can record fringes with wavelengths as small as 5 microns at the film plane, this corresponds to 5 microns divided by the camera magnification on the specimen.

SECTION 5

Comparison of Results

5.1 Numerical and Experimental Comparison

The in-plane displacement contours on the free surface of the specimen as generated from the numerical program at 3950 N are shown in Figures 51 and 52. The contours are only shown for a region of 5 cm in diameter which represents the maximum field of view of the experiment.

The corresponding contours as measured from the experiment are shown in Figures 53 and 54. It can be seen that there is noise in the measured contours near the crack tip. These contours are determined from the experimental photograph which has the largest field of view. The data from the photograph with higher magnification should reduce the noise, but at the cost of a smaller field of view. This comparison is made for the purpose of examining the large scale match of the numerical and experimental data.

The contours from the experiment and the numerical model are shown together in Figures 55 and 56. As seen in Figure 55, the contours of constant U displacement match well except ahead of the crack ($x > 1.5$). This appears to be due to a misalignment of the saddle point in the fringe pattern. At this low load level, the small number of fringes makes it difficult to align the saddle point precisely. (An

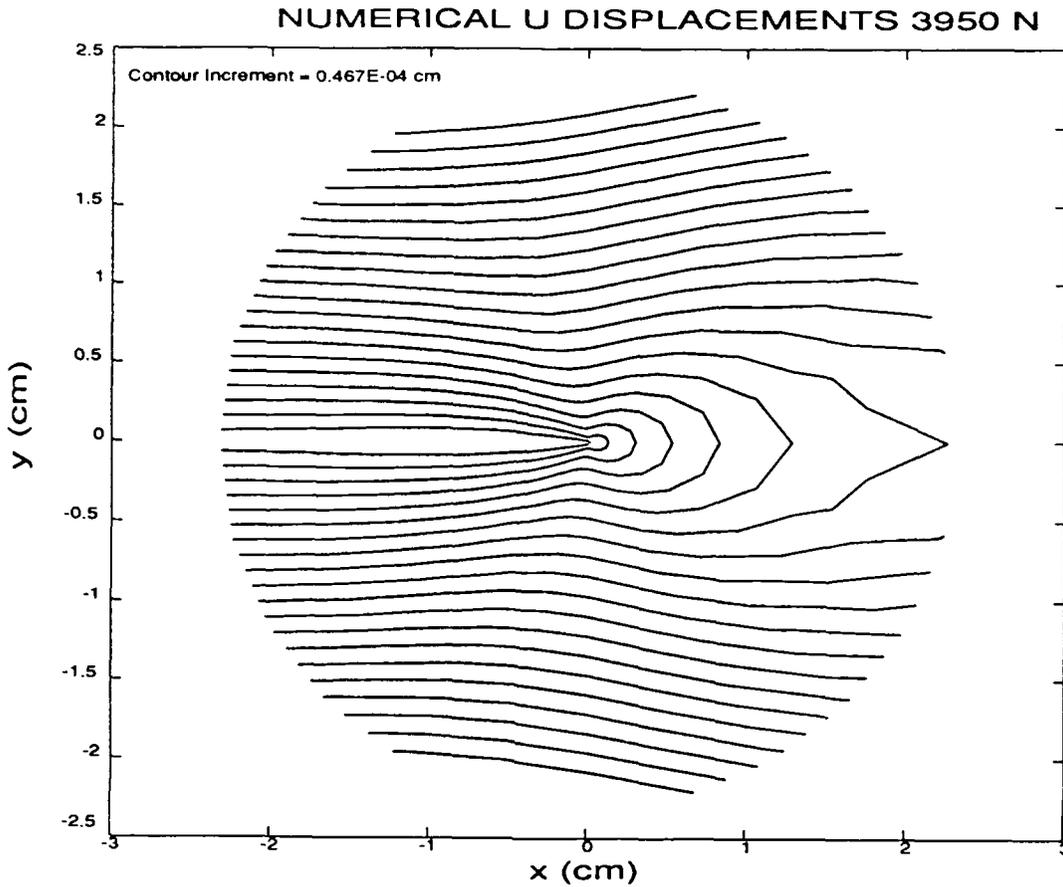


FIGURE 51. Contours of constant U displacement at the free surface at 3950 N. These contours are generated from the three dimensional finite element model. The crack lies along the line $x < 0$, with the crack tip at $x = 0$, $y = 0$. The increment of displacement between contours is 4.67×10^{-5} cm.

attempt was made to align the fringe saddle point with the saddle point predicted by the numerical model for comparison.) The contours of constant V displacement also match well at this low load level.

The deformation at 3950 is mostly elastic, with any plastic deformation contained in a region of about 0.01 cm around the crack tip (as determined from the numerical data). This is the lowest load for which meaningful experimental data is available. The ultimate failure load of the specimen is about 75000 N.

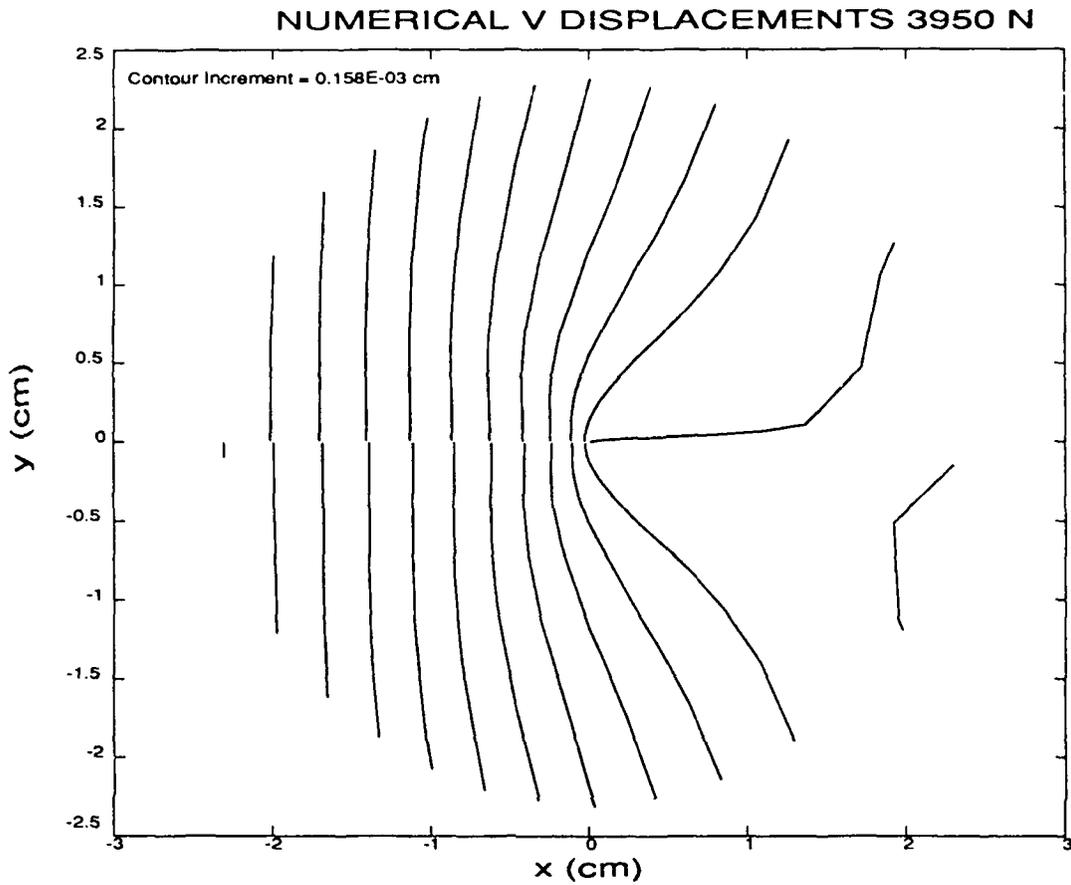


FIGURE 52. Contours of constant V displacement at the free surface at 3950 N. These contours are generated from the three dimensional finite element model. The crack lies along the line $x < 0$, with the crack tip at $x = 0$, $y = 0$. The increment of displacement between contours is 1.58×10^{-4} cm.

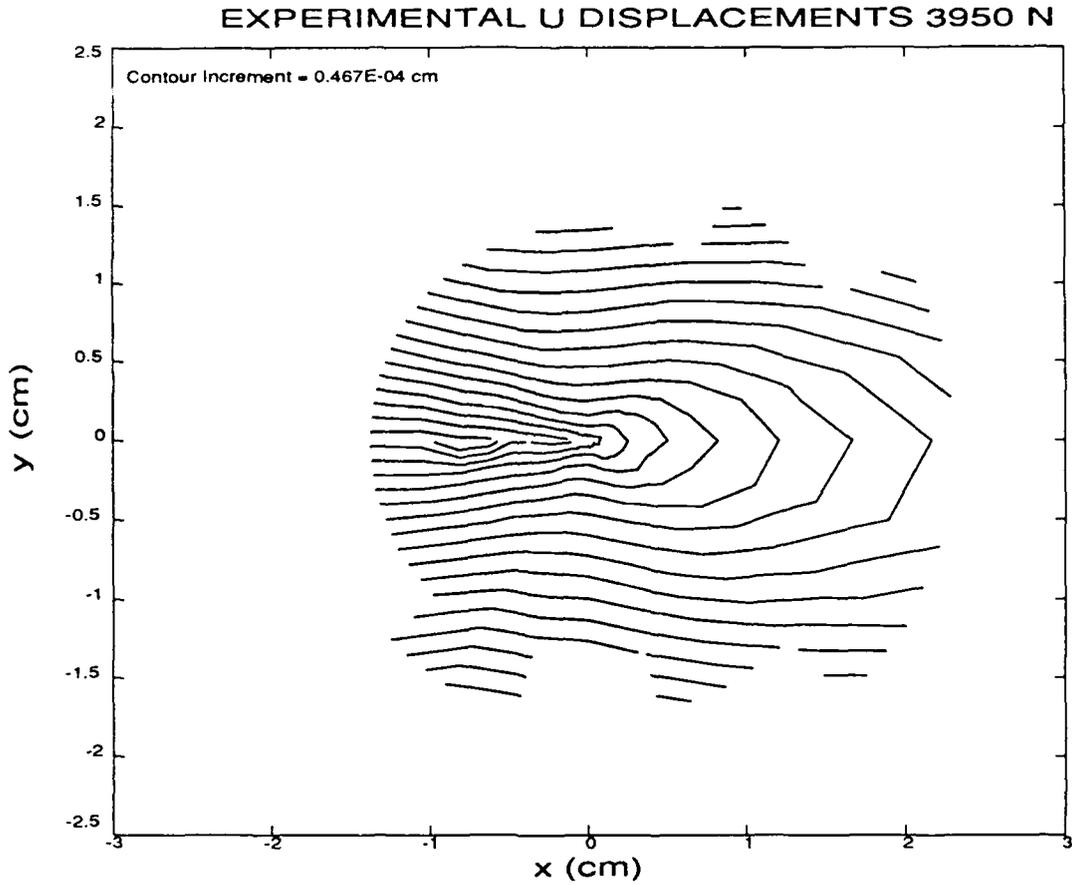


FIGURE 53. Contours of constant U displacement at the free surface at 3950 N. These contours are measured from the experiment. The contour increment is the same as in Figure 51, which shows contours from the numerical model. The crack lies along the line $x < 0$, with the crack tip at $x = 0, y = 0$. The increment of displacement between contours is 4.67×10^{-5} cm.

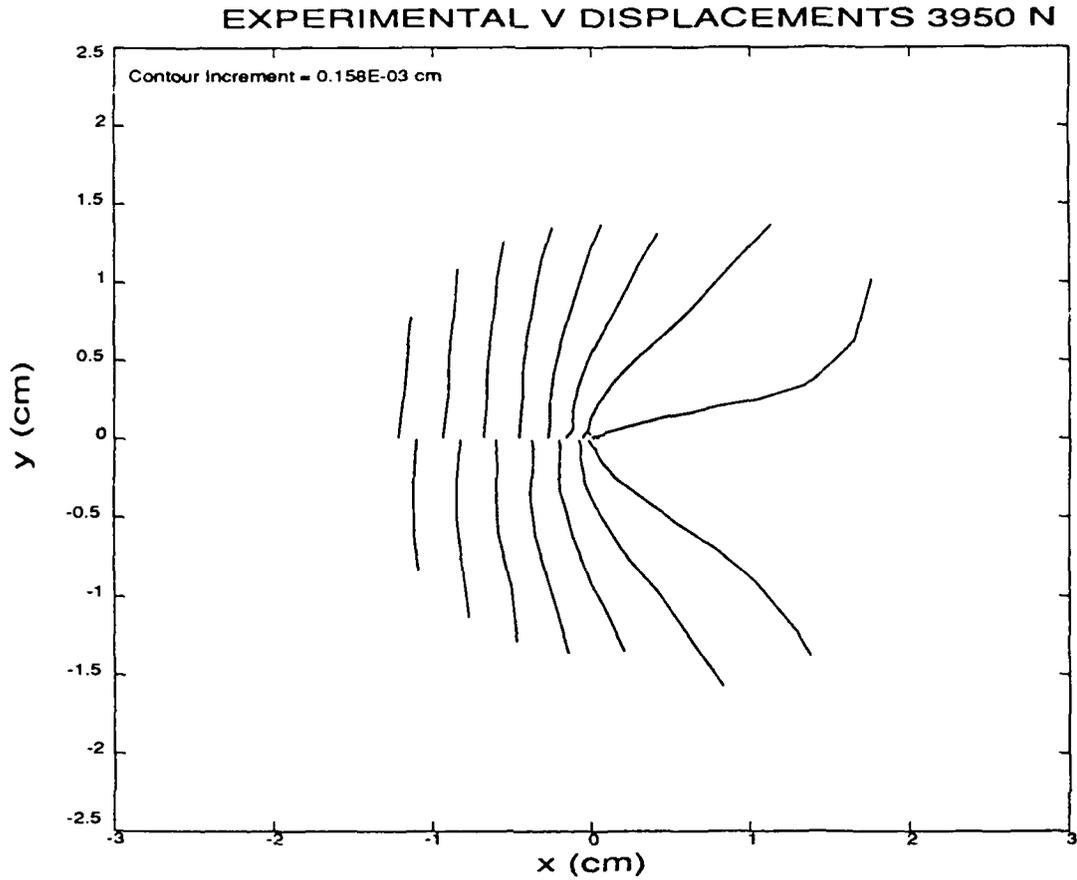


FIGURE 54. Contours of constant V displacement at the free surface at 3950 N. These contours are measured from the experiment. The contour increment is the same as in Figure 52, which shows contours from the numerical model. The crack lies along the line $x < 0$, with the crack tip at $x = 0, y = 0$. The increment of displacement between contours is 1.58×10^{-4} cm.

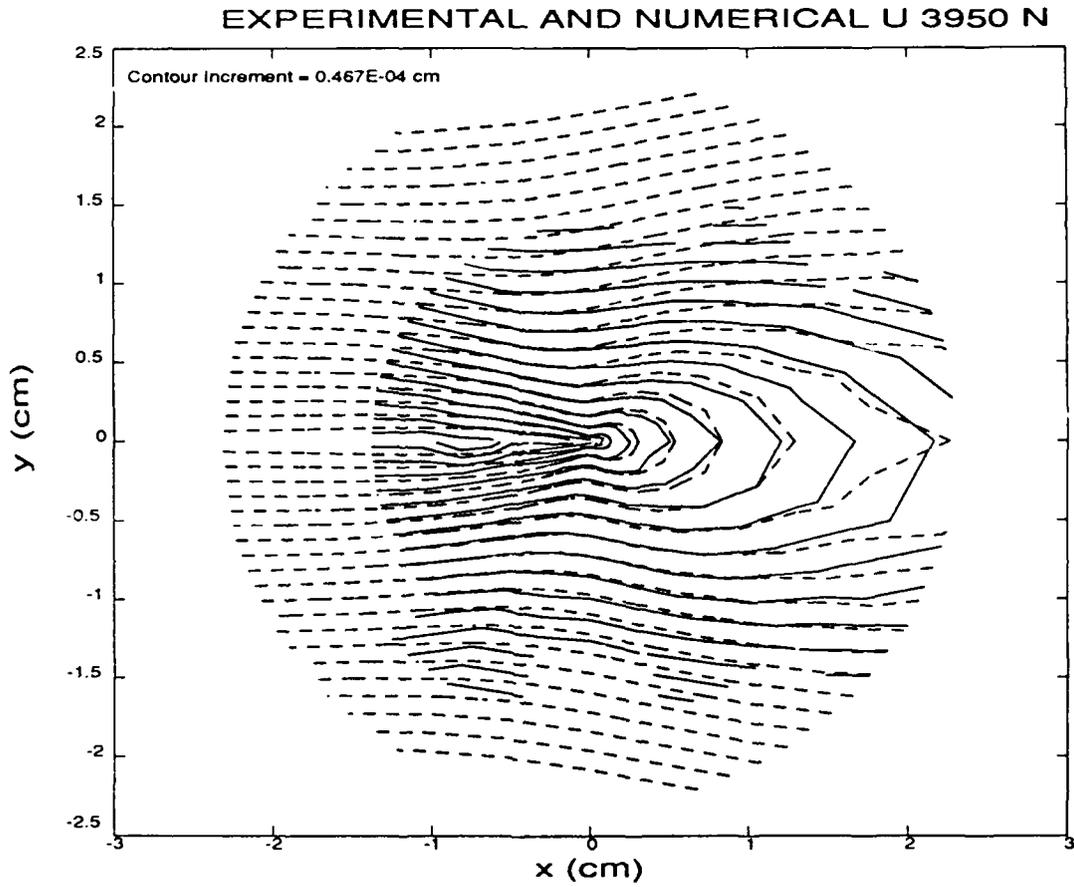


FIGURE 55. Contours of constant U displacement at the free surface at 3950 N. The contours measured from the experiment are solid lines, and the contours from the numerical model are dashed lines. The increment of displacement between contours is 4.67×10^{-5} cm.

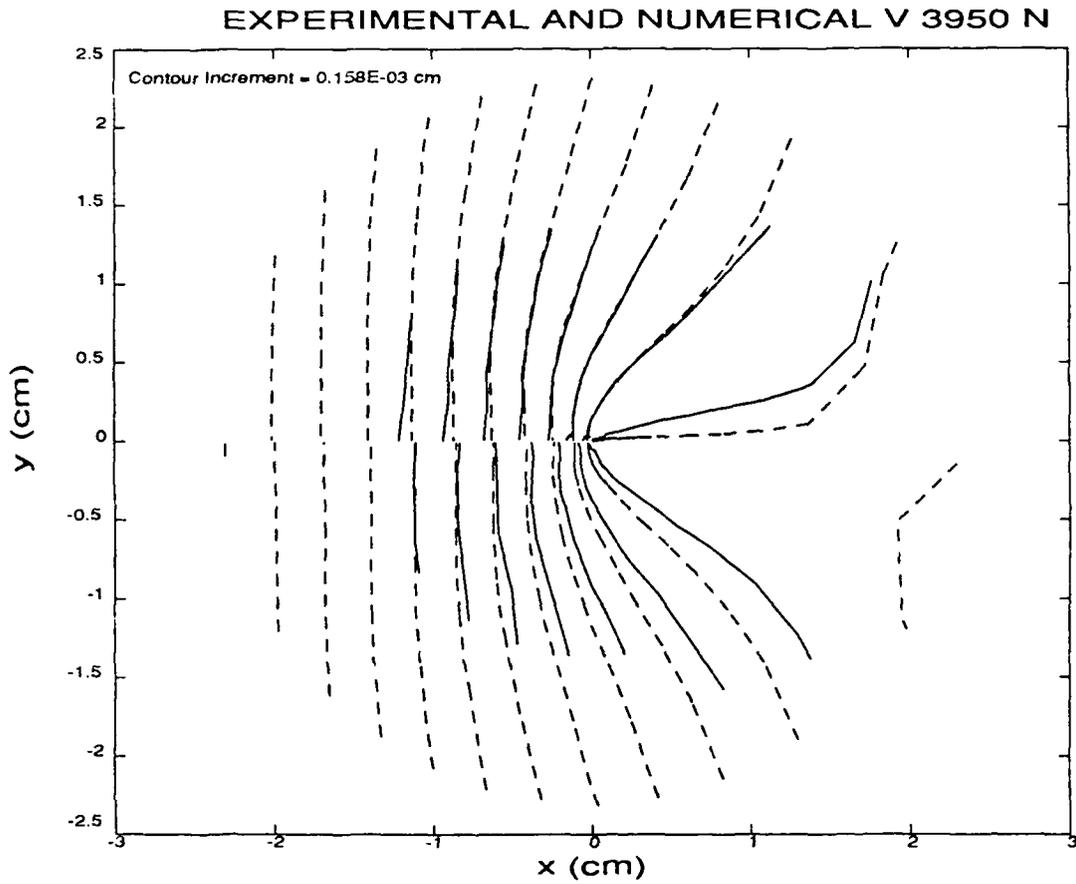


FIGURE 56. Contours of constant V displacement at the free surface at 3950 N. The contours measured from the experiment are solid lines, and the contours from the numerical model are dashed lines. The increment of displacement between contours is 1.58×10^{-4} cm.

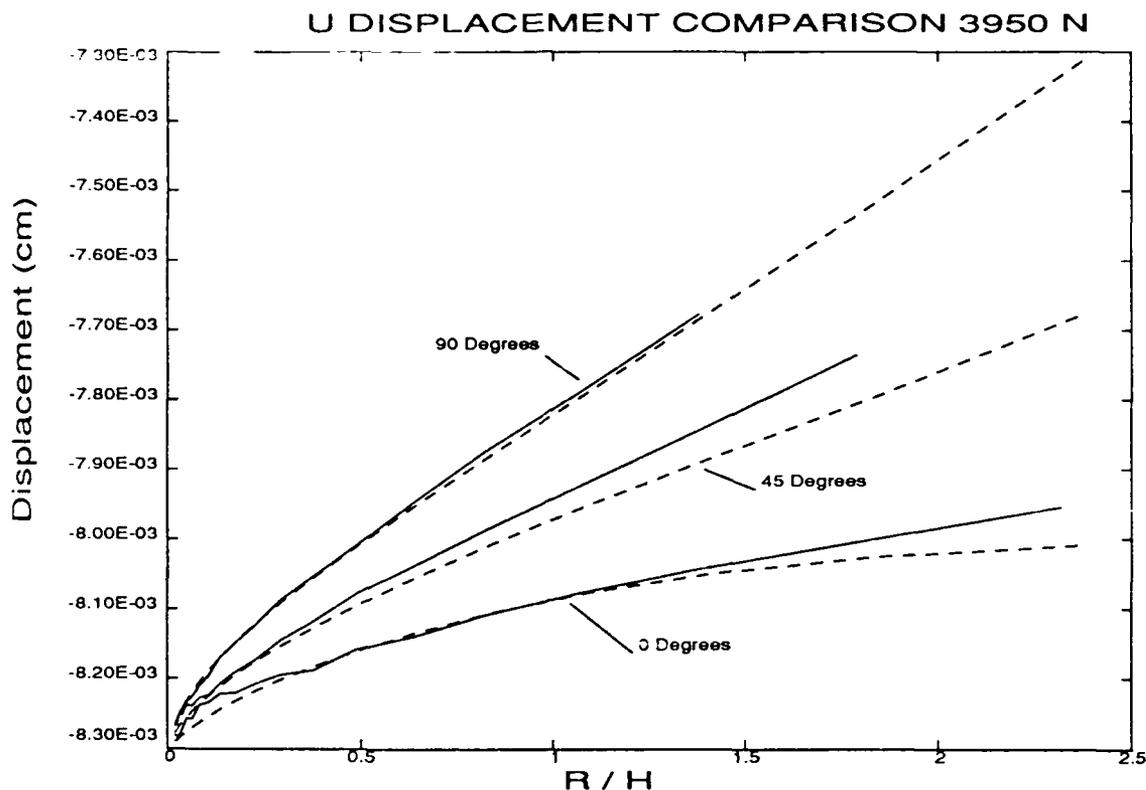


FIGURE 57. Comparison of the U displacement from the experiment and the numerical model. The experimental data is shown in solid lines, the numerical data is shown in dashed lines.

Comparison of the experimental and numerical displacements are shown in Figures 57 through 64. The displacements are plotted as a function of r/h , the ratio of the radius from the crack tip to the thickness of the plate. These plots are made at angular increments of 45 degrees, where the angle is measured from the positive x axis ahead of the crack. Since the experimental data is measured from fringes which only determine relative displacement, the experimental data was forced to match the numerical data on the x axis at r/h of 0.8. This offset value is kept constant for all the data.

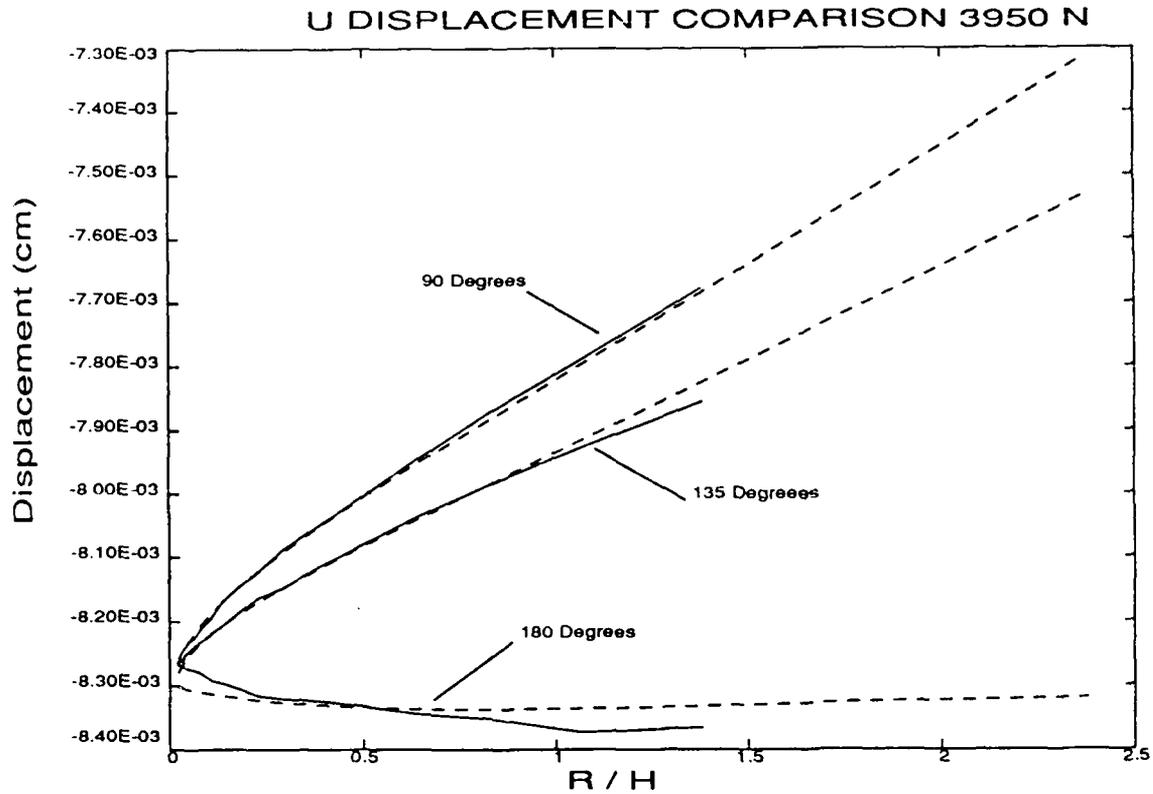


FIGURE 58. Comparison of the U displacement from the experiment and the numerical model. The experimental data is shown in solid lines, the numerical data is shown in dashed lines.

It can be seen from the figures that the U component of displacement matches well except along the line $\theta = 45^\circ$, and along $\theta = 0^\circ$ for $r/h > 1.5$. Again, this is most likely due to a slight misalignment of the saddle point. The noise in the experimentally determined displacements is most visible for $\theta = \pm 180^\circ$. In the comparison of the V displacement components, the noise in the experimental measurements is quite evident near the crack tip.

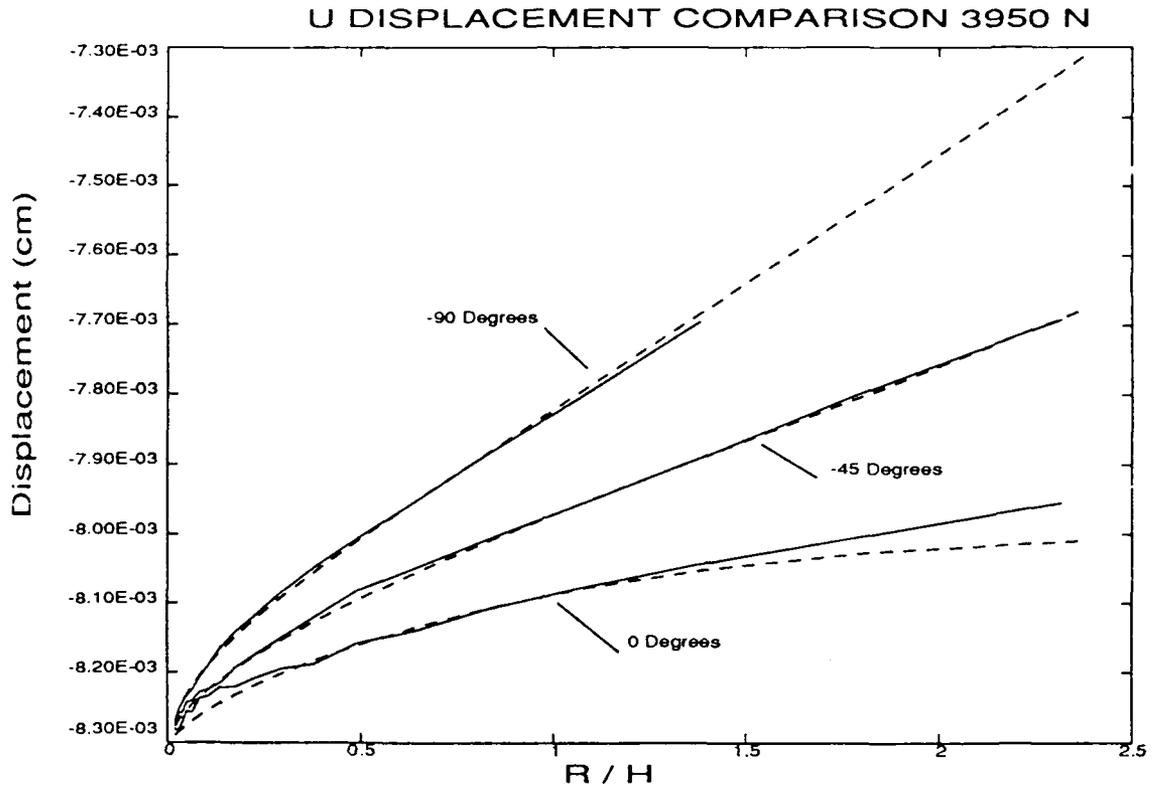


FIGURE 59. Comparison of the U displacement from the experiment and the numerical model. The experimental data is shown in solid lines, the numerical data is shown in dashed lines.

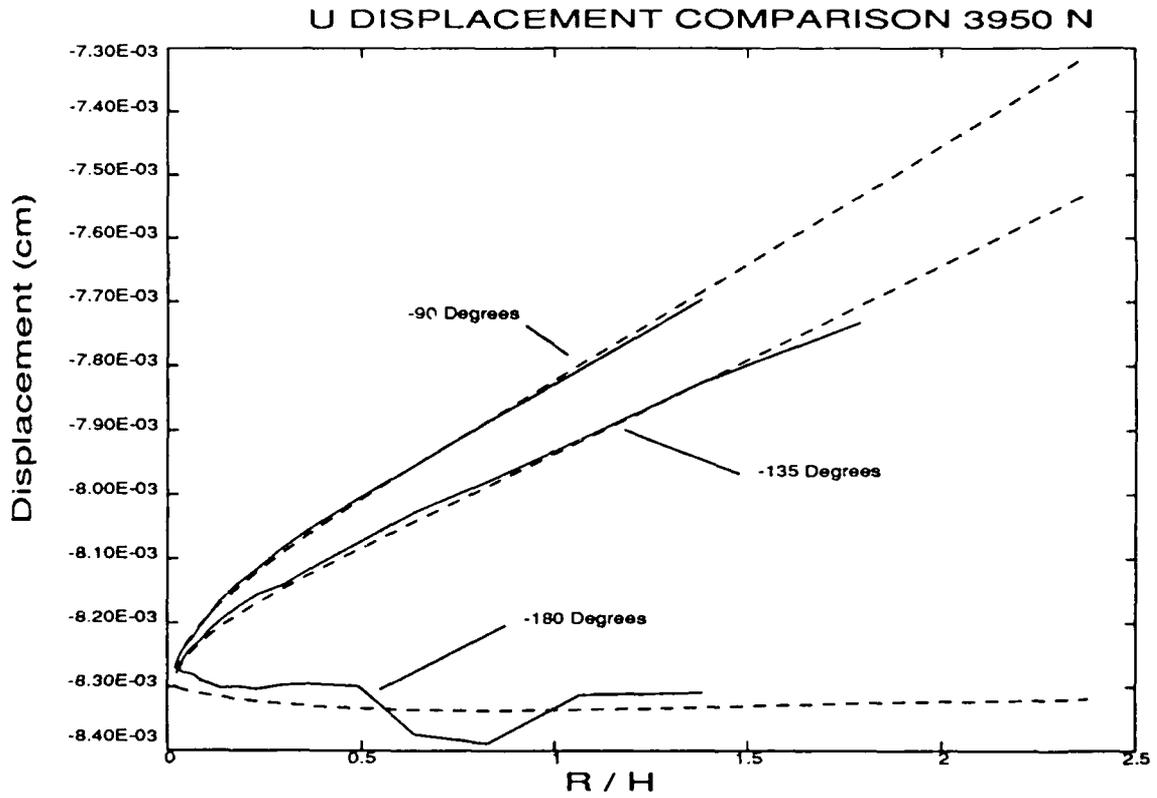


FIGURE 60. Comparison of the U displacement from the experiment and the numerical model. The experimental data is shown in solid lines, the numerical data is shown in dashed lines.

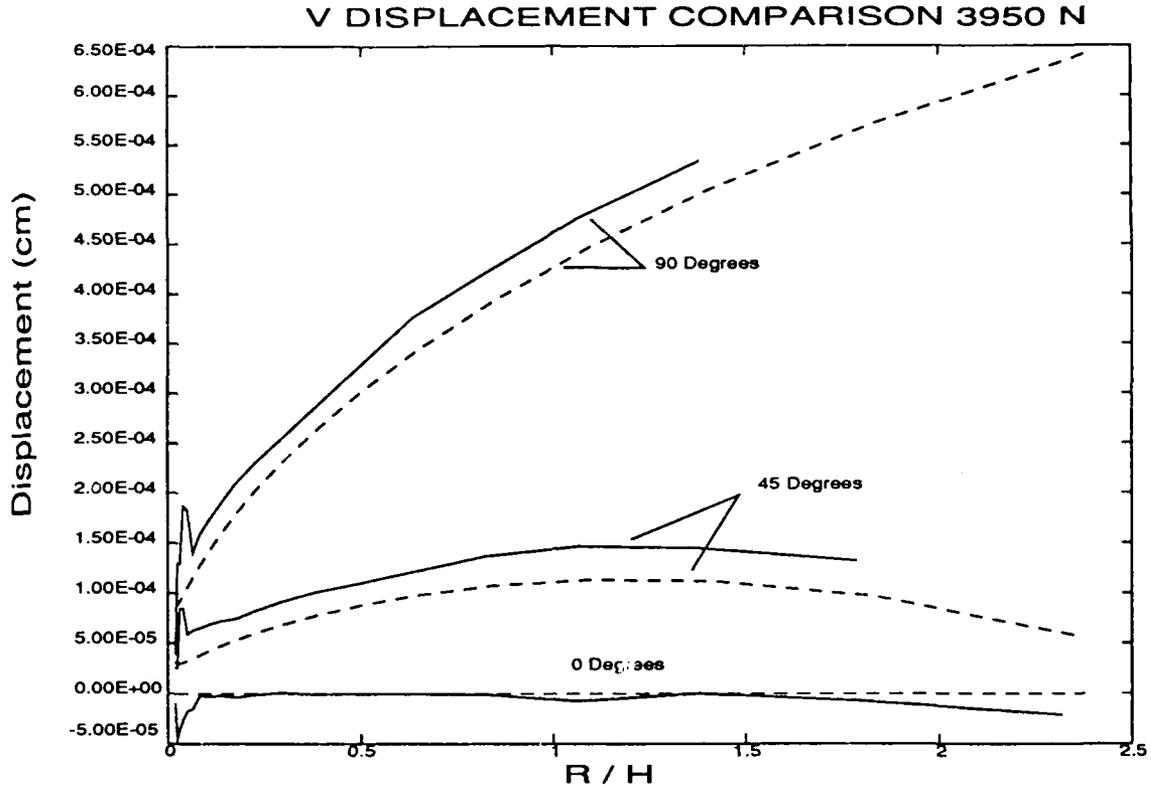


FIGURE 61. Comparison of the V displacement from the experiment and the numerical model. The experimental data is shown in solid lines, the numerical data is shown in dashed lines.

The agreement for the V displacements does not appear as good except along $\theta = \pm 180^\circ$. This is also most likely the result of a slight misalignment of the saddle point in the experiment. The shapes of the curves agree quite well, however, as in the case of $\theta = \pm 45^\circ$, where both the experiment and numerical model show first a deviation from zero displacement and then a trend back toward zero displacement as r increases.

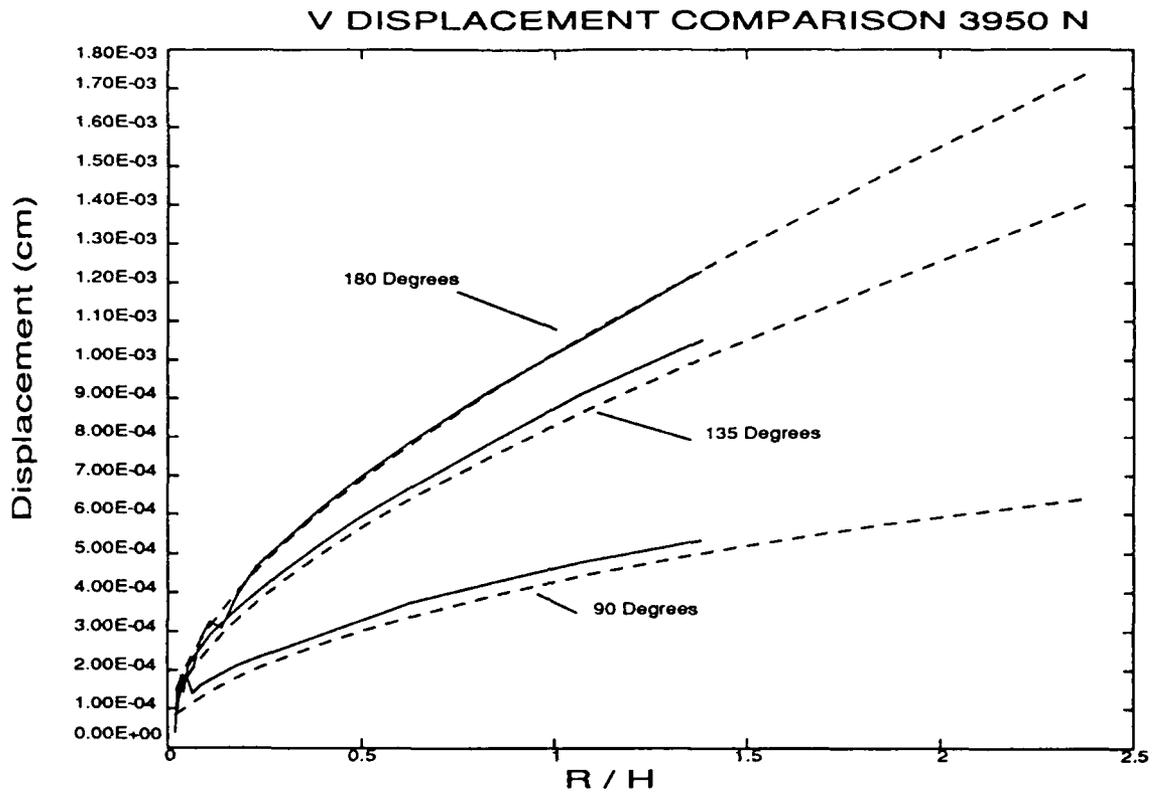


FIGURE 62. Comparison of the V displacement from the experiment and the numerical model. The experimental data is shown in solid lines, the numerical data is shown in dashed lines.

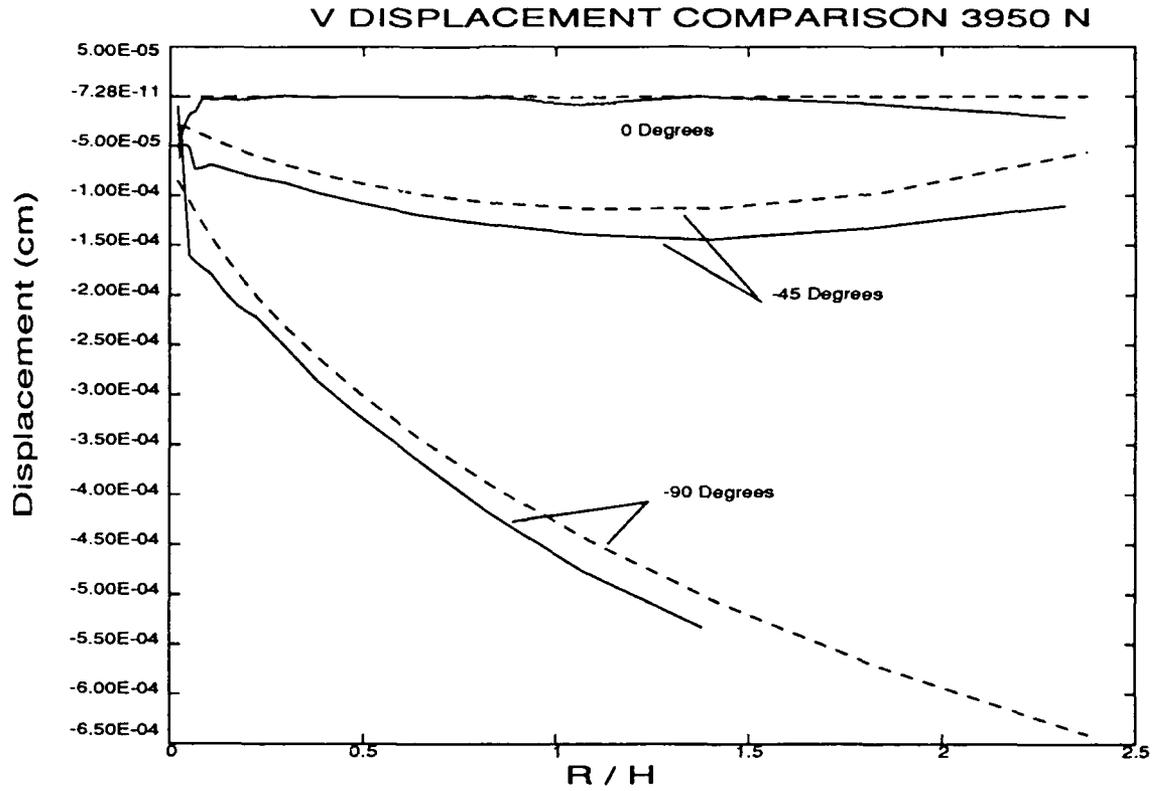


FIGURE 63. Comparison of the V displacement from the experiment and the numerical model. The experimental data is shown in solid lines, the numerical data is shown in dashed lines.

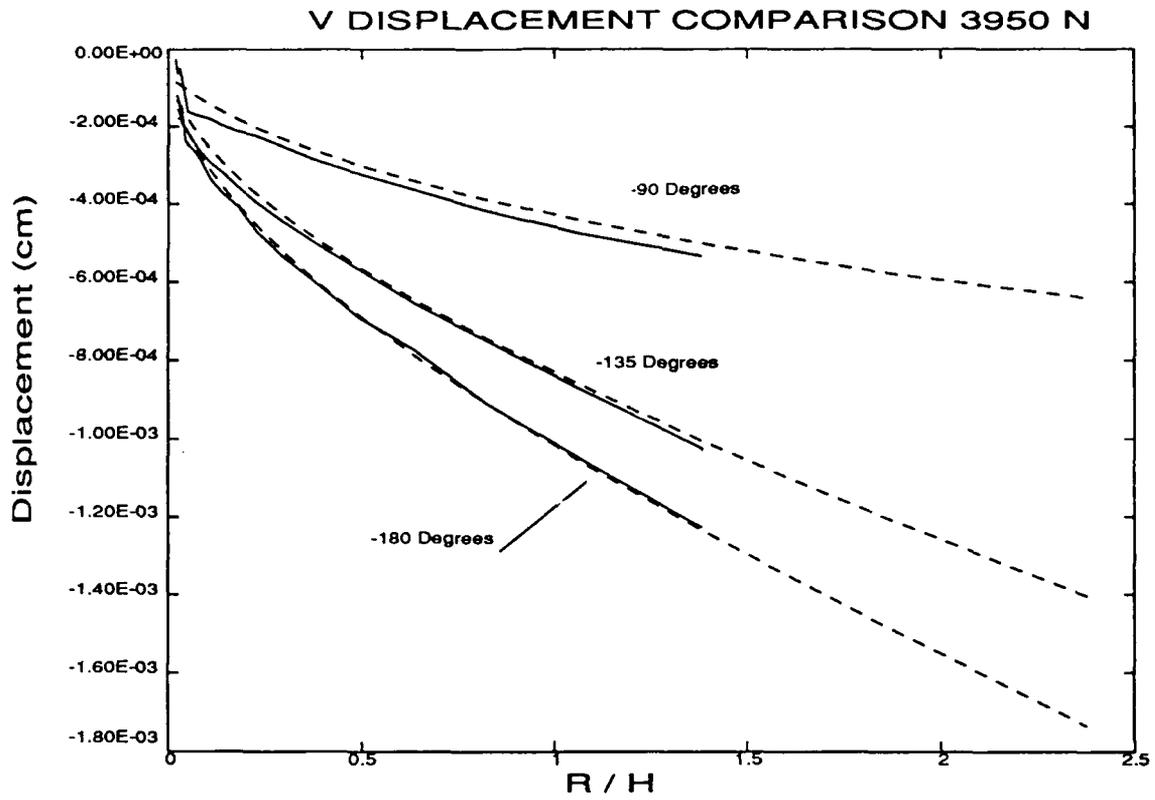


FIGURE 64. Comparison of the V displacement from the experiment and the numerical model. The experimental data is shown in solid lines, the numerical data is shown in dashed lines.

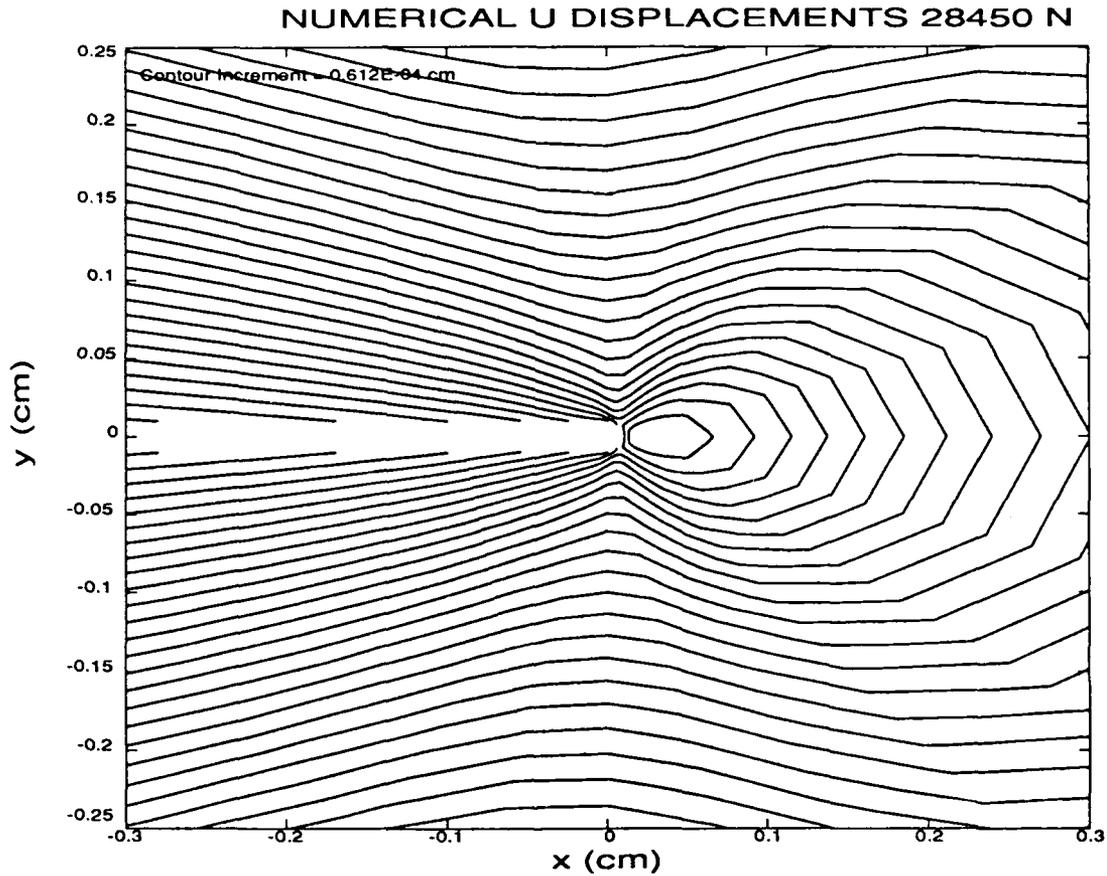


FIGURE 65. Contours of constant U displacement at the free surface at 28450 N. These contours are generated from the three dimensional finite element model. The crack lies along the line $x < 0$, with the crack tip at $x = 0$, $y = 0$. The increment of displacement between contours is 6.12×10^{-5} cm.

Numerically generated in-plane displacement contours are shown for a load level of 28450 N are shown in Figures 65 and 66. Note that the field of view in these figures is much smaller than the view in Figures 51 and 52. The finite shape of the notch is evident in Figures 65 and 66.

The equivalent contours as measured from the experiment are shown in Figures 67 and 68. Again, there is considerable noise in the measured displacement field near

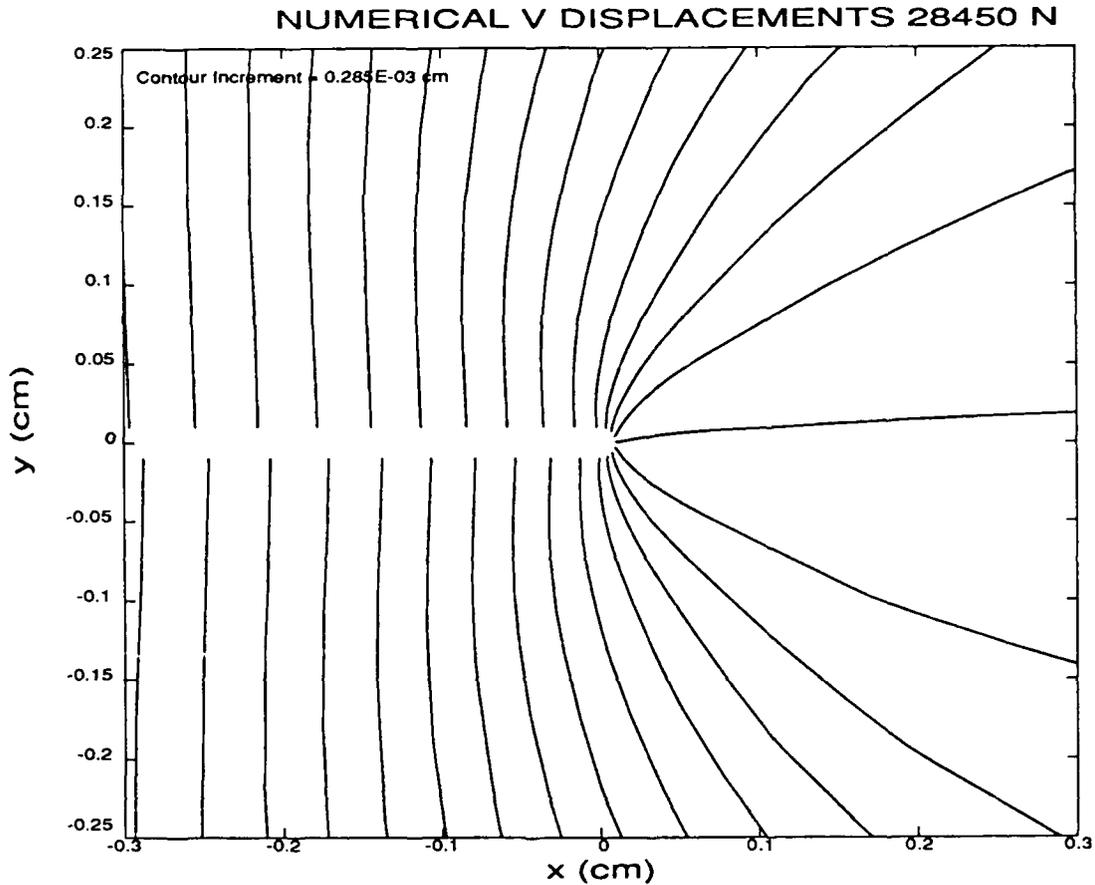


FIGURE 66. Contours of constant V displacement at the free surface at 28450 N. These contours are generated from the three dimensional finite element model. The crack lies along the line $x < 0$, with the crack tip at $x = 0$, $y = 0$. The increment of displacement between contours is 2.85×10^{-4} cm.

the crack tip. This is the largest load step for which there is essentially no shadow spot in the experimental photographs, and the data is taken from the pictures with the highest magnification and are blown up approximately twenty times for printing. The reproduction of the negative by the enlarger to paper for scanning is not particularly good, owing to the limitations of the enlarger. In order to produce a better print, the negative will have to be blown up through a microscope and then printed.

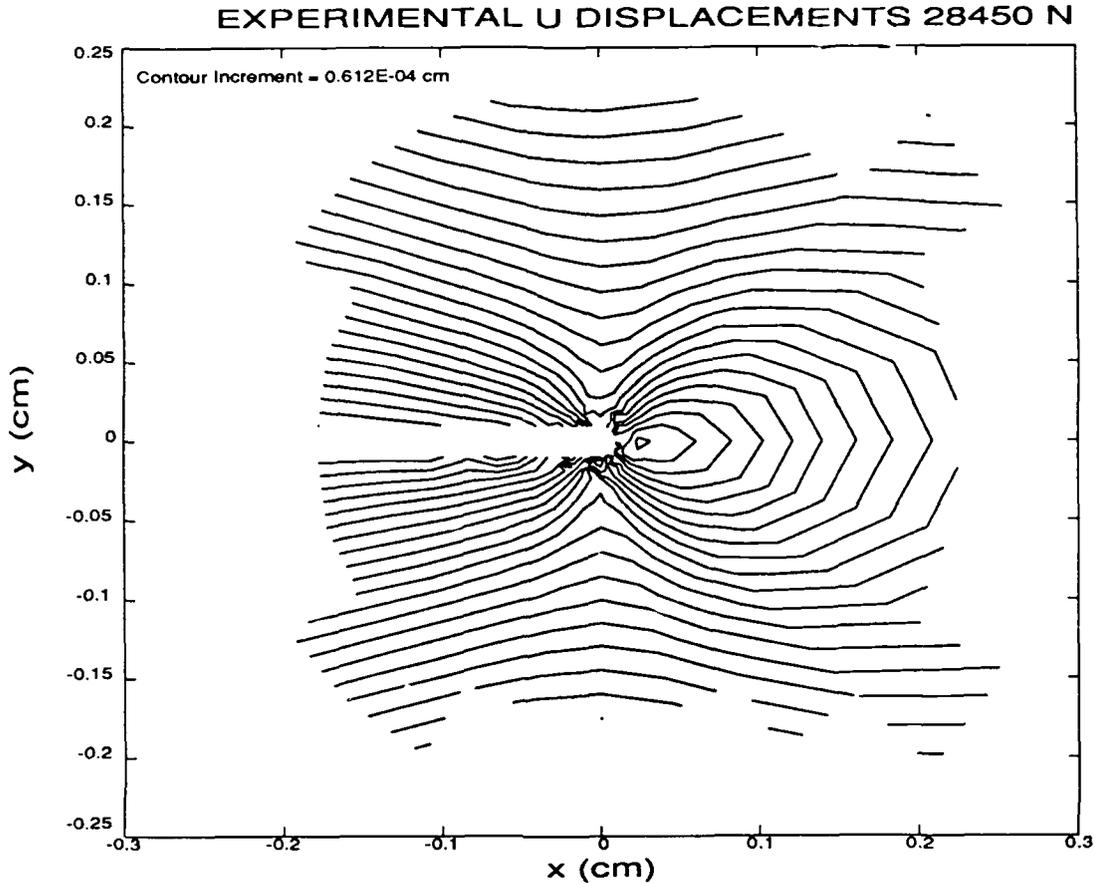


FIGURE 67. Contours of constant U displacement at the free surface at 28450 N. These contours are measured from the experiment. The contour increment is the same as in Figure 65, which shows contours from the numerical model. The crack lies along the line $x < 0$, with the crack tip at $x = 0, y = 0$. The increment of displacement between contours is 6.12×10^{-5} cm.

The contours from the experiment and the numerical model are shown together in Figures 69 and 70. The U displacement contours appear to match quite well in shape and spacing except in regions near $\theta = \pm 90^\circ$. It is interesting that there is a closed loop in the U displacement field at the tip of the crack in both the experimental and numerical contours, as this implies that the displacement is not monotonic, as would be expected from either the linearly elastic or the HRR analytic solutions. The match of the V displacement contours does not appear as good from

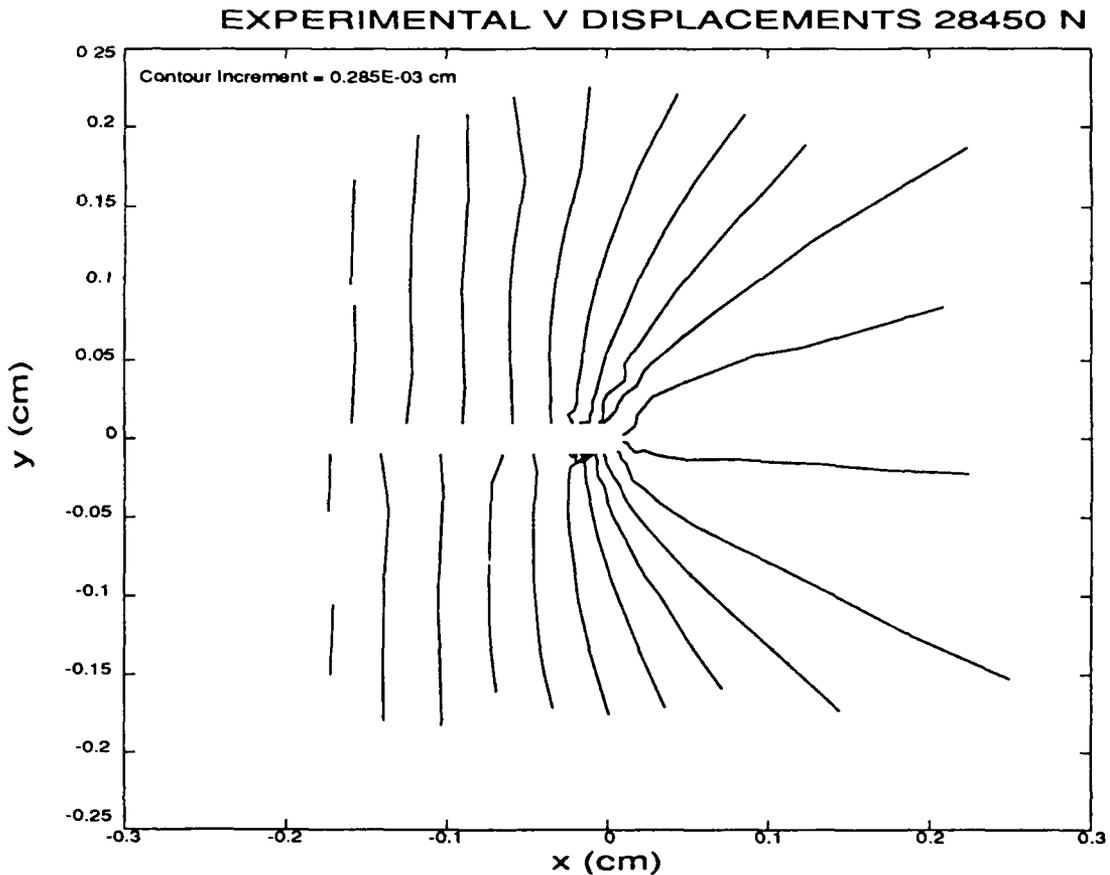


FIGURE 68. Contours of constant V displacement at the free surface at 28450 N. These contours are measured from the experiment. The contour increment is the same as in Figure 66, which shows contours from the numerical model. The crack lies along the line $x < 0$, with the crack tip at $x = 0, y = 0$. The increment of displacement between contours is 2.85×10^{-4} cm.

Figure 70, although do seem to correspond approximately in shape.

At 28450 N, the plastic zone is on the order of 0.2 cm surrounding the crack tip. This is the highest load for which there is no shadow spot visible in the experimental photographs. The ultimate failure load of the specimen is about 75000 N.

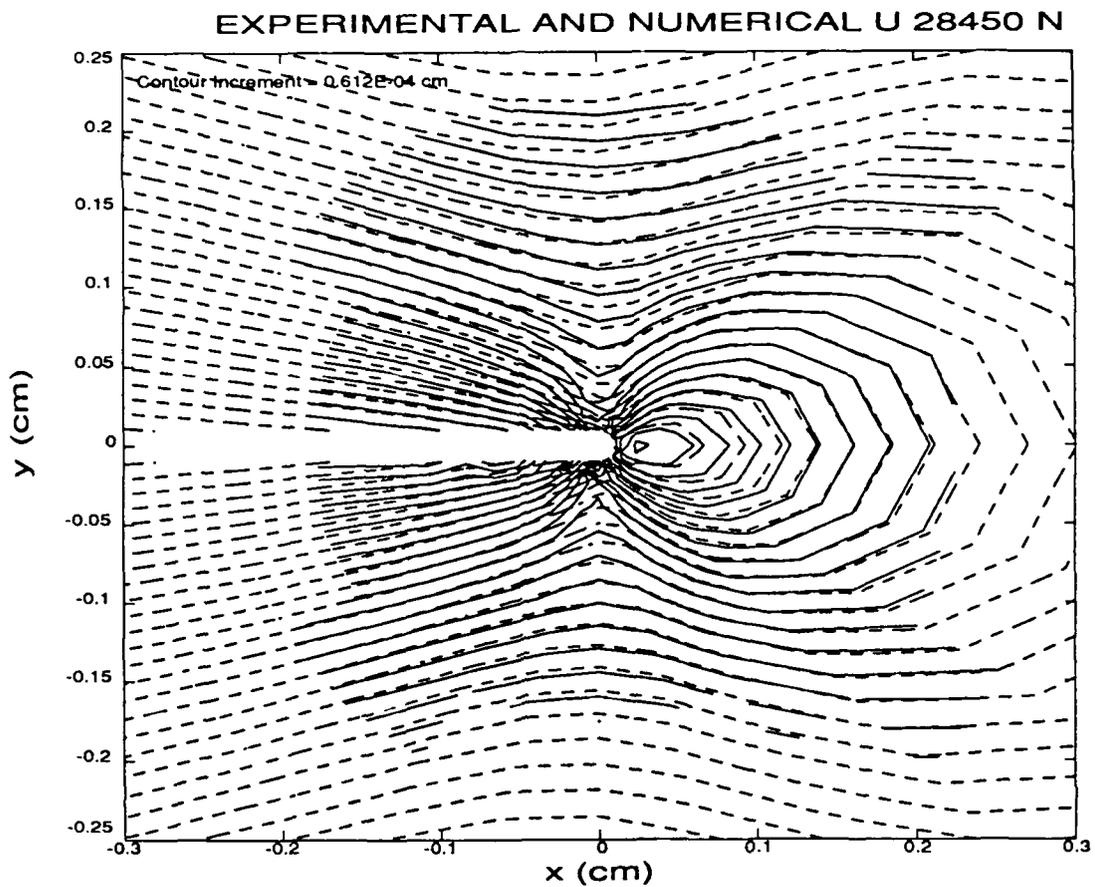


FIGURE 69. Contours of constant U displacement at the free surface at 28450 N. The contours measured from the experiment are solid lines, and the contours from the numerical model are dashed lines. The increment of displacement between contours is 6.12×10^{-5} cm.

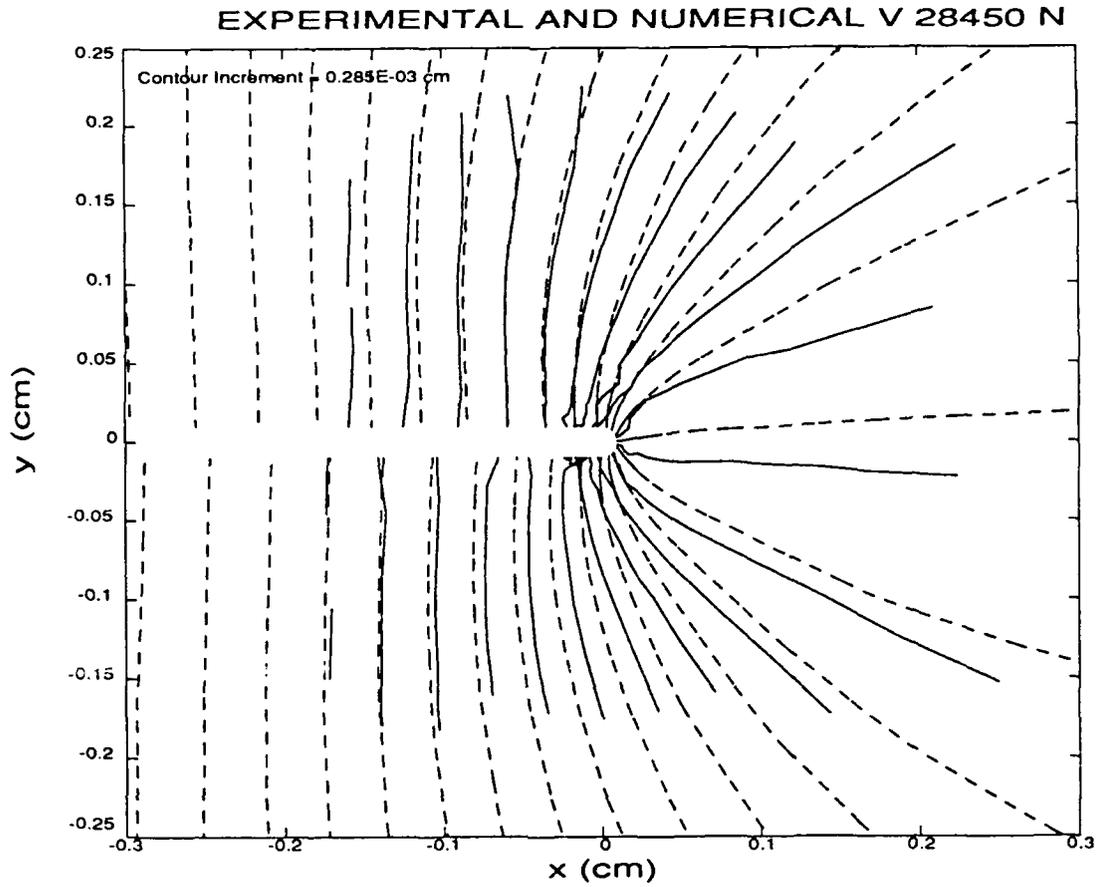


FIGURE 70. Contours of constant V displacement at the free surface at 28450 N. The contours measured from the experiment are solid lines, and the contours from the numerical model are dashed lines. The increment of displacement between contours is 2.85×10^{-4} cm.

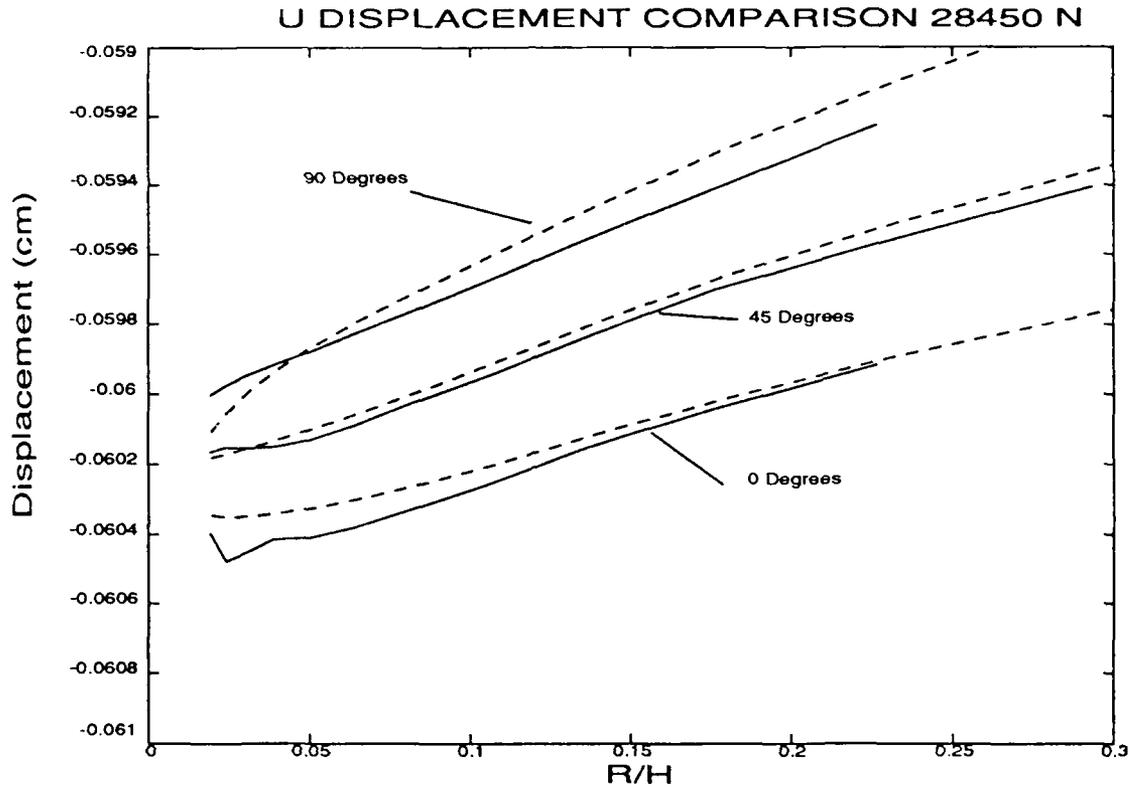


FIGURE 71. Comparison of the U displacement from the experiment and the numerical model. The experimental data is shown in solid lines, the numerical data is shown in dashed lines.

Comparison of the experimental and numerical displacements for the load of 28450 N are shown in Figures 71 through 78. The displacements are plotted as a function of r/h , the ratio of the radius from the crack tip to the thickness of the plate. These plots are made at angular increments of 45 degrees, where the angle is measured from the positive x axis ahead of the crack. Since the experimental data is measured from fringes which only determine relative displacement, the experimental data was forced to match approximately the numerical data on the x axis at r/h of 0.2. This offset value is kept constant for all the data.

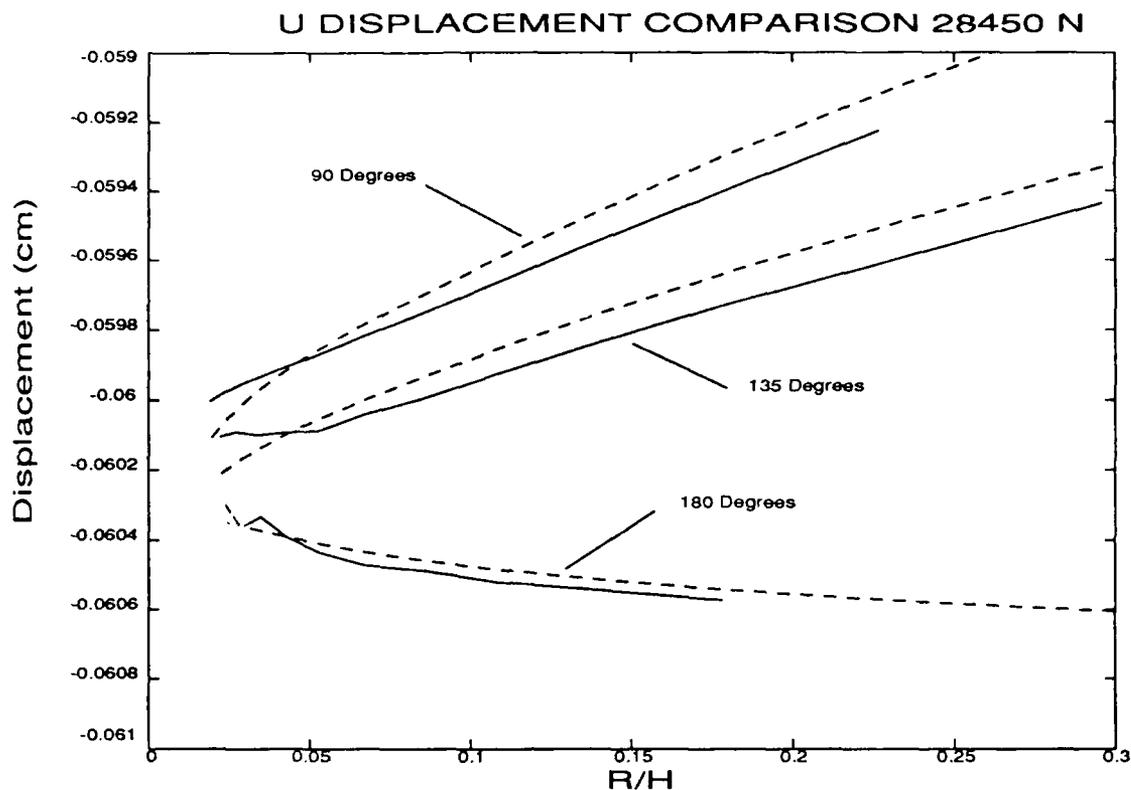


FIGURE 72. Comparison of the U displacement from the experiment and the numerical model. The experimental data is shown in solid lines, the numerical data is shown in dashed lines.

The data for the U displacements seem to match quite well away from the crack, except along the lines $\theta = 90^\circ$ and $\theta = 135^\circ$. Again the noise in the experimental measurement is evident in the vicinity of the crack tip. Again, although the match is not exact, the trends of the data are similar for the numerical model and the experiment. Two features in the U displacement field which may be important are seen along $\theta = 0^\circ$ and $\theta = \pm 90^\circ$. Along $\theta = 0^\circ$, there is a local minimum away from the crack tip, which gives rise to the closed loop contours noted above. Along $\theta = \pm 90^\circ$, it is seen that the experimental data does not drop off as the crack tip is approached in the same way that the numerical data does. This may also be

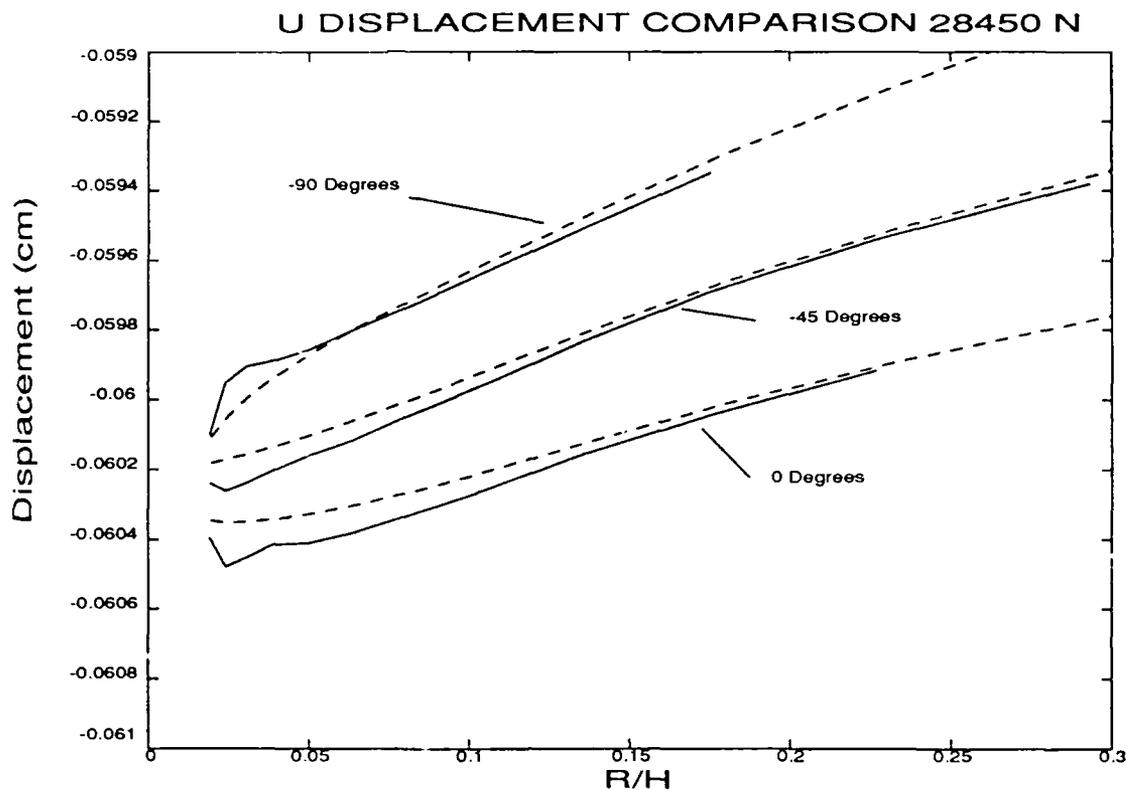


FIGURE 73. Comparison of the U displacement from the experiment and the numerical model. The experimental data is shown in solid lines, the numerical data is shown in dashed lines.

seen for $\theta = \pm 135^\circ$. It will be interesting to see if these features are apparent for comparisons at other loads.

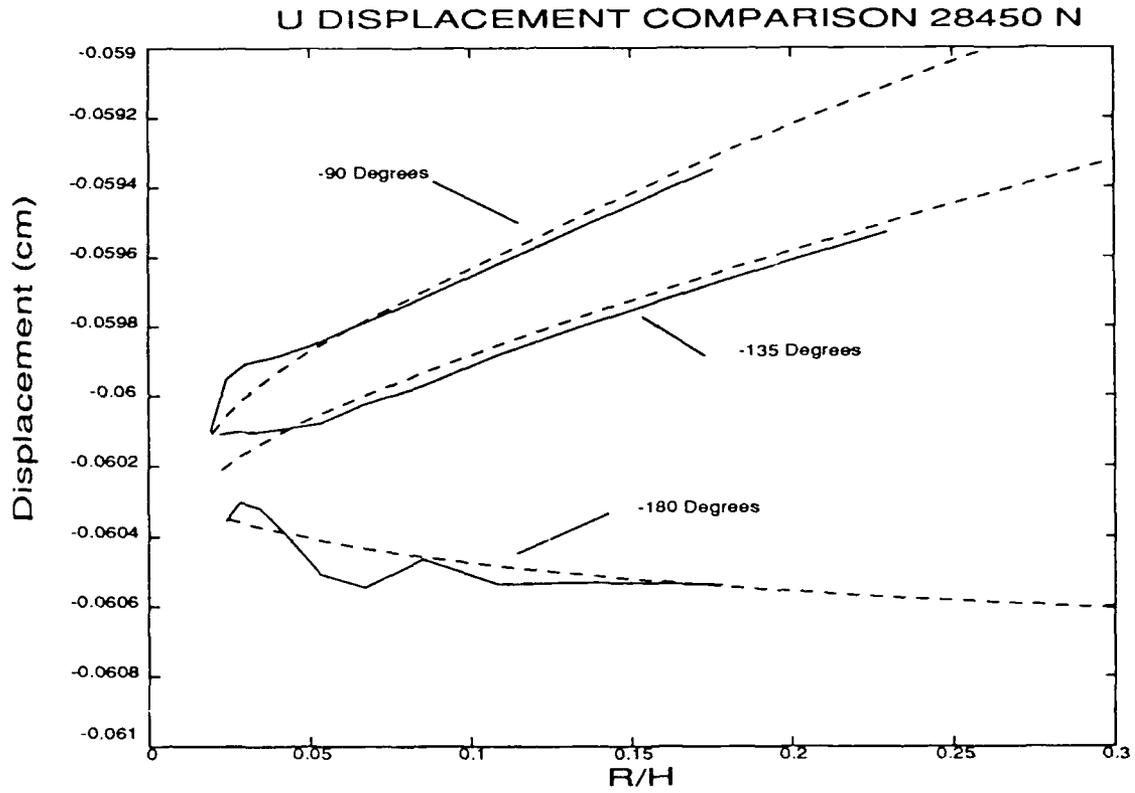


FIGURE 74. Comparison of the U displacement from the experiment and the numerical model. The experimental data is shown in solid lines, the numerical data is shown in dashed lines.

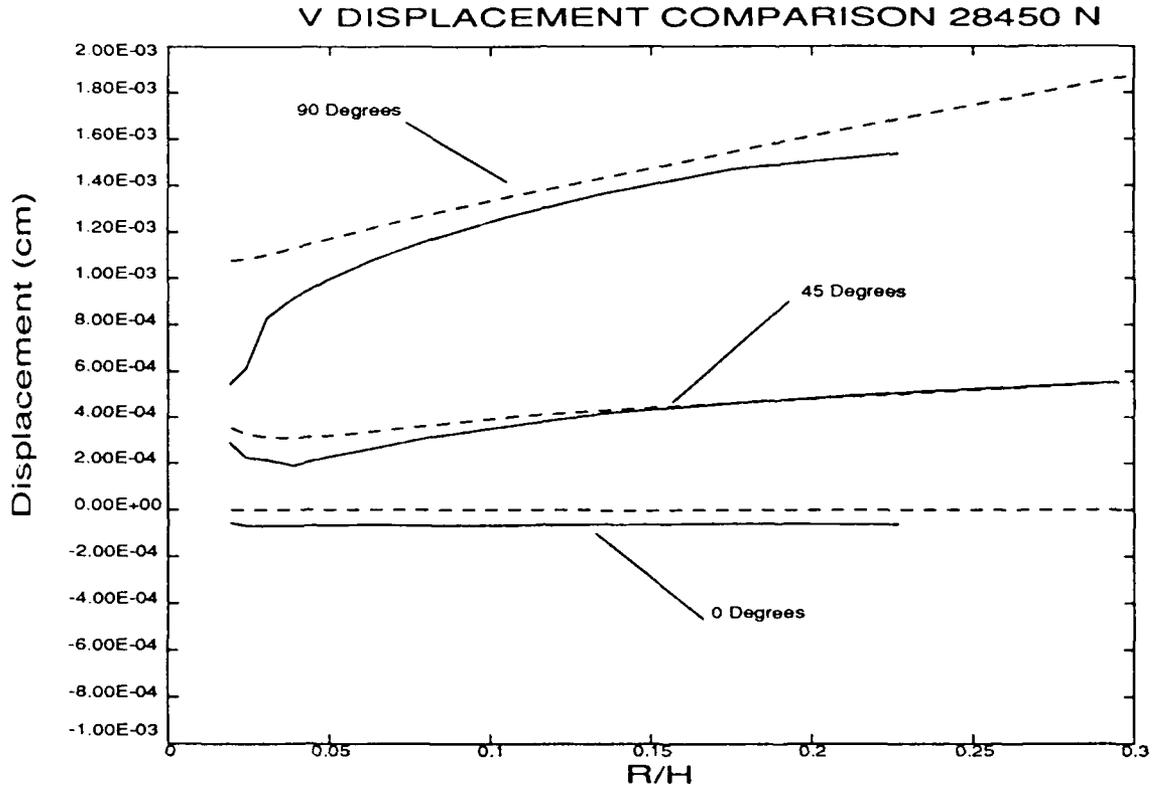


FIGURE 75. Comparison of the V displacement from the experiment and the numerical model. The experimental data is shown in solid lines, the numerical data is shown in dashed lines.

The data for the V displacements seem to match well in the region ahead of the crack as r/h increases. It is interesting that along $\theta = \pm 45^\circ$ there is a local minimum (or maximum) away from the crack, as in the U displacement field. Again, this is not predicted in either the elastic K -field solution or the HRR field solution. This feature is apparent in both the experimental and the numerical data. As in the U displacements, the largest deviations between the two data sets occurs along $\theta = \pm 90^\circ$ near the crack.

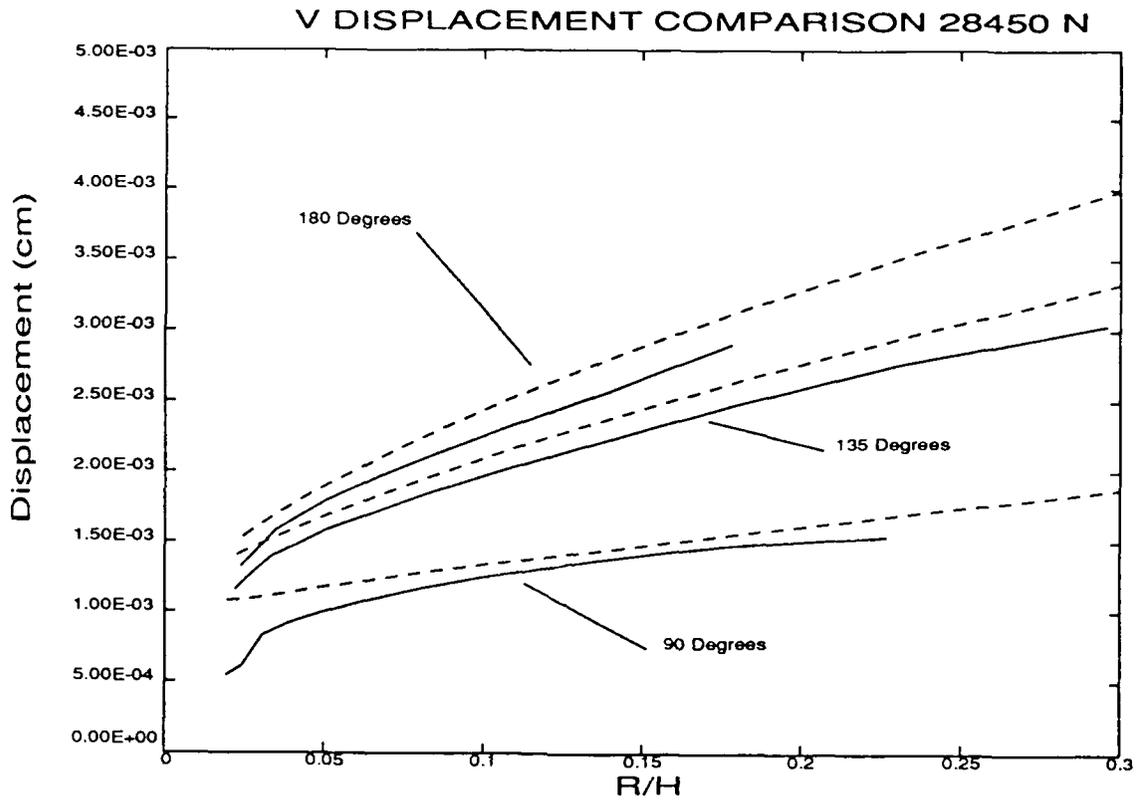


FIGURE 76. Comparison of the V displacement from the experiment and the numerical model. The experimental data is shown in solid lines, the numerical data is shown in dashed lines.

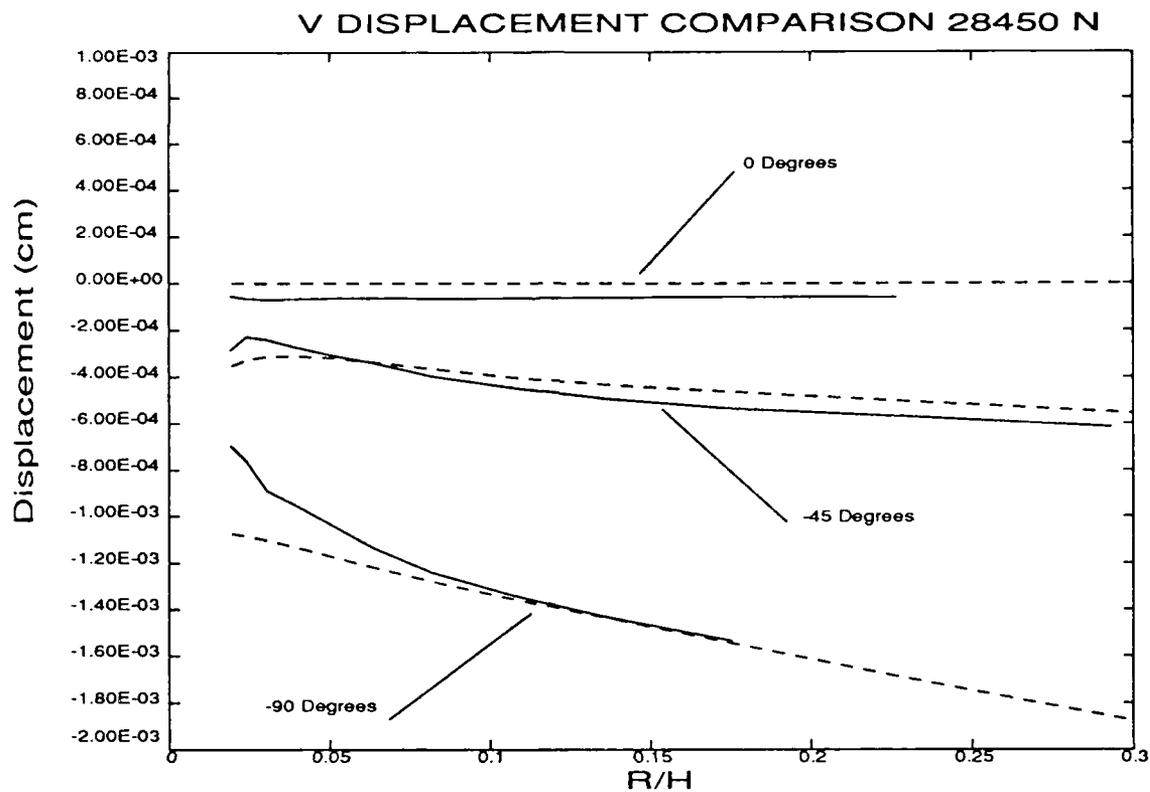


FIGURE 77. Comparison of the V displacement from the experiment and the numerical model. The experimental data is shown in solid lines, the numerical data is shown in dashed lines.

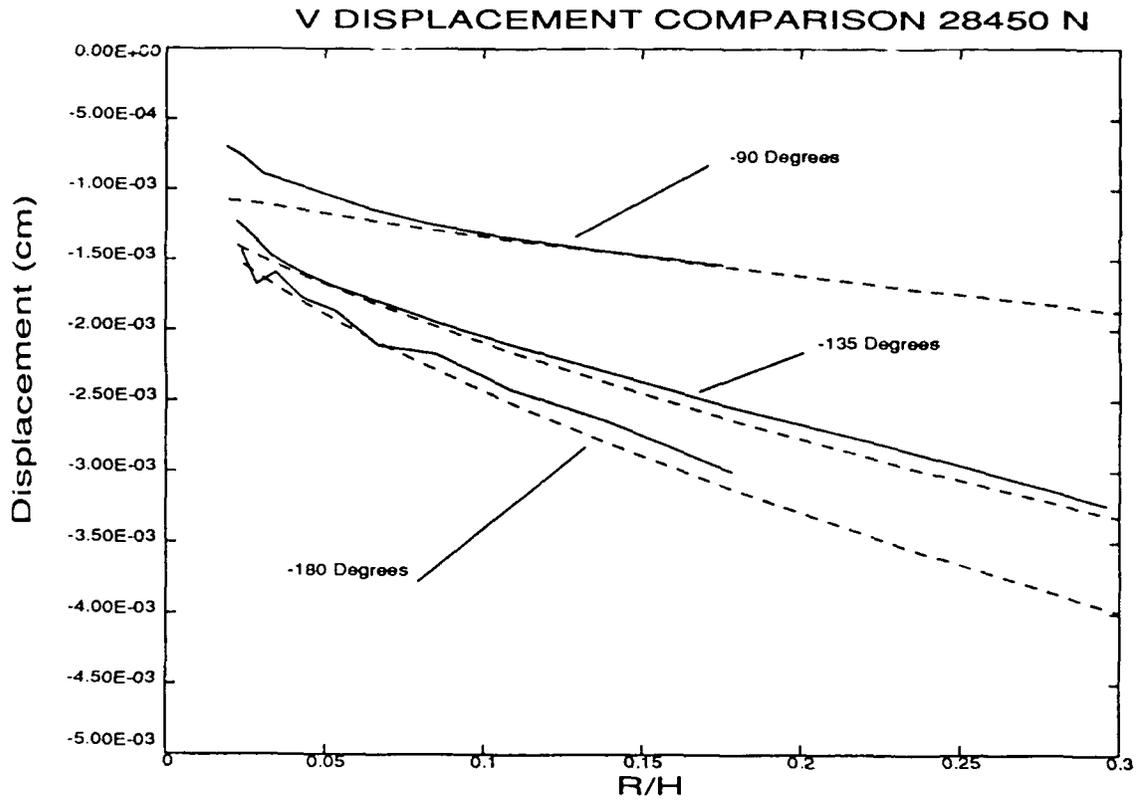


FIGURE 78. Comparison of the V displacement from the experiment and the numerical model. The experimental data is shown in solid lines, the numerical data is shown in dashed lines.

These two load values (3950 N and 28450 N) were chosen as the lowest and highest loads at which data could be acquired over the entire field. Unfortunately, the experimental data for these loads has two different problems which make comparison difficult. For the lower load, the sparse fringe pattern made it difficult to align the saddle points of the displacement fields, and any misalignment would represent a considerable discrepancy in the measured displacements. At the higher load, the high fringe density made it difficult to analyze anything except a very small field of view, and the difficulty in enlarging the photograph made that analysis more troublesome than it should have been. The next step will be to investigate the data available from the second lowest and second highest load levels, to see if the features apparent in this data, particularly for the higher load, are also visible at other loads.

A planned second set of experimental measurements will push the data collection optics so that data from higher load levels will be captured, at least in the region surrounding the crack tip.

References

1. Bathe, K-J., *Finite Element Procedures in Engineering Analysis*, Prentice-Hall, Inc., Englewood Cliffs, N.J., 1982.
2. Durelli, A.J. and Parks, V.J., *Moire Analysis of Strain*, Prentice-Hall, Inc., Englewood Cliffs, N.J., 1970.
3. Guild, J., *The Interference Systems of Crossed Diffraction Gratings*, Oxford at the Clarendon Press, 1956.
4. Guild, J., *Diffraction Gratings as Measuring Scales*, Oxford University Press, London, 1960.
5. Hughes, T.J.R., "Generalization of Selective Integration Procedures to Anisotropic and Nonlinear Media," *International Journal for Numerical Methods in Engineering*, Vol. 15, pp. 1413-1418, 1980.
6. Hutchinson, J.W., "Plastic Stresses and Strain Fields at a Crack Tip," *Journal of Mechanics and Physics of Solids*, Vol. 16, pp. 337-347, 1968.
7. KODAK Pamphlet P-110, *KODAK Materials for Holography*, Eastman Kodak Company, Rochester, New York, 1976.
8. KODAK Pamphlet P-255, *KODAK Technical Pan Film 2415/6415*, Eastman Kodak Company, Rochester, New York, 1985.
9. Livnat, A. and Post, D., "The Governing Equations of Moire Interferometry and Their Identity to Equations of Geometrical Moire," *Experimental Mechanics*, Vol. 25, No. 4, 1985.

10. Matthies, H. and Strang, G., "The Solution of Nonlinear Finite Element Equations," *International Journal for Numerical Methods in Engineering*, Vol. 14, pp. 1613-1626, 1979.
11. McDonach, A., McKelvie, J. and Walker, C.A., "Stress Analysis of Fibrous Composites Using Moire Interferometry," *Optics and Lasers in Engineering*, Vol. 1, pp. 85-105, 1980.
12. Nakamura, T., Shih, C.F. and Freund, L.B., "Analysis of a Dynamically Loaded Three-point-bend Ductile Fracture Specimen," *Engineering Fracture Mechanics*, Vol. 25, pp. 323-339, 1986.
13. Narasimhan, R. and Rosakis, A.J., *Influence of Hardening on Quasi-Static Extension of Mode I, Plane Stress Cracks in Isotropic Elastic-Plastic Materials*, SM Report 86-35, California Institute of Technology Report, Pasadena, 1986.
14. Narasimhan, R. and Rosakis, A.J., "A Finite Element Analysis of Small-Scale Yielding Near a Stationary Crack Under Plane Stress," *Journal of Mechanics and Physics of Solids*, Vol. 36, No. 1, pp. 77-117, 1988.
15. Narasimhan, R. and Rosakis, A.J., *Three Dimensional Effects Near a Crack Tip in a Ductile Three Point Bend Specimen, Part I: A Numerical Investigation*, SM Report 88-6, California Institute of Technology Report, Pasadena, January, 1988.
16. Pirodda, L., "Strain Analysis by Grating Interferometry," *Optics and Lasers in Engineering*, Vol. 5, No. 1, 1984.
17. Post, D., "Sharpening and Multiplication of Moire Fringes," *Experimental Mechanics*, Vol. 7, No. 4, April, 1967.

18. Post, D., "Moire Interferometry at VPI & SU," *Experimental Mechanics*, Vol. 23, No. 2, pp. 203-210, June, 1983.
19. Post, D. and Baracat, W.A., "High Sensitivity Moire Interferometry — A Simplified Approach," *Experimental Mechanics*, Vol. 21, No. 3, March, 1981.
20. Rice, J.R. and Rosengren, G.F., "Plane Strain Deformation Near a Crack Tip in a Power-law Hardening Material," *Journal of Mechanics and Physics of Solids*, Vol. 16, pp. 1-12, 1968.
21. Schultheisz, C., *Moire Interferometry*, SM Report 84-4, California Institute of Technology Report, Pasadena, 1984.
22. Shih, C.F., *Tables of Hutchinson-Rice-Rosengren Singular Field Quantities*, MRL E-147, Materials Research Laboratory Report, Brown University, Providence, 1983.
23. Theocaris, P.S., *Moire Fringes in Strain Analysis*, Pergamon Press, Oxford, 1969.
24. Tada, H., Paris, P.C., and Irwin, G.R., *The Stress Analysis of Cracks Handbook*, Del Research Corporation, Hellertown, Pennsylvania, 1973.
25. Vest, Charles M., *Holographic Interferometry*, John Wiley & Sons, New York, 1979.
26. Williams, M.L., "On the Stress distribution at the Base of a Stationary Crack," *Journal of Applied Mechanics*, Vol. 24, pp. 109-114, 1957.
27. Zehnder, A.T., "Dynamic Fracture Initiation and Propagation in Metals: Experimental Results and Techniques," Ph.D. Thesis, California Institute of Technology, 1987.

28. Zehnder, A.T. and Rosakis, A.J., *Three Dimensional Effects Near a Crack Tip in a Ductile Three Point Bend Specimen, Part II: An Experimental Investigation Using Interferometry and Caustics*, SM Report 88-7, California Institute of Technology Report, Pasadena, March, 1988.