

LOUGHBOROUGH
UNIVERSITY OF TECHNOLOGY
LIBRARY

AUTHOR/FILING TITLE

CLARK, S E

ACCESSION/COPY NO.

020888/02

VOL. NO.

CLASS MARK

LOAN COPY

30 JUN 1989

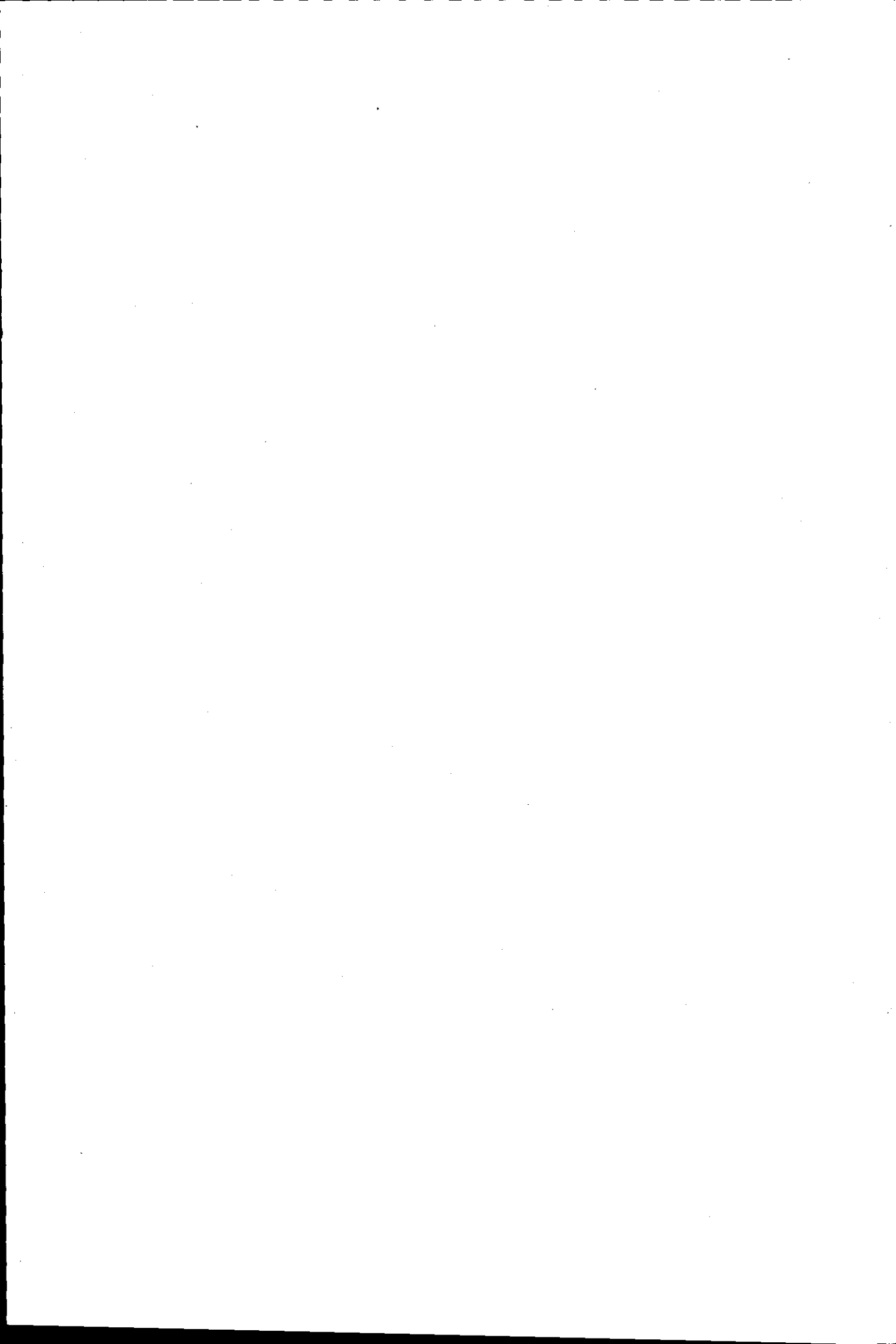
~~5 JUL 1990~~

~~6 JUL 1990~~

25 JUN 1999

002 0888 02





Excimer Laser Induced Modifications of Optical Surfaces

by

Stuart Eric Clark

A Doctoral Thesis

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

April 1988



by S.E.Clark 1988

Loughborough University	
of Techn. Library	
Date	Aug 88
Class	
Acc. No.	020888/02

Acknowledgements

I would like to thank all those involved with this project for their help and encouragement. In particular I would like to thank my supervisor David Emmony for his guidance and for the many fruitful discussions that took place.

Thanks are also due to all of my colleagues, in particular to Ian Laidler for his help with the setup of experimental equipment and to Alan Holloway with whom many useful conversations took place. I would also like to thank Basil Omar and Noel Kerr for respectively providing background information on laser induced fluorescence and help with the vacuum related experiments.

The guidance and help provided by the University microscopy workers namely Anita Hood, John Bates and Frank Page is gratefully acknowledged.

Acknowledgement is also due to Dr. G.A. Brooker of Oxford University whose teaching of optics and related phenomena inspired me to this work.

Lastly and certainly not least, I wish to thank my parents for their unstinting support and faith in me.

Abstract

The work in this thesis is a study of the interactions of a well characterized high power ultra violet excimer laser beam with solid optical quality surfaces, principally Al and Ge but with limited work on dielectrics.

In order to characterize the beam and to monitor the changes induced on test surfaces by the beam, a computer controlled image processing system was developed. The system used the fluorescence from a UV absorbing glass to image the excimer beam and Schlieren imaging of the test surfaces with cw HeNe or pulsed dye laser illumination to monitor the induced changes.

The imaging of the induced fluorescence enabled for the first time a 2 dimensional profile of a pulsed laser beam to be obtained in a single shot which, when subsequently processed would yield all the required spatial characteristics of the beam. The monitoring of test surfaces by Schlieren imaging allowed both accurate determination of the onset of permanent changes and more significantly, for the first time, correlation to be made between transient and permanent changes.

The characterised beam was used to study effects in 2 main areas...

- a) Laser Induced Periodic Surface Structures..LIPSS
- b) Early Detection of Laser Induced Damage....EDLID

a) LIPSS are permanent modifications to a surface essentially caused by interference between the electric field of the incident light and the fields induced on the test surface by the incident light which, can result in periodic "ripple" patterns being formed on the test surface. A detailed study of the spacing and morphology of the patterns as a function of various parameters including the angle of incidence and polarization of the incident light was made. A comparison of these results, results from other workers and those predicted by a new theory developed in Canada yielded not only new patterns and effects but, showed that theory and experiment were in excellent agreement.

b) EDLID as developed in this work is the process whereby a non damaging laser pulse is used to enable for the first time accurate predictions as to the spatial location of potential damage sites to be made. The technique is based upon imaging the transient modifications (heating) of a test surface by a non damaging laser pulse and the subsequent determination of areas of anomalous absorption which, it is argued are the areas most likely to damage. The sample is then damaged and a comparison of the areas that damaged and those of anomalous absorption i.e. the predicted areas is made. Excellent agreement is found between predicted and observed damaged areas.

A limited amount of work was performed on the determination and characterization of laser damage to a small number of novel dielectric coating materials based on the so called "sol-gel" technology.

Contents

	page
Acknowledgements	i
Abstract	ii
Dedication	ix
<u>1 Review of Laser Induced Damage (LID)</u>	1
1.1 General Introduction	1
1.2 Sources of Information on LID	3
1.3 Historical Review	4
1.4 Comment	5
1.5 The Role of Defects - a First View	5
1.6 LID in Metals	6
1.6.1 Multiple Pulse	6
1.6.2 Single Pulse	8
1.7 LID in Dielectrics	10
1.7.1 Introduction	10
1.7.2 Bulk Damage	11
1.7.3 Surface Damage	13
1.7.4 Absorbing Inclusion Damage	14
1.7.5 Multiple Pulse and Spot Size Effects	14
1.7.6 Front and Back Surface Effects	15
1.8 LID in Thin Film Coatings	15
1.9 Dependence of LID on Laser and Material Parameters	17
1.10 Laser Annealing	18
1.11 UV LID	19
1.12 Definition of	21
1.12a Fluence	21
1.12b Damage threshold	22
1.13 Photoacoustic and Photothermal Effects	23
1.14 Conclusions and comments	23
References	24

<u>2</u>	<u>Laser Beam Profiling and Damage detection</u>	26
2.1	Introduction	26
2.2	Review	26
	2.2a Laser Beam Profiling	26
	2.2b Damage Detection	29
2.3	Damage Facility	30
2.4	Beam Profiling - Initial technique	32
2.5	Beam Prfoiling Using a Video Framestore	32
	2.5.1 Introduction	32
	2.5.2 Video Basics	33
	2.5.3 Video Cameras	33
	2.5.4 Video Framestore	36
	2.5.5 Computer Control	37
	2.5.6 Timing Requirements for the System	37
	2.5.7 Video Copy Processor	37
	2.5.8 Noise Considerations	40
	2.5.9 Fluorescence and Fluorescers	40
	2.5.10 Operation of the System	41
	2.5.11 Computer Routines	43
	2.5.11a Computer Processing on the Apple IIe	43
	2.5.11b Data Aquisition/Processing Routines	44
	2.5.11c Profiling routines	44
	2.5.11d Statistical Routines	45
	2.5.12 Results	47
	2.5.12a 1	47
	2.5.12b 2	52
	2.5.12c 3	55
	2.5.12d 4	56
	2.5.13 Profiling and Damaging Simultaneously	56
	2.5.14 Comments	56
2.6	Damage Detection	58
	2.6.1 Introduction	58
	2.6.2 Schlieren Techniques	58
	2.6.3 Detection using the Framestore and Schlieren Techniques	59
	2.6.4 Damage Processing Routines	62
	2.6.5 Imaging and Alignment of the Sample	63
	2.6.6 Damage Testing	66
	2.6.7 Laser Pre-cleaning	66

2.6.8	Results	67
2.6.9	Comments	67
2.7	Conclusions	67
	References	70
3	<u>Laser Induced Ripple Structures</u>	71
3.1	Introduction	71
3.2	Laser Induced Thermomechanical Structures (LITS)	72
3.2.1	Introduction and Review	72
3.2.2	Experimental Configuration	75
3.2.3	Procedures for Analysis of Results	75
3.2.4	Samples Tested	77
3.2.5	Experimental Procedure	77
3.2.6	Results	77
3.2.7	Analysis of Results	78
3.2.8	Conclusions	78
3.3	Laser Induced Capillary Waves (LICW)	80
3.3.1	Introduction and Review	80
3.3.2	Samples Tested	82
3.3.3	Experimental Procedure	82
3.3.4	Comment	83
3.3.5	Results	83
3.3.6	Laser Heating theory	90
3.3.7	Analysis of Results	93
3.3.8	Conclusions	97
3.4	Laser Induced Periodic Surface Structures (LIPSS)	99
3.4.1	Introduction	99
3.4.2	Review of	101
3.4.2a	Theories	101
3.4.2b	Experimental Work	103
3.4.3	Coordinate System Used	106
3.4.4	Experimental Procedure	106
3.4.5	Results	108
3.4.6	Theory	117
3.4.6a	Introduction	117
3.4.6b	Spacing and Orientation	117
3.4.6c	Resolidification	122
3.4.6d	Energy of formation	125
3.4.7	Comparison of Theory and Experiment	127

3.4.7a Spacing and Orientation	127
3.4.7b Ripple Morphology	141
3.4.7c Formation Fluence	145
3.4.8 Conclusions	147
3.5 Anomalous LIPSS	147
3.5.1 Introduction and Review	147
3.5.2 Propagation in a Plasma Theory	147
3.5.3 Experimental Results	148
3.5.4 Analysis of Results	148
3.6 Comment	150
3.7 Overall Conclusions	150
References	152
<u>4 Early Detection of LID</u>	154
4.1 Introduction	154
4.2 Review	154
4.3 System Requirements	157
4.4 The Use of Interferometry	157
4.5 A Pulsed Laser Schlieren Imaging System	163
4.5.1 System Configuration	165
4.5.1a Hardware	165
4.5.1b Optics	173
4.5.1c Computing	174
4.5.2 Samples	178
4.5.2a Samples Used	178
4.5.2b Pre-imaging Cleaning of the Samples	178
4.5.3 Experiments Performed	178
4.5.4 Experimental Operation of the System	179
4.5.5 Selection of Excimer Fluence	180
4.5.6 Noise Considerations	180
4.5.7 Results 1	181
4.5.7a Al Single Shot	181
4.5.7b Al Multiple Shot	187
4.5.7c Ge Single Shot	194
4.5.7d Dielectric Single Shot	197
4.5.8 Time Dependence of Transient effects	201
4.5.8a Photomultiplier Monitoring	201
4.5.8b Image Monitoring	203
4.5.9 Results 2 - Transient Variations	203

4.6	The Physical Basis of the deflection of the Probe Beam	207
4.6.1	Damaging Fluences at Atmospheric Pressure	207
4.6.2	Beam Deflection Dependence on Pressure	208
4.6.3	Beam deflection theory	210
4.6.3a	Surface Temperature of the Metal	211
4.6.3b	Determination of the Air Temperature	212
4.6.3c	Equation of Optical Deflection	213
4.6.4	Results of Calculations	216
4.7	Conclusions	216
	References	220
<u>5</u>	<u>Colloidal Silica as a High Damage UV Antireflection Coating</u>	221
5.1	Introduction	221
5.2	Antireflection Theory	221
5.3	Electric Field Considerations	224
5.4	Review	225
5.5	Coatings Tested	226
5.6	Structure of Colloidal Silica Coatings	226
5.7	Damage Thresholds	227
5.8	Analysis and Comments on the Damage Results	227
5.9	The "Breath Test"	232
5.9.1	Introduction	232
5.9.2	Nucleation theory	233
5.9.3	Effects of Surface Properties on the Nucleation Rate	235
5.9.4	Results	235
5.9.4a	Observed Effects	236
5.9.4b	LIMA	236
5.9.4c	TEM	236
5.9.5	The "Breath Test" - Eplanation	243
5.10	Dielectric UV LIPSS	243
5.11	Conclusions	247
	References	248
<u>6</u>	<u>Conclusions and Recommendations</u>	249

To my parents

Chapter 1

Review of Laser Induced Damage

1.1 General Introduction

In the author's view four main questions need to be asked and answered. These are...

- 1) What is Laser Induced Damage (hereafter known as LID or damage)?
- 2) What causes LID?
- 3) What are the consequences of LID?
- 4) How can LID be avoided?

To an extent, answering these questions fully would take more than a thesis in itself however, this section will attempt to provide concise yet meaningful answers to them. Further details can be found in subsequent parts of this and following chapters.

- 1) The author defines LID to be..

"A detectable permanent and detrimental change in the optical properties of a sample caused by interaction of the sample with a laser beam."

A detailed discussion of exactly how the damage threshold i.e. the fluence (Jcm^{-2}) or power (Wcm^{-2}) required to cause LID is defined is undertaken later however, it will suffice at this stage to define the damage threshold as..

"The lowest fluence at which any area tested on a sample suffered LID."

Before proceeding to answer the other questions it is worthwhile to stop and consider carefully the words chosen in the definition of LID.

Only detectable changes are important since if a change has occurred but is too small to detect then for all practical purposes it may as well never have happened. This is relevant to multiple pulse (shot) experiments where it is found that the damage thresholds are almost always lower than single shot values which must imply that there are changes occurring each shot which, on a single shot basis are undetectable but, after enough shots become

observable.

With today's high power lasers it is to be reasonably expected that components irradiated by a single pulse from such lasers will undergo at least transient changes in their properties. Provided that these properties return exactly to their original values before the next laser pulse there will be no cumulative effect which could lead to permanent changes i.e. LID. These transient changes which probably result in a laser producing less power than might have been ideally expected are a "fact of life" for users of such high performance lasers that has to be lived with.

It is possible with a laser to induce changes that actually improve the performance of a sample (e.g. laser annealing) leading to higher damage thresholds although often there is only a fine line between improving the sample by irradiation and actually damaging it.

Thirdly because the laser only interacts with a sample's optical properties it is only changes in these properties that are relevant to LID. However, changes in other properties (e.g. thermal or mechanical) cannot be ignored if they lead to changes in the optical properties of a sample.

2) Answering the question as to what causes LID is by far the most difficult and involved of the four. However, it is generally accepted that in optically opaque samples e.g. metals LID is caused mostly by simple linear rather than complex non linear e.g. 2 photon absorption of the incident light which leads to either mechanical failure such as buckling or cracking or to thermal failure such as melting or vaporization. In normal working conditions the absorption process and subsequent initiation of damage usually occurs on surface defects/contaminants such as microscopic highly absorbing inclusions or residual polishing compound rather than on the intrinsic sample.

The situation in optically transparent samples e.g. dielectrics is much more complicated and there is widespread disagreement in the literature as to the exact details of the damage process. However, there is a consensus that damage in such materials is related to the electric field (E field) induced either on the surface or in the bulk of a sample by the incident light.

On the surface of transparent materials the enhanced electric field associated with the presence of intrinsic defects such as scratches, voids, cracks and assorted extrinsic defects e.g. dust and embedded polishing compound is generally assumed to be responsible for LID.

Most of the disagreements in the literature centre around the details of damage in the bulk of optical transparent samples. There are two competing

and indeed according to some workers interacting mechanisms which, can couple sufficient energy from the incident laser pulse into the lattice of the sample to cause LID. These processes are electron avalanche and multiphoton ionization, both of which will be discussed in detail later.

3) LID can produce a variety of physical results such as the creation of voids and defects in the bulk of a material. In the case of LID on the surface of a sample such things as the removal (ablation) of material from the surface and the creation of totally new surface morphologies (such as ripple patterns) are possible. The consequences of LID are generally catastrophic since a sample that has been damaged will transmit or reflect far less light than before (i.e. show increased scatter and or absorption) and thus the performance of the system is greatly if not completely reduced.

4) Avoidance of LID requires quite simply although often rather impractically that the power a sample is subjected to is kept sufficiently low. If as is often the case in reality this is not possible then some form of early warning of damage should ideally be available. In this context it is with a degree of cautious optimism that the author points the reader to the work in chapter 4.

The rest of this chapter will be divided into several sections dealing firstly with the basic damage mechanisms and then with some of the factors that influence both damage threshold and morphology. Finally there will be a brief discussion of some non damaging laser related effects that are used as diagnostic tools for damage work.

It is not intended for this chapter to be an exhaustive review of LID but to highlight what in the author's view are the most important points and to provide the reader with suitable references to further his or her understanding of LID.

1.2 Sources of Information on LID

A very good basic introduction to LID and related effects is contained within the book by Wood¹. The main source of detailed information on LID and indeed the main forum for exchange of such information is the annual conference on "Laser Induced Damage in Optical Materials" held in Boulder, Colorado, USA. This gathering of workers in damage and damage related areas started in 1970 and continues (to date, 1988) publishing through the National Bureau of Standards (in Washington) a comprehensive set of proceedings for each conference. Essentially the sum of knowledge of damage

is contained in these books to which the reader is referred not only to augment the work herein but to cover the many areas that the author hasn't.

1.3 Historical Review

Back in the sixties when the first high power lasers (Q switched Ruby) were developed it soon became clear that many new effects such as self focusing, second harmonic generation etc were possible with the increased power (by several orders of magnitude) of these pulsed lasers as compared to classical light sources. However, along with the advantages of high power came the soon observed problem that it was possible by "merely" exposing a sample to such a beam of light to cause what is now defined as LID. Marker et al² were the first to report damage in a transparent dielectric and the production of a spark in air by focussing their pulsed ruby laser.

The fact that a beam of light could cause LID may well have come as a shock to workers back in those days however, it is a result of not only new non linear mechanisms that can couple light into a material but also..

1) simple linear absorption, since even absorbing a small fraction of a lot of light can result in a large (by classical light source standards) amount of light and thus energy being coupled into a sample and

2) the presence of defects in the samples used enhancing the coupling of light into the material.

Early work on LID notably that by Bloembergen³ was greatly hampered by such things as irreproducibility of the spatial and temporal profiles of the beam, poor quality samples in terms of their purity and, particularly in dielectrics, the new phenomenon of self focussing whereby the refractive index was found to increase with intensity causing the light to continually converge, eventually producing catastrophic material failure. The exact definition of damage was in those days unclear and to an extent even today this uncertainty persists.

From humble empirical beginnings the subject of laser induced damage in optical materials has developed into a separate area of work involving many workers worldwide. A great deal of money and effort both theoretical and experimental is being applied not only to understand damage but, to design new optical materials which are more damage resistant since ultimately the advance of high power laser systems is limited by the failure of the materials used within it.

1.4 Comment on the Progress on LID

It soon became clear that for the purposes of damage, not only do metals, dielectrics and to an extent thin film coatings form separate classes of material but, that even within a class, different damage mechanisms can and do operate.

Before discussing the mechanisms of damage in the various classes due to their importance in the damage of samples in reality, the role of defects will be discussed briefly to provide some general background. Further details on defects are contained within each class of material.

Following the discussion of defects and before discussing effects common to all classes of materials such as laser annealing, photothermal and photoacoustic effects together with ways of increasing the damage threshold and exactly how the damage threshold is defined, the damage mechanisms operative in each class will be discussed.

1.5 The Role of Defects - a First View

On uniform samples it would be expected that..

1) damage would first occur at the point irradiated by the highest fluence and

2) that if the fluence is sufficient to damage a point, then all points irradiated by that fluence should damage.

In practice unless extreme measures are taken such as working with ultra-pure samples in ultra-clean conditions e.g. high vacuum, what is found is that when a sample damages...

1) the damage occurs in isolated spots which are often well away from the peak of the beam and

2) not all points exposed to a given fluence damage.

It is unanimously agreed by workers in LID (including the author) that this behaviour is due to defects i.e. "weak points" in the material which damage at lower fluences than the surrounding material.

There are a variety of possible defects such as microscopic highly absorbing inclusions e.g. residual polishing compound and surface absorbed contaminants, areas of increased surface roughness and enhanced electric field and, physical imperfections such as scratches, grain boundaries, cracks and voids. In the bulk (for transparent samples) the term defect can also encompass such things as variation in the refractive index due to

either impurities or a variation in stoichiometry.

The essentially random distribution of these defects accounts for damage occurring in isolated spots away from the peak in the beam and the fact that if the fluence used just causes damage on a defect then on a "good" area it will have no effect so accounting for why not all points exposed to the same fluence damage.

The reduction in damage threshold of a surface over the bulk value for a given sample is attributed to the increased likelihood of defects at the surface of a sample particularly such things as scratches, pits and absorbed contaminants being present.

For the purposes of academic study the problem of defects can be greatly reduced if not eliminated by performing experiments using only very small volumes or areas of a sample and tightly focussed laser beams. However, whilst for academic purposes this technique may work it is totally impractical for "real" applications of lasers with the result that in practice defects are almost always the limiting factor in the performance of a sample.

1.6 LID in Metals

Damage in pure metals⁴ for laser pulses longer than the electron relaxation time ($\approx 10^{-14}$ s) is undoubtedly thermal in origin. As to whether the mechanism is melting or some form of thermally induced stress fracture is much harder to answer. It is certainly the case that the fluence required to induce mechanical slip is less than that required to produce melting.

It is very probable that different damage mechanisms operate depending upon whether single or multiple pulse thresholds are being studied and consequently they will be discussed separately.

1.6.1 Multiple Pulse

It has been well documented that the multiple shot damage threshold is lower, often substantially lower than the single shot value. It was initially suggested⁵ that this effect was due to an accumulative heating whereby the temperature of the sample gradually increased shot to shot i.e. the heating effect from a given shot did not have enough time to diffuse away before the next pulse. This would thus require less fluence per shot to raise the sample to melting than attempting to do so in a single shot would. Whilst this is certainly possible (and known as thermal runaway) if the repetition frequency is too high as detailed by Wood et al¹, the work by Lee et al⁶⁻⁷ was performed at a sufficiently low repetition rate that any heating effect from a pulse had decayed before the next pulse was applied.

In their work performed on high quality (but not necessarily defect free) samples, Lee et al observed and explained several important effects namely that..

- 1)the multiple shot threshold was lower than the single shot value
- 2)the single shot threshold decreased as the laser spot size increased from a few microns to about 250 μm whereafter it remained essentially constant
- 3)the multiple shot threshold increased with increasing spot size reaching in the limit of large spot size the same value as the single shot threshold for that spot size.

The single shot spot size effect is attributed to the probability of the laser beam interacting with a defect which, increases with spot size and becomes constant when the spot size is much larger than the average distance between defects. The greater the probability of interacting with a defect the lower the damage threshold will be as the defects damage at lower fluences than the host material.

The effects for multiple pulses are attributed to thermally induced mechanical stress (slip/plastic deformation) as was initially suggested by Musal⁸ and subsequently developed by Lee et al.

The thermal stresses induced in the surface of a metal by heating relax to nearly zero in a direction normal to the surface as there is no constraining mechanism in this direction. However, the stresses can be very large in directions parallel to the surface since displacement in these directions is inhibited by the surrounding unheated region. If these stresses exceed the elastic limit of the metal, the surface will suffer plastic deformation.

Multiple pulse irradiation at sufficiently high fluence will thus create a succession of cycles of compressive (heating) and tensile (cooling) plastic deformations which will eventually result in the formation of large slip bands, which introduce surface roughness that can degrade the optical quality of the surface by increasing nonspecular scatter and or absorptance and ultimately lead to failure of the surface by cracking.

In general there will be two types of stress induced by the laser which are..

- 1)dynamic stress directly induced by laser heating and
- 2)quasistatic stress induced due to the temperature gradient associated with the non uniform spatial distribution of the laser beam.

Dynamic stress will relax by launching stress waves in the material

propagating at the speed of sound ($\approx 5,000 \text{ ms}^{-1}$) whereas quasistatic stress can only relax by the relatively slow process of thermal diffusion.

In his work Musal assumed a uniform spatial distribution of the energy in the laser beam i.e. he considered only dynamic stress and concluded that if the region heated by the laser is much larger than the distance the stress waves propagate during the laser pulse then, dynamic stress would cause damage. He further showed that under these conditions higher fluences can be tolerated at smaller spot sizes since the dynamic stress is able to relax sufficiently quickly.

In their work Lee et al ensured the stress waves were able to propagate much further than the dimensions of the heated region so that quasistatic rather than dynamic stress was the most significant mechanism. By applying linearized thermoelastic theory they were able to show that the critical stress σ for plastic deformation (slip) is given by

$$\sigma = At_p/w \quad 1)$$

where A is a material constant, t_p the laser pulse duration and w the radius of the laser (which is assumed to have a Gaussian distribution spatially) at the $1/e^2$ points. Thus the critical stress is inversely dependent on the spot size which accounts for the multiple shot damage threshold increasing with spot size until the fluence equals that of the single shot threshold where upon the sample will fail by the appropriate single shot mechanism.

Whilst 1) agrees with their experimental observations they commented on the fact that the spot size dependence of the critical stress and therefore the damage threshold depends upon the relative magnitudes of the pulse duration t_p and spot size w specifically, on the relative magnitudes of w^2 and $4kt_p$ where k is the thermal diffusivity. In the short pulse large spot case i.e. $w^2 \gg 4kt_p$ the critical stress is given by 1) but for the long pulse small spot case i.e. $w^2 \ll 4kt_p$ it is given by an equation of the form

$$\sigma = A'w^2 \quad 2)$$

1.6.2 Single Pulse

The mechanism of single pulse damage is not as clearly defined as that for the multiple pulse case as it is possible to produce slip bands with a single laser pulse⁷ and it is unclear as to what if any effect these slip bands produced after only 1 shot would have on the optical properties of the sample.

A more widely accepted criterion⁹⁻¹⁰ for damage involves raising the temperature of a sample's surface to at least the melting point. However, the presence of defects on the surface considerably complicates the situation as there is the likelihood of severe localised heating that can result in damage either by

- 1) straight forward localised melting or
- 2) thermal expansion of the defect area resulting in sufficient stress being induced in either the defect or host material to cause failure by mechanical fracture.

Consider first the experimental work by Porteus et al¹¹ together with the theoretical work by Sparks and Loh⁹⁻¹⁰. In their work Porteus et al took exceptional precautions to try and avoid the problem of defects and used single crystal samples of Al and Cu which were spark cut, orientated, mechanically polished and electropolished before being placed in the test chamber. Final cleaning was performed by argon ion sputtering after bake out and residual argon was removed by annealing at 500° C for one hour prior to testing. Surface cleanliness was verified by Auger analysis but, it should be noted that the surface was of low optical finish due to the sputtering and showed the characteristic "orange peel" effect of rough surfaces.

The most important feature of their threshold results was that the surface when it damaged did so by melting uniformly over the entire irradiated area rather than at isolated spots as in almost all previous experiments thus indicating that the defect problem had been overcome but only at the expense of the optical quality of the surface. This work also showed that the threshold on such samples was more than twice that on "normal" samples under "normal" conditions i.e. ones with defects operated in air.

In their theoretical work Sparks and Loh calculated the energy required to melt the surface of various metals. They also commented on and allowed for in their calculations the large change i.e. a factor of 5 -6 increase in the optical absorption with increasing temperature. Their results showed excellent agreement between the fluence needed to melt Al and Cu and the damage thresholds measured by Porteus et al thus indicating that for "clean" samples melting of the surface is the damage mechanism.

Applying the criterion of surface melting to "normal" samples i.e. samples in air and of good optical quality, resulted in poor agreement between theory and experiment. It was suggested and indeed it is the author's view that this discrepancy is due to the presence of defects on "normal" samples that are absent on the ultraclean samples.

In summary single shot damage in reality is almost certainly defect dominated i.e. extrinsic, which can result in either thermal or mechanical failure whereas on ultraclean samples intrinsic thermal failure is the damage mechanism.

1.7 LID in Dielectrics.

1.7.1 Introduction

As a result of the widespread use of dielectrics both in bulk form and as thin film coatings more research has been performed on this class of material than on any other and thus there is a wealth of literature on the subject. The author would like to draw the reader's attention to several review articles which serve as excellent introductions to this area, these are the papers by Bloembergen³, Smith¹², Sparks et al¹³, Lowdermilk and Milam¹⁴ and finally that by the Soviet authors Manenkov and Prokhorov¹⁵.

Since under their intended operating conditions most dielectrics have very small absorptions (well under 1%), the defect problem (particularly where the defect is in the form of a highly absorbing inclusion such as Platinum specks in glass) is more severe in this class of material than in metals. Unless exceptionally pure (dielectric) samples are used the damage will almost certainly be defect dominated. However, for the moment the presence of highly absorbing inclusions that lead to damage through thermally induced mechanical fracture of either the defect or host material will be ignored but, will be discussed later.

It is universally agreed that the damage mechanism in dielectrics is electric field dependent. Essentially as the incident fluence and hence E field increase, there comes a point where the E field is strong enough not only to accelerate any electrons initially in the conduction band but also via collisional excitation promote valence band electrons to the conduction band. The electrons promoted to the conduction band are provided either by the host material or more likely by impurities with energy levels near to the bottom of the conduction band. Once in the conduction band these electrons can help not only to produce more electrons but, will cause the absorption of the sample to increase as effectively the sample is becoming more and more metallic due to the increasing numbers of "free" electrons. As the absorption increases so will the sample's temperature (Joule heating) and at some point irreversible changes such as melting or fracture i.e. damage will occur.

The above picture is a very simplistic description of what is obviously a very complex interaction and indeed one that is by no means understood. There

are two mechanisms generally considered to be capable of providing enough conduction band electrons to cause damage...

- 1) Avalanche ionization and
- 2) Multiphoton ionization

In the avalanche model it is normally assumed that there are a few electrons initially in the conduction band perhaps as a result of impurity contamination. These electrons are then accelerated by the incident E field until they have enough energy to produce secondary electrons. These secondary electrons will similarly be accelerated and yield yet more electrons causing an "electron avalanche", with the process continuing until the absorption is so high that the sample damages through excessive heating.

The multiphoton model assumes that all the required electrons are provided by non linear optical processes whereby 2 or more low energy photons can combine to ionize an atom (i.e. provide a conduction band electron) where the ionization energy is greater than that of a single photon. This process, as in the avalanche model, continues until thermal failure occurs.

To proceed further it is useful to discuss bulk and surface damage separately.

1.7.2 Bulk Damage

Early in the development of the theory of bulk damage in dielectrics the ideas of Bloembergen were paramount and strongly favoured the avalanche model. He did however point out that as the energy of the incident photon approached that of the bandgap of the sample, multiphoton ionization may become a means of generating large numbers of conduction band electrons. In the mid seventies and early eighties Sparks and his co-workers took up the development of a suitable theory of bulk damage.

Sparks¹⁶ was able to show that for the high quality samples used in damage experiments the probability of a photon interacting with an electron initially in the conduction band (i.e. starting electrons) was ~~negligible~~ ^{negligible} thus showing that all the electrons required for damage had to be generated by the laser pulse. He further showed in detail that neither the avalanche or multiphoton models individually could account for experimental observations of the damage threshold but that by using a theory which combined elements of both models substantial progress could be¹³ made

Essentially the new theory allowed for the generation of the starting electrons by multiphoton ionization of electrons from the valence to conduction bands. The electrons were then accelerated to an energy equal to

that of the bandgap by inter conduction band transitions and or by multiple Holstein processes. The Holstein process involves the interaction of an electron with both a lattice phonon and an incident photon and allows the electron to absorb almost all of the photon energy whilst using the phonon to conserve the wavevector. The main reasons for the good agreement between their work and experiment was the use of more realistic electron phonon scattering rates i.e. energy loss rates by the inclusion of Umklapp processes and the use of temperature dependent lattice constants and phonon frequencies. The E field required to ensure that after allowing for losses an electron would gain energy from the incident field and thus ultimately cause an avalanche effect was then calculated.

Unfortunately, as they commented on, in their calculations they used a mathematical formulation that requires the photon energy to be "small" which, they defined to correspond to wavelengths longer than about $1\mu\text{m}$ with the consequence that their theory fails in the UV. To the author's knowledge there has been no attempt to extend these types of calculations into the UV with the result that no comprehensive UV theory appears to exist. This is not as serious as it might appear since unless exceptionally pure samples are used the damage mechanism is almost certainly attributable to the presence of defects in the sample.

Sparks et al¹³ showed that there was excellent agreement between theory and experimental results on ultra pure samples for the dependence of the breakdown field (and hence fluence) on temperature, pulse duration and wavelength (for $\lambda \geq 1\mu\text{m}$) even though the theory contained no adjustable parameters. This agreement marked a major advance in the understanding of bulk damage in dielectrics in the infra red.

At shorter wavelengths, multiphoton ionization is expected to become increasingly important not only in generating the starting electrons but more importantly in providing generated electrons with sufficient energy to yield more electrons.

Vaidyanathan et al¹⁷ performed a comparison of the critical (damaging) E field predicted by either the avalanche or multiphoton models as a function of laser wavelength from $0.355\mu\text{m}$ to $1.06\mu\text{m}$. They also studied the predictions of a combined avalanche multiphoton model by Vinogradov¹⁸. They concluded that individually neither the avalanche or multiphoton models could account for experimental observation but that the combined model showed better agreement.

It should be noted that at $1.06\mu\text{m}$ far better agreement between theory and experiment was obtained by Sparks et al as opposed to Vaidyanathan et al and leads the author to speculate that if Sparks et al had extended their

calculations into the UV (as they stated in their paper that they were going to do) then, a detailed understanding of damage in dielectrics from the UV to the IR would probably have been obtained.

1.7.3 Surface Damage

Since at an interface it is almost impossible to avoid defects either in the form of physical imperfections or impurities, surface damage more than bulk damage is defect limited.

Consider for the moment samples that are impurity free. It is well known as initially pointed out in Bloembergen's¹⁹ classic paper that, on such samples the surface damage threshold is lower (often by an order of magnitude or more) than that of the bulk. In this paper he attributed the reduction in threshold to the presence of submicroscopic cracks and pores in the surface that cause local enhancement of the electric field which leads to damage at lower incident fluences.

It was shown that the field inside a number of possible types of defect could be written in the form

$$E_{ins} = 1/\{1+[(1-\epsilon)/\epsilon]L\}E_0 \quad 3)$$

where E_0 is the incident field, E_{ins} the field in the defect, ϵ the dielectric constant and L the depolarization factor appropriate to the geometry of the defect in question.

The parameter L has the value $1/3$ for a spherical pore, $1/2$ for a cylindrical groove and ≈ 1 for a crack that is much longer than it is thick. Hence for long thin cracks the field in the defect is enhanced by a factor of ϵ over the incident field which for high index materials can result in enhancements of up to a factor of 10 (e.g. for CdTe or ZnSe at $10.6 \mu\text{m}$) in the E field and hence 100 in intensity. This results in lowering the apparent surface damage threshold by up to a factor of 100 compared to that of the bulk. It was further shown that the predicted lowering of the damage threshold due to E field enhancement was in good agreement with that found on samples of ultraclean and pure glass and Al_2O_3 .

Even nominally transparent solids will have a small absorption and, by taking reasonable values of this small absorption together with a crude estimate of the temperature rise associated with this absorption for a defect of size A , Bloembergen was able to show that the presence of defects of less than $\approx 0.01 \mu\text{m}$ in diameter is unlikely to lead to damage.

1.7.4 Absorbing Inclusion damage

Now that the intrinsic damage mechanisms i.e. non impurity mechanisms have been discussed the role of highly absorbing inclusions (impurities) which, lead to what can be termed extrinsic damage can now usefully be discussed. It should be noted that for bulk damage the use of ultra-pure samples essentially removes the problem of impurities but, for surface damage it is much harder to eliminate such impurities as the surface is always going to be contaminated by the surrounding environment.

Providing an impurity does not absorb at the laser wavelength it's presence will not affect the damage threshold but if it absorbs, then, as shown in the classic paper by Hopper and Uhlmann²⁰ it can lead to damage via mechanical failure due to thermally induced stresses around the defect. This failure mechanism can occur both in the bulk and on the surface of a sample and, in practice is almost certainly the dominant damage mechanism. In general the most likely types of absorbing inclusion are metallic (notably Pt) and oxide particles.

In their paper Hopper and Uhlmann considered the stresses induced in crown glass as a result of the presence of specks of Pt of varying sizes. By solving the heat conduction equation for various boundary conditions they were able to show that the heating of the particle goes through a maximum as the particle size increases. By subsequent application of elastic theory they concluded that for particles between 100 nm and a few microns the induced tensile stresses exceed the theoretical strength of the glass and thus mechanical failure will occur. They also concluded that the presence of oxide impurities such as FeO could lead to damage through excessive thermally induced stress.

It should be noted as pointed out by Walker et al²¹⁻²² that Hopper and Uhlmann used certain approximations (in particular the assumption that the thermal diffusivity is large) in solving the heat conduction equation that are valid for good conducting impurities such as metallic particles but not for poorly conducting impurities such as dielectric materials. In the case of dielectric impurities an exact solution to the heat conduction equation such as that by Goldenberg and Tranter²³ is required.

1.7.5 Multiple Pulse and Spot Size Effects

In the cases where the damage is dependent only upon the electric field then provided thermal runaway does not occur the damage threshold will be independent of both the number and frequency (rep rate) of laser pulses used. The presence of absorbing defects can yield a failure mechanism similar to that responsible for the multiple pulse failure of metals i.e. progressive

slip and plastic deformation at low rep rates and increase the likelihood of thermal runaway at higher rep rates. As commented on by Lange et al²⁴ (to which the reader is referred for a much more detailed discussion) there is a great deal of confusion in the understanding of multiple pulse effects.

As with metals, as the spot size is increased from a few microns the damage threshold apparently decreases reaching a constant value for sizes typically greater than 300 μm . This is explained as before by the increased probability of interacting with a defect as the laser spot size is increased. The author has not found any work which shows a non defect related spot size effect.

1.7.6 Front and Back Surface Effects

Early on in the study of damage it became evident that in a transparent solid where the incident and exit surfaces had been identically prepared, the exit (back) surface damaged at a lower fluence than the entrance (front) surface. This effect was explained by Crisp et al²⁵ on the basis of interference between the incident and reflected waves at each surface.

At the front surface the reflection is at a more dense media so the reflected and incident waves are out of phase whereas at the back surface the reflection is at a less dense media so the reflected and incident waves are in phase. This leads to the intensity at the rear surface being given by

$$I_r = 4n^2/(n+1)^2 I_i \quad 4)$$

where n is the refractive index and I_i the intensity at the front surface. Hence for a refractive index of 1.5 the rear surface intensity is 1.44 that at the front surface and so the rear surface threshold should be 69% of the front surface, a value in good agreement with that found experimentally.

1.8 LID in Thin Film Coatings

Given that thin film coatings consist of many layers and hence interfaces between often very different materials deposited under conditions that are often far from ideal (clean) it is not surprising that it is this class of materials that suffers the greatest defect problem and thereby has the most inconsistencies in the literature. Excellent introductions to the area can be found in the work and review by Lowdermilk and Milam¹⁴ together with the papers by Walker et al²¹⁻²².

In their work Walker et al performed a detailed comparison between experimentally measured damage thresholds at various laser wavelengths and pulse widths and the three main theoretical mechanisms

(avalanche/multiphoton ionization and absorbing inclusions) in an attempt to determine the mechanism operative in thin films. Of the 15 effects they observed, the absorbing inclusion (impurity) model was able to explain accurately 14 whereas the other two mechanisms could not explain any of the observations.

Table 1.1 shows a comparison of experimental data for the variation in damage threshold with wavelength, pulse length and film thickness and the 3 main damage mechanisms together with a combined avalanche/multiphoton ionization model where it can be seen clearly that the dominant damage mechanism is absorbing inclusions i.e. impurities.

Their impurity model though based on either the Hopper and Uhlmann or Goldberg and Trantor solutions to the heat conduction equation included for the first time a variable absorption cross section. As they commented, the use of a constant for the absorption cross section is only valid when the impurity is much larger than the wavelength of the incident radiation whereas in thin films the size of impurity is usually limited by the film thickness and is therefore of the same order as the wavelength of the incident light. To allow for this size effect the constant cross section was replaced with the cross section obtained by using Mie scattering theory which has a dependence on both size of the impurity and the wavelength of the incident light. The breakdown criterion was taken to be melting of the host material.

Two of the strongest pieces of evidence to support their contention that impurities are the dominant damage mechanism are the observations that the damage threshold decreases as the incident laser wavelength decreases (see later) and the damage threshold increases as the film thickness is reduced. The latter is explained on the basis of the exclusion of larger, easier to damage impurities as the film thickness is reduced.

Since almost all coatings are based on layers that are optically a quarter wave thick, the nodes and antinodes of the E field occur at the interfaces between the high and low refractive index materials. Given that interfaces suffer the most from defects and physical imperfections, the occurrence of nodes and antinodes at such interfaces is intrinsically undesirable.

Work by Apfel et al²⁶ showed good correlation between damage threshold and the E field in multi layer silica/titania stacks. There has been a limited amount of work to attempt to improve the damage threshold by the use of non quarter wave (optically) layers but the results have in the main been inconclusive which the author suspects is related to not overcoming the impurity problem.

Table 1.1

Evaluation of the Avalanche, Multiphoton and Impurity Models
when Compared to the Experimental Data (after ref 22).

Model	Parameter	Wavelength	Pulse Length	Film Thickness
Multiphoton		reasonable	poor	poor
Avalanche		poor	poor	poor
Combined		inconsistent	poor	poor
Impurity		good	excellent	excellent

1.9 Dependence of LID on Laser and Material Parameters

There is a lot of confusion and much contradictory evidence in the literature as to the dependence (i.e. so called scaling factors) of the damage threshold on such things as laser wavelength, pulse duration, spot size and on material parameters such as surface roughness and methods of sample preparation. Since each of the 3 main damage mechanisms generally predict different scaling factors, it is the author's view that, at least some of the confusion in the literature arises from incorrectly attributing damage to a particular mechanism however, in some cases it is possible to extract general trends as will now be done.

Generally as the laser wavelength decreases so does the damage threshold however, it is by no means clear as to the functional dependence of the threshold on wavelength. For alkali halide crystals²⁷ the threshold shows little if any variation from 10.6 μm to about 1 μm whereafter it decreases rapidly. For metals it is likely that the increase in linear absorption (i.e. the decrease in reflectivity) as the wavelength decreases is responsible for the reduction in damage threshold.

It is easier to obtain an expression for the dependence of the threshold on the duration of the laser pulse than for laser wavelength. Both the impurity model²² for thin films/dielectrics and the thermal heating model for metals⁹ (even if allowance is made for the variation in optical absorption with

temperature) predict that the threshold varies with pulse width t as $t^{1/2}$. In each case experimental observations confirm this dependence. It should however be pointed out that occasionally other dependences predicted by either of the other 2 main damage mechanisms or indeed by different formulations of the impurity model appear to be more appropriate to the results obtained by some workers. It should also be noted that for very short (<150 ps) or very long (>300 ns) pulses the situation can become much more complex and these simple models may well fail.

The effect of spot size, because of its complications and ease of misinterpretation was discussed previously for each class of materials individually.

In the literature (and from the author's own experience) there is a general "belief" that the damage threshold decreases as the surface roughness increases but almost no work to quantify this effect except that by House et al²⁸ (experimental) and by Babu²⁹ (theoretical) appears to have been performed.

In their work House et al showed that the damage threshold decreased with increasing surface roughness and that a relation of the form...

$$E\sigma^m = \text{constant} \quad 5)$$

can be used to describe the relationship between damage threshold and surface roughness where E is the breakdown field, σ the rms surface roughness and m is a parameter of value ≈ 0.5 . The parameter m was found to vary somewhat with the method of surface preparation (e.g. polishing, etching etc).

On the premise that the surface damage threshold depends upon the total exposed area i.e. the area including all surface "hills and valleys" and with some further simplifying approximations Babu was able to theoretically predict the above relation with $\sigma = 0.5$.

In the author's view much work both experimental and particularly theoretical needs to be performed to properly quantify and understand the effect of surface roughness on the damage threshold.

1.10 Laser Annealing (Polishing)

Although laser annealing usually refers to the processing of semiconductors on which vast amounts of work have been performed (see chapter 3), in the context of LID, laser annealing (polishing) refers to irradiating samples at non damaging fluences with the intention of raising the damage threshold.

In most cases the test surface is irradiated by either multiple laser

pulses³⁰ below the (non irradiated) single shot damage fluence or a cw beam³¹ for typically a few seconds. As an illustration of this effect consider figs 1.1a-b which are of the damage on colloidal silica induced by an excimer laser. Fig 1.1a is that produced by irradiating the sample with 1000 (non damaging) pulses (at 1hz.) at 0.15 Jcm^{-2} followed by a single (damaging) pulse at 6 Jcm^{-2} whereas fig 1.1b is the damage produced by a single 6 Jcm^{-2} pulse with no low fluence pre irradiation. The processed sample shows far less damage than the unprocessed one thus showing the improvement in the sample as a result of low level irradiation.

Swain et al³⁰ report upto a factor of 2 increase in the single shot threshold of colloidal silica antireflection (AR) coatings if they are processed by typically 10 pulses at non damaging fluences. Temple et al³¹ reported more than a factor of 2 increase in threshold on some samples of fused silica processed by a cw CO₂ laser and then damaged by a pulsed CO₂ laser.

Two mechanisms are thought to be responsible for this effect, either the cracks, scratches etc (i.e. physical imperfections) left in the surface after mechanical polishing are reduced if not removed by the heating effect of the laser or, the heating effect removes surface contaminants (laser desorption). The work by Swain et al tends to support the laser desorption mechanism whilst that by Temple et al supports the reduction of physical imperfections.

It is the author's view that almost certainly both mechanisms occur simultaneously and that the condition and nature of the surface dictates which mechanisms dominates. If the surface easily absorbs water vapour etc then the desorption process probably dominates whereas on say a clean metallic surface the reduction of physical imperfections mechanism probably dominates.

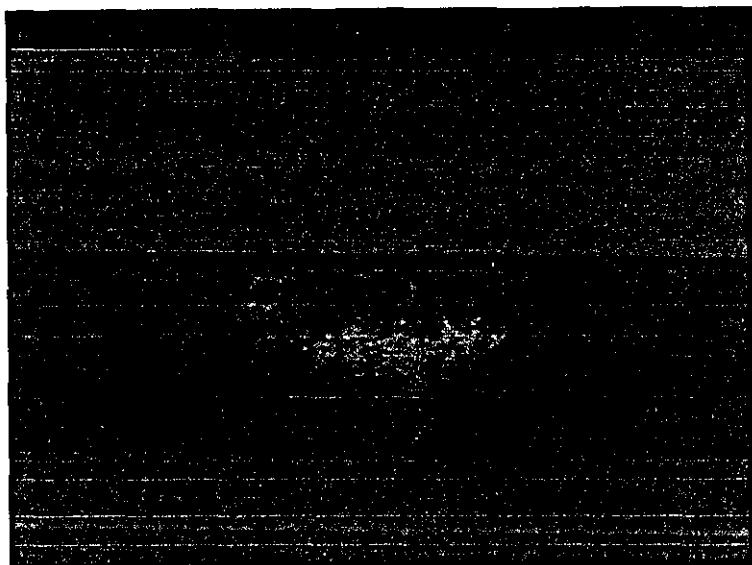
1.11 UV LID

Whilst nearly all of the above applies to UV LID, the reduction in incident wavelength brings many extra problems and imposes far more severe conditions on sample preparation.

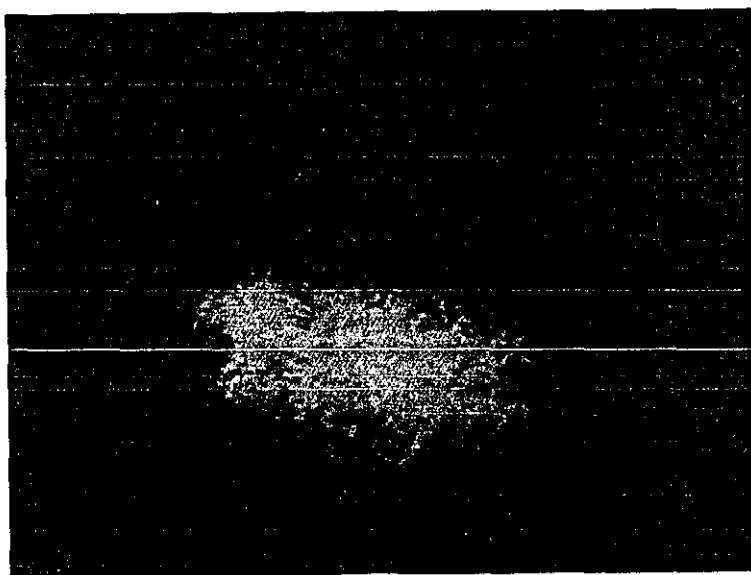
In practice and almost certainly due to increases in one or more of

- a) absorption
- b) the importance of smaller and more numerous physical imperfections
- c) the probability of multiphoton ionization

the damage threshold of materials decreases as the incident wavelength



A



B

shortens.

Whilst a reasonable amount of work has been performed on UV LID the author would like to draw the reader's attention to volume 476 of the conference proceedings³² published by the Society of Photo-optical Instrumentation Engineers (SPIE) which deals exclusively with excimer lasers and UV LID.

In the far UV (i.e. below about 300 nm) due to the relative magnitudes of the photon energy and the bandgap in most materials, the main problem is finding transmitting materials with the result that in practice UV transmissive optics are almost always based on metallic fluorides (e.g. CaF₂ and MgF₂) and oxides (e.g. SiO₂ and Al₂O₃).

Reflecting optics can be based on multilayer dielectric stacks or on metals but, it should be noted that in the far UV the reflectivity of metals can decrease significantly. Al is probably the most widely used metal but, due to the rapid formation of its oxide cannot be used below about 220 nm unless covered with a protective overcoat such as MgF₂.

The reduction in laser wavelength means not only that absorption (either intrinsically in the material or from absorbing inclusions) increases but that physical imperfections which are allowable at say 1.06 μm cannot be tolerated in the UV and so far more care in and control of sample preparation is required.

In the author's view these effects together with what often appears to be a very "casual" attitude by manufacturers to the conditions in which samples are prepared have resulted in the relatively poor damage thresholds to date of UV optics. However, it appears now that manufacturers are making a serious attempt to rectify these problems.

1.12 Definition of

1.12a Fluence

Up until a few years ago it was common practice to obtain the fluence in a beam by dividing the beam energy by the area at the $1/e^2$ points. Whilst this is reasonable for a near uniform beam it is inappropriate for a beam with significant variation in spatial intensity such as a true Gaussian beam where the peak fluence is 2.31 that calculated by the above method.

With improved beam profiling techniques not only the peak fluence but the actual fluence at any point in the beam can easily be found and it is now becoming much more common in quoting damage thresholds to use the actual fluence (rather than using the beam energy divided by the $1/e^2$ area) that did not cause damage.

1.12b Damage Threshold

Confusion as to the definition of the damage threshold arises because the impurity/defect mechanism is in practice the dominant damage mechanism. If the laser spot size used is not much larger than the average distance between defects then the fluence at which an area damages will vary depending upon whether a defect was or was not in the irradiation area. Even if the spot size is much larger than the inter-defect distance problems can arise if the intensity of the laser beam varies significantly across the irradiated area.

Hence given the above it is not surprising that if several sites are irradiated at the same fluence not all of them damage and thus to fully quantify what is happening both the fluence used and the probability of damage must be specified.

The confusion in the literature occurs as to what probability of damage should be associated with the definition of the damage threshold. It is probably true that most reports on LID use the 50% probability level as the definition of damage i.e. the fluence at which 50% of the sites tested at this fluence damaged. The 0% level i.e. the highest fluence at which no site tested at this fluence damaged is sometimes used with only a very few workers using the 100% level i.e. the lowest fluence at which all sites tested at this fluence damaged. To an extent, exactly which level is most suitable depends upon both the desired application of the component and the laser spot size used.

In the case of a spot size comparable to or smaller than the average inter-defect distance the maximum non damaging fluence corresponds to the intrinsic (defect free) damage threshold of the host material. For spot sizes much larger than the average inter-defect distance the probabilistic nature of the damage is associated with the location of a defect relative to the most intense part of the beam and with the uniformity of the beam over the irradiated area. In this case a true measure of the intrinsic threshold can only be obtained by finding the maximum fluence that does not cause damage on sites irradiated by the most intense part of the beam. In both cases, the 0% level is indicative of the defects present on the sample.

Various authors e.g. Porteus and Sital³³ or Foltyn³⁴ have proposed statistically based theories of defect dominated damage to explain the observed spot size effect and to extract the absolute (0%) onset level of damage as well as the intrinsic threshold.

The use of small spot sizes whilst desirable for academic work is impractical in reality where, components generally have to withstand large area irradiation. Thus the author prefers to use large area irradiation (to

simulate reality and eliminate spot size effects) and to define the damage threshold as the 0% level as this gives the maximum fluence that can be used in practice with no likelihood of damage to the sample.

1.13 Photoacoustic and Photothermal Effects

Although not directly related to LID, but because of their increasing use as diagnostic tools for LID, the photoacoustic (PA) and photothermal (PT) effects will be briefly discussed. The reader's attention is drawn to the detailed review of both of these effects by McDonald³⁵ and to any one of the many papers by Rosencwaig³⁶ who pioneered most of the work in this area.

The PA effect is the generation of acoustic (sound) waves in a solid or in the ambient gas by the absorption of heat usually from a periodic heat source such as a chopped (repetitively pulsed) laser beam. The waves occur in any sample with a nonzero coefficient of thermal expansion due to the stress-strain conditions induced in the sample.

The PT effect is the deflection of a probe laser beam due to interaction with a heated region (solid and or gaseous) produced by a second more powerful pump laser. The probe beam can be deflected by direct interaction with a heated surface although in practice the deflection due to the heated gas e.g. air surrounding the sample is usually orders of magnitude bigger.

1.14 Conclusions and Comments

LID, whilst still largely empirical is gradually becoming better understood theoretically with the reproducibility of samples and experiments slowly improving. Whilst much work both experimental and in particular theoretical remains to be performed, the outlook for LID appears bright.

References

- 1)"Laser damage in Optical Materials" by R.M.Wood, pub. by Adam Hilger 1986
- 2)P.D.Marker et al, Proc. 3rd.Int.Conf. on Quantum Electronics, ed by P.Grivet and N.bloembergen, Paris, Dunnel 1964 p 1559
- 3)N.Bloembergen, IEEE J.Quant.Elec. 10 375 (1974)
- 4)D.L.Decker et al, NBS Spec.Pub. 669 p178 (1984)
- 5)A.B.Callendar, NBS Spec.Pub. 435 p202 (1977)
- 6)C.S.Lee et al, Appl.Phys.Lett. 41 625 (1982)
- 7)C.S.Lee et al, J.Appl.Phys. 54 5727 (1983)
- 8)H.M.Musal Jr.,NBS Spec.Pub. 568 p159 (1980)
- 9)M.Sparks and E.Loh Jr., J.Opt.Soc.Am 69 847 (1979)
- 10)M.Sparks and E.Loh Jr., J.Opt.Soc.Am 69 859 (1979)
- 11)J.O.Porteus et al, Appl.Phys.Lett 29 156 (1976)
- 12)W.L.Smith, Opt.Eng. 17 489 (1978)
- 13)M.Sparks et al, Phys.Rev.B 24 3519 (1981)
- 14)W.H.Lowdermilk and D.Milam, IEEE J.Quant.Elec. 16 1888 (1981)
- 15)A.A.Manenkov and A.m.Prokhorov, Sov.Phys.Usp. 29 104 (1986)
- 16)M.Sparks, NBS Spec.Pub. 435 p331 (1977)
- 17)A.Vaidyanathan et al, IEEE J.Quant.Elec. 16 89 (1981)
- 18)A.V.Vinogradov and F.S.Faizullof, Sov.J.quant.Electron 7 650 (1977)
- 19)N.Bloembergen, Appl.Opt. 12 661 (1973)
- 20)R.W.Hopper and D.R.Uhlmann, J.Appl.Phys. 41 4023 (1970)
- 21)T.W.walker et al, IEEE J.Quant.Elec 17 2041 (1981)
- 22)T.W.Walker et al, IEEE J.Quant.Elec 17 2053 (1981)
- 23)H.Goldenberg and J.C.Tranter, Brit.J.Appl.Phys. 3 296 (1952)
- 24)M.R.Lange et al, NBS Spec.Pub 727 p382 (1986)
- 25)M.D.Grisp et al, Appl.Phys.Lett 21 364 (1972)
- 26)J.H.apfel et al, NBS Spec.Pub. 462 p301 (1976)
- 27)A.Schmid et al, Phys.Rev.B 16 4569 (1977)
- 28)R.A.House et al, IEEE J.Quant.Elec 13 361 (1977)
- 29)S.Babu, IEEE J.Quant.elec 17 533 (1979)
- 30)J.E.swain et al, NBS Spec.Pub. 669 p292 (1984)
- 31)P.A.Temple et al, Appl.Opt. 21 3249 (1982)
- 32)Volume 476 in the series of conference proceedings of SPIE entitled "Excimer lasers, their applications, and new frontiers in lasers" published in 1984.
- 33)J.O.Porteus and S.C.Seitel, Appl.Opt. 23 3796 (1984)
- 34)S.R.Foltyn, NBS Spec.Pub. 669 p368 (1984)
- 35)F.A.McDonald, Can.J.Phys. 64 1023 (1986) and references therein

36)A.Roscnwaig and J.Opsal, IEEE Trans. Ultrasonics, Ferroelectrics and Frequency Control 33 516 (1986) and references therein

Chapter 2

Laser Beam Profiling and Damage Detection

2.1 Introduction

It has often been said that the Laser was a solution in search of the problem, however in connection with LID it is usually the laser itself that is the problem, not only in terms of trying to detect the effects of the laser beam but in particular, trying to characterize the beam itself.

Laser beam profiling is the characterization of the spatial and temporal distribution of the energy within a beam. This type of analysis is essential if meaningful values of beam fluence are to be used in LID experiments. Once performed it is possible to quote such figures as peak fluence and average fluence (J cm^{-2}) together with the "spot size" of the beam taken herein to be the area of the beam at the points where the intensity is reduced to $1/e^2$ that of the peak value.

Damage detection is the process of determining firstly if a sample has changed its optical properties in anyway and secondly, if this change conforms to a criteria that defines LID in which case the sample is said to have "damaged".

2.2 Review

2.2a Laser Beam Profiling

In general the temporal profiling of a cw or long (i.e. $>1\text{ns}$) laser beam is straight forward and accomplished by using suitably fast photodiodes and oscilloscopes although for short pulses i.e. $<1\text{ns}$ other techniques such as the use of streak cameras have been developed whilst for ultra short pulses ($<1\text{ps}$) techniques based on the auto-correlation of the beam with a temporally delayed version of itself are essentially the only way to obtain temporal information about the beam.

The spatial profiling of a laser beam is considerably more complicated than temporal profiling and, at least in the past, the techniques used have depended strongly on the type of laser i.e. wavelength of operation involved.

Some of the first techniques¹⁻³ for spatial profiling used photographic recording of the beam in some form. The basis of all these techniques was to image the beam with suitable attenuation onto pieces of film. The major problem with these methods is the variation of optical density (i.e. $\log_{10} T$ where T is the fraction of light transmitted) of the film with exposure

(where exposure is the product of intensity of light and the time for which the light was incident on the film) which, as was shown by the classic paper of Hurter and Driffild⁴ be highly non linear. This non linearity which can vary enormously from film to film requires that some form of calibration, which is very time consuming, be performed before each profile can be recorded.

All of these techniques in some way attempted to reduce the calibration problem by using such things as a multiple lens element¹ with a different neutral density filter attached to each lens to obtain a series of images of the beam or, using multiple beamsplitting techniques²⁻³ to provide a set of images of the beam. The aim in providing multiple images was to obtain a set of images with a known intensity ratio between them, covering a total intensity range of typically two decades, so that by using a microdensitometer a linear profile of the beam could be obtained from the non linear photographic recording. Whilst these techniques do provide a spatial profile of the beam they suffer from having to use chemicals to develop the film (wet processing) and consequently they cannot provide real time information about the beam.

To avoid wet processing and in an attempt to obtain at least near real time beam profiling beam scanning techniques⁵ were developed based on translating some form of detector spatially across the beam. A single detector such as a photodiode and a suitable aperture are translated across the beam and its output recorded as a function of position. When a pulsed laser is being profiled several values at each position would be taken to reduce the problem of shot to shot variation of the laser energy. The output of the detector and its physical meaning depend greatly on the type of aperture used, typical apertures being pinholes slits and knife edges. A pinhole provides the best spatial resolution but suffers from the small amount of light that actually reaches the detector through the pinhole. Scanning techniques are most suited to beams that are axially symmetric or nearly so since, because the entire area is not profiled hot spots or other irregularities can easily be missed and cause significant errors in the beam parameters.

In theory and to a large extent in practice systems based on imaging detectors (arrays) can solve all of the problems that photographic and beam scanning techniques have, and yield very simply, quickly and accurately a spatial profile of the beam. The basis of these systems is that an image of the beam ideally, but not always optical is formed, which can be subsequently processed (ideally by computer) to yield the required beam parameters. Devices used to perform the imaging can be divided into 2

classes...

- 1) discrete element arrays such as pyroelectric arrays⁶ and CCD/CID devices⁷ (usually in the form of video cameras) and
- 2) integrated target detectors such as vidicon tubes⁸.

The use of pyroelectric arrays will yield the spatial profile of the beam but since a direct optical image of the beam is not formed, large amounts of computer processing are required to generate the profile in an easily recognisable form.

Vidicon systems will form a direct optical image of the beam. These systems suffer from a variety of noise problems that severely limit their dynamic range (i.e. the intensity range over which the device responds in a linear fashion) to a factor around 100. However they have the advantage of larger total image area when compared to discrete element arrays. Systems based on charge coupled devices (CCD) and charge injection devices (CID) will also form direct optical images of the beam and suffer very little noise problems enabling dynamic ranges of 10^3 or more to be readily obtainable. Standard commercial devices have pixel (i.e. each discrete element) sizes of around $20 \mu\text{m}$ square and come in the form of 2 dimensional arrays of typically 400 by 600 elements. For specialist applications e.g. military use it is possible to get (at substantially higher cost) arrays of 10000 by 10000 elements with pixel sizes smaller than $20 \mu\text{m}$.

Lasers operating outside the direct response range of vidicon and CCD/CID devices which is typically 400-1100nm can be profiled by using such devices as fluorescing glasses (for UV lasers) and graphite blocks (for CO_2 lasers) to produce "visible" images that can be readily detected by the sensing device.

Once the data for the profile of a beam has been obtained there is still the problem of exactly which parameters are defined and quoted for the beam. The most important is some measure of the beam fluence but unfortunately there is widespread disagreement in the literature about exactly which measure to quote. Two forms are widely used, either the peak fluence or an "average" fluence. The peak fluence corresponds to the value at the most intense part of the beam whilst the average fluence is usually taken to be the energy of the beam divided by the spot size area, and hence the average fluence relates either to the $1/e^2$ area or half power area. It is trivial to show that for a perfect Gaussian beam the peak fluence is 2.31 times the average fluence when the averaging is performed over the $1/e^2$ area. Note that in what follows unless specified otherwise all quoted

fluences are peak fluences.

2.2b Damage Detection

It has already been discussed in chapter 1 that laser induced damage (LID) can be defined as a permanent change in the optical properties of a sample. This change is usually associated with degradation of these properties but phenomena such as laser polishing and annealing (see chapter 1) show that laser induced changes can improve the optical properties of a sample. Given that in most cases, damage is caused by defects e.g. dust particles, residual polishing material, inclusions, voids, cracks and scratches in a sample and therefore is statistical in nature it is not sufficient to merely say that at fluence x the sample damaged, there must be a probability figure associated with the quoted threshold fluence. For example 0% would indicate that the sample never damaged at this fluence, whereas 100 % would mean the sample always damaged at this fluence. A value of 50% means that 1 out of every 2 sites damaged. In the literature there is a general but by no means conclusive trend towards defining the threshold for LID as the fluence at which there is a 50% chance of damage.

Little appears to have been written specifically on methods for detecting LID except for the work by Milam⁸ and there is little consistency in the methods used by different workers. Essentially the techniques used can be divided into 2 classes, imaging and non imaging techniques.

The simplest of the non imaging techniques is the scatter probe⁹ in which a photodiode and lens combination, where the lens has a stop in its focal plane to block specular light, is used to monitor changes in the scattered light coming from the test sample. The output of the photodiode is recorded on an oscilloscope for subsequent analysis. The photoacoustic effect¹⁰ in which a suitable transducer has been attached to the sample has been used to detect LID on a real time basis. Whilst non imaging techniques involve little equipment and are simple to perform they give no information as to the spatial location and extent of the damage and it is impossible to tell from them if damage has been caused by blowing particles such as "dust" and "grease" off the surface or if the surface itself has damaged which is extremely important to know in some experiments particularly those involving LID on thin film coatings.

Traditionally, imaging techniques such as direct unaided observation of the surface, optical microscopy in all its forms (e.g. Nomarski, total internal reflection¹¹, dark and brightfield illumination) and Schlieren/photography¹² based systems have required the user to make a real time judgement as to whether any change of a sample has occurred, as it has not been possible to

store the images of the sample before and after the laser pulse in order that comparisons can be made. On samples that have significant numbers of scattering sites, scratches etc before exposure to the laser (as most real samples do) it can be almost impossible to make reliable real time judgements on any changes that may have occurred. The recent availability of video recorders and in particular the very recent development of computer controlled video framestores has meant that it is now possible to save before and after images of a sample for either manual comparison or computer processing e.g image subtraction and enhancement. Passing reference to the use of video recorders in detection is made in the literature but to the author's knowledge the use of video framestores and computer processing as developed in 2.6.3 has not been reported.

2.3 Damage Facility

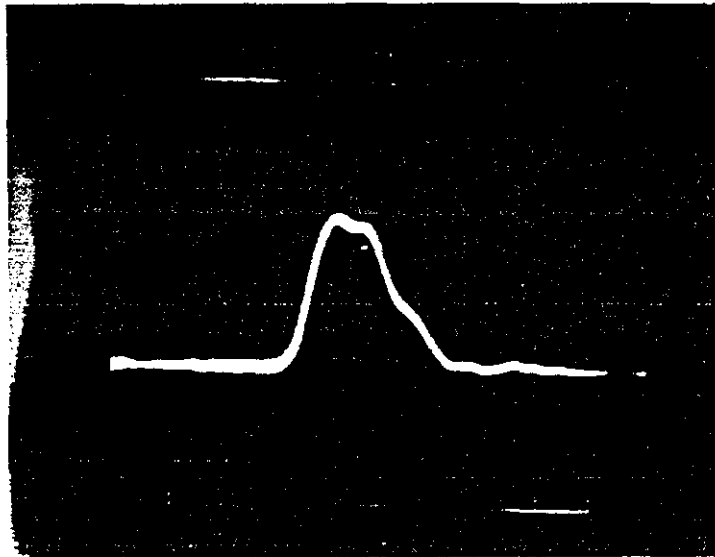
The work performed in this thesis is based upon a Krypton Fluoride (KrF) excimer laser operating at 249 nm i.e. well into the UV. The laser, a Lambda Physik model EMG 200, when operated with KrF gas is capable of producing a nominal 1 J pulse of duration ≈ 20 ns (to the half power points) as shown in fig 2.1. Pulse energies were measured with a Laser Instrumentation Calorimeter and were found to vary shot to shot by approximately 5%.

Two forms of attenuation were routinely used to reduce the laser energy to the desired values, these were the use of special filter glasses of varying optical density and a purpose built variable length dye cell attenuator¹³. The filter glasses depending upon their physical thickness gave fixed attenuations of 0.5 to ≈ 0.01 . The dye cell consisted of 2 tubes with appropriate windows and seals so that one tube could slide within the other thereby producing a variable length between the input and output windows. The volume between the windows was filled with Rhodamine 6G dye (R6G) in distilled water. As R6G absorbs strongly at 249 nm, by varying the concentration of dye and or the length of the dye cell any desired attenuation could be produced.

Fused silica was used as the material for all of the excimer optics as it has essentially no absorption at 249 nm and is resilient to any heating effects caused by high repetition operation of the excimer laser.

To reach fluences sufficient to cause damage, the excimer beam was focussed by a fused silica lens. All of the work performed in this thesis was done using a nominal 20 cm focal length lens.

Due to space considerations and the requirements of colleagues, the excimer beam had to be steered to the focussing lens and hence the plane of operation by various combinations of mirrors and or beamsplitters. Since all



time

Fig 2.1
Temporal profile of excimer beam
timebase 10 ns/div

energy measurements were made with the calorimeter immediately after the dye cell attenuator, an allowance for losses in the beam due to the steering mirrors and beam splitters was required.

This was done by measuring the average laser energy over 10 shots with the calorimeter in its normal position and then repeating the process with the calorimeter close to the focal plane of the focussing lens. The ratio of these two values represents a correction factor which allows determination of the energy in the operating plane from a measurement further back in the system at a far more convenient place.

2.4 Beam Profiling - Initial Technique

When the damage test facility was initially setup beam profiling was performed by means of a scanning pinhole. In this technique a small ($\approx 1\mu\text{m}$) pinhole was mounted on the front of a UV photodiode and then the combination was placed in the beam such that the plane of the pinhole was co-incident with the plane in which the profile was required. By means of an "x y" translator equipped with micrometer drives the pinhole/photodiode was scanned in $\approx 20\ \mu\text{m}$ steps along the horizontal and vertical central axes of the beam i.e. the beam was scanned along only one horizontal line and one vertical line. To reduce the problem of shot to shot variation in the excimer energy the average photodiode output from 5-10 excimer pulses were used. Not only was this very time consuming taking typically 2 hrs per profile but since the beam is profiled only along its axis, non axial hot spots would not be detected causing serious errors in beam parameters.

2.5 Beam Profiling using a Video Framestore

2.5.1 Introduction

To improve the quality of beam profile and reduce the time required to obtain it, the use of a video imaging system in which the beam is directly imaged was investigated. Since the response range of video cameras is around 400 to 1100 nm a medium was required to convert the UV light to detectable light. The obvious choice for UV light and the one used is fluorescence.

The intention was to use suitable fluorescers to produce a visible image of the UV beam which the video camera could detect and to have the camera linked to a computer controlled video framestore in which the video signal corresponding to the beam image is digitised and stored electronically pending image processing.

2.5.2 Video Basics

In the UK the tv screen and hence the picture consists of 625 lines of information, with each line running essentially left to right across the screen and lasting 64 μ s. Each group of 625 lines is known as a frame and subdivided into 2 groups of 312.5 lines known as fields. The first field consists of lines of the odd lines e.g. 1,3,5 etc and the second the even lines e.g. 2,4,6 etc. The image as viewed, is built up by a process known as interlacing in which all the lines of one field e.g. the even field are displayed before all those in the other field. This conveniently results in a complete scan of the tv screen being performed every 20 ms with a frame taking 40 ms to build up.

To "tell" the tv where the end of each line or field is, a video signal consists of two parts, the actual data in the form of voltage levels between 0.3 and 1 volt together with sync pulses of 0 to 0.3 volts. To distinguish between line and field sync pulses the time duration of the pulse is varied being about 5 times longer for field sync pulses than line sync pulses. Figs 2.2a and b show schematics of video signals. Each time the tv "sees" a line sync it moves the electron beam back to the left hand side and starts scanning left to right again whilst when a field sync is seen the electron beam is reset to the top left hand corner and scanning recommences.

2.5.3 Video Cameras

In the first instance for simplicity and economic reasons standard commercially available video cameras were investigated for suitability for use in the system. Due to their high noise and low dynamic range vidicons were immediately ruled out. The most suitable devices were found to be CCD based.

CCD devices are based on purpose built chips which can be thought of as a two dimensional array of photodiodes in which each pixel i.e. individual light detector is isolated from its neighbours and converts incident light to electric charge. This electric charge is then "read" on a point by point basis and converted to the video signal. CCD devices are produced in two forms, so called line or frame transfer devices. The designations of each device refers to how the charge on each pixel is read and the subsequent video image formed.

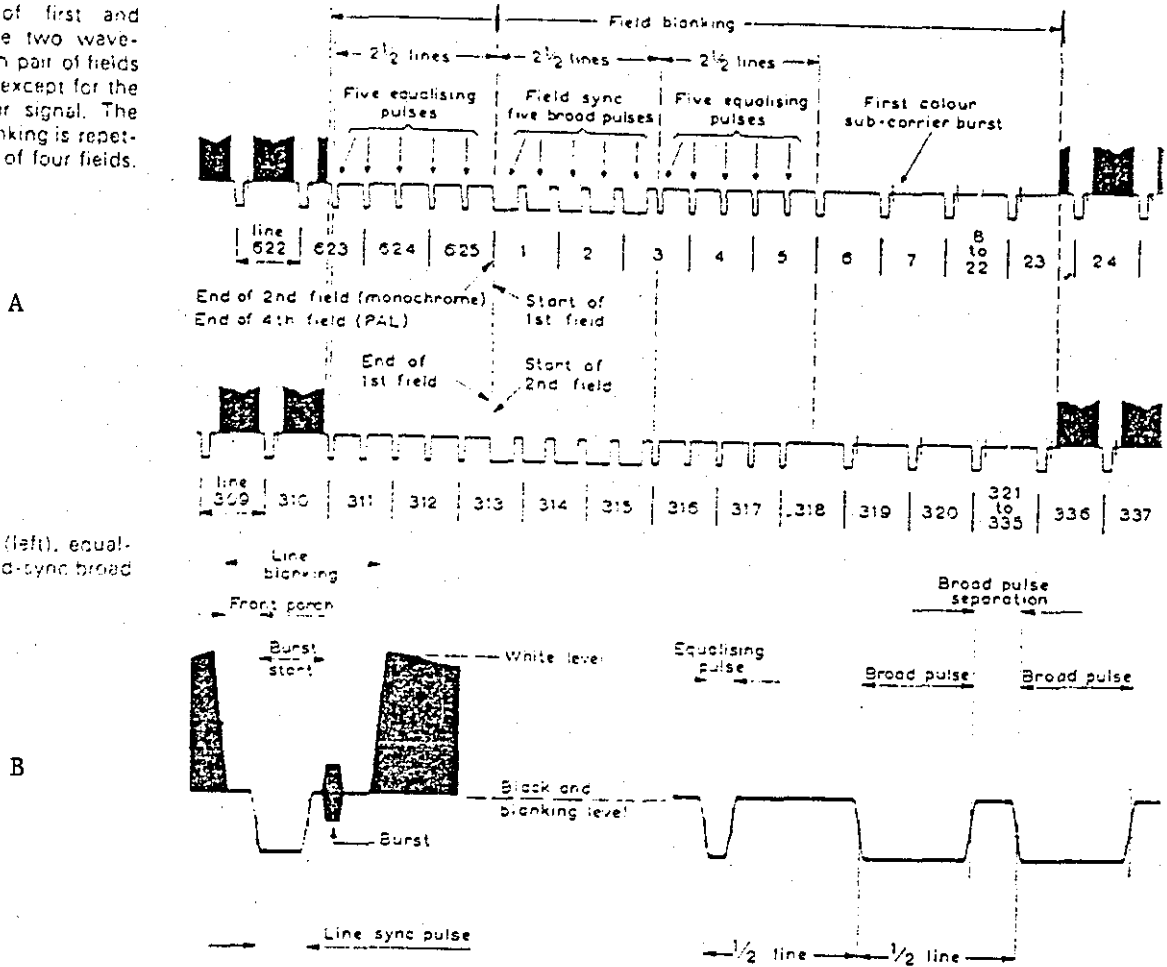
In line transfer devices (LTD) all of the pixels image for typically 2ms and then charge readout and conversion into the video signal starts at the top of the chip and works down so that by the end of each 20ms period every pixel has been read and cleared for the next image. In these devices

Fig 2.2

Video format (supplied by camera manufacturers)

625-Line Monochrome and PAL Colour Waveforms

Commencement of first and second fields. The two waveforms recur in each pair of fields making a picture, except for the bursts in a colour signal. The timing of burst blanking is repetitive in sequences of four fields.



Waveform Durations (between half-amplitude points)

Line period	64 µs nominal.
Line blanking	12.05 ± 0.25 µs.
Front porch	1.55 ± 0.25 µs.
Line-sync pulse	4.7 ± 0.1 µs.
Field period	20 ms.
Field blanking	25 lines + 12.05 µs.
Equalising pulse	2.35 ± 0.1 µs.
Separation between field-sync broad pulses	4.7 ± 0.1 µs.

Burst Parameters (Colour Waveform)

Waveform	10 ± 1 cycles of colour subcarrier.
Start	5.5 µs. + 0.2 - 0.0 µs. after the leading edge of the line-sync pulse. The phase of the subcarrier wave has no specified relationship with the line-sync pulse.
Amplitude	0.3 V p-p ± 0.25 dB.

Burst Blanking and Burst Phase Sequence (Colour Waveform)

Bursts are omitted from nine-line periods commencing in advance of or coincident with field blanking, and are introduced in successive series of four fields as follows:

- First field: on lines 7 to 309
- Second field: on lines 319 to 621
- Third field: on lines 6 to 310
- Fourth field: on lines 320 to 622

Times of Rise and Fall : 10-90°, full amplitude

Line-sync pulse	0.25 ± 0.05 µs.
Equalising pulse	0.25 ± 0.05 µs.
Field-sync broad pulse	0.25 ± 0.05 µs.
Line blanking	0.3 ± 0.1 µs.
Field blanking	0.3 ± 0.1 µs.

Amplitudes, Relative to Blanking Level

White level	+0.7 V ± 0.25 dB.
Sync level	-0.3 V ± 0.25 dB.

Frequencies

Line (f _L)	15.625 Hz.
Field (f _F)	50 Hz.
Colour subcarrier (f _{sc})	4.433.618.75 ± 1 Hz.

$$f_{sc} = \left(\frac{567}{2} + 1 \right) f_L + \frac{1}{2} f_F$$

The first burst waveform in each field is at 135 degrees relative to phase reference, and following bursts are alternately at 225 i.e. -135; degrees and 135 degrees.

imaging i.e. the conversion of light to charge occurs at the same time as other pixels are being read which means that in order to obtain a complete image in one field, of an object that covers all of the CCD chip when using a pulsed light source, the light pulse must occur in the first 2ms of the given field.

Frame transfer devices (FTD) have all of the pixels imaging for 18 ms and then imaging is stopped by applying suitable biasing pulses to the pixels. During the last 2 ms of a field, the charge on each pixel is transferred to what is known as the storage area on the CCD chip on a point by point basis. The storage area is identical to the imaging area of the chip except it is protected from light. In these devices conversion to and readout of the video signal is performed from the storage area and takes place during the frame after the imaging, i.e. the video signal coming from a FTD is 1 field (20ms) behind what is happening in real time. Hence if the event to be imaged occurs at time t the video signal of the image occurs at time $t+20$ ms.

Initially a Hitachi KP 120U LTD was used for the beam profiling system but later on this was replaced by an EEV P4310 FTD for 2 reasons. Since a FTD images for 18 ms the requirements on the synchronization of the excimer pulse to the camera are far less severe than for LTD's and secondly the EEV FTD's are approximately 50% more sensitive and suffer far less from blooming. Blooming is the effect whereby charge from a pixel that has exceeded its maximum charge holding capacity (i.e. saturated due to too much light) leaks out to corrupt adjacent pixels and produce a distorted image.

The cameras used all had dynamic ranges greater than 200 and the EEV P4310 had no automatic gain control (agc) i.e. built in electronics to change the amplification of the output signal depending upon the amount of incident light. The Hitachi KP 120U had an agc but averages for over ≈ 2 seconds before changing the gain. Hence given that the camera always sat in darkness before the excimer pulse, and that the produced fluorescence was of $\approx 1\mu$ s duration i.e. much less than 1 field the agc did not have time to respond and so could not effect the results.

Given that the monitor used to view the images had a screen 280 by 210 mm and that the CCD chip is 8.8 by 6.6 mm the images as viewed are effectively magnified 32 times (neglecting any initial optical magnification).

As discussed in 2.5.2 a tv picture (frame) is made up of 2 interlaced fields each 20 ms long. To obtain this interlacing, each pixel on the CCD chip is actually made into 2 separate sensors so that the data for successive fields comes from different halves of the pixel. This splitting of a pixel into 2 which, is done along a horizontal line, causes the data for

successive fields to be obtained from sensors that are vertically displaced from each other by $\approx 10 \mu\text{m}$.

This then means that for successive fields the data for the same point in the video image has come from a different point in space and therefore is slightly different and results in what is termed "the half pixel shift" whereby, if 2 images of the same surface with identical illumination are captured on successive fields or such that one is on the odd field and the other on the even then, due to this shift subtraction of the images yields a non zero result.

Obviously for the detection of changes on a sample this shift must be eliminated and so related images are always captured with time delays that are multiples of 40 ms which ensures the same field is captured each time.

2.5.4 Video Framestore

The heart of the system is an ImageIII video framestore produced by Eltime Ltd in England. The framestore is based upon the VI500 single board field store developed by British Telecom Research Division and is a computer controlled digital picture store with variable picture resolution and 6 bit A/D and D/A converters giving 64 intensity levels known as greylevels. For beam profiling purposes the store is operated in the "2 image" mode where it is capable of storing 2 separate images each of 512 (horizontal) by 256 (vertical) points whereas for the work in chapter 4 the board was run in "4 image" mode with each of the four possible images being 256 by 256 points. By suitable circuitry the board can be interfaced to either Apple II or IBM PC type computers.

The board when requested, either via the computer or from an external trigger source will perform a real time full field capture (snatch). The snatch is synchronized to and starts on the first field sync pulse after the request, and the entire incoming video signal for that field is captured and digitised. The digitisation is performed so that level 0 corresponds to 0.3 volts on the video signal and level 63 to 0.7 volts.

Once "in" the store an image can be thought of as being in the form of a 2 dimensional array with parameters x (horizontal) and y (vertical). The image is arranged so that $x=y=0$ is the top left hand point in the image, with increasing x running left to right and increasing y top to bottom.

By suitable computer routines it is possible to write to or read from any point in an image and thereby perform such operations as image processing and reading/writing images to computer disc storage.

2.5.5 Computer Control

At the time when the beam profiling system was initially written the only available computer was an Apple IIe based on a 6502A 8 bit processor with 64k of ram and twin floppy discs each of which could hold 170k of data. Later on IBM PC type machines became available and parts of the original profiling system were rewritten for use on these faster and more sophisticated machines.

2.5.6 Timing Requirements for the System

Since the excimer induced fluorescence lasts of the order of 1 μ s (see 2.5.9) i.e. all of the information is contained within 1 video field, there must be appropriate synchronization of both the excimer pulse and the framestore snatch request to the field sync pulses coming from the camera.

To perform this synchronization an electronic trigger box (later modified for the work in chapter 4) based upon previous work by a colleague¹⁴ was build. Fig 2.3a shows the basic circuit diagram for this box.

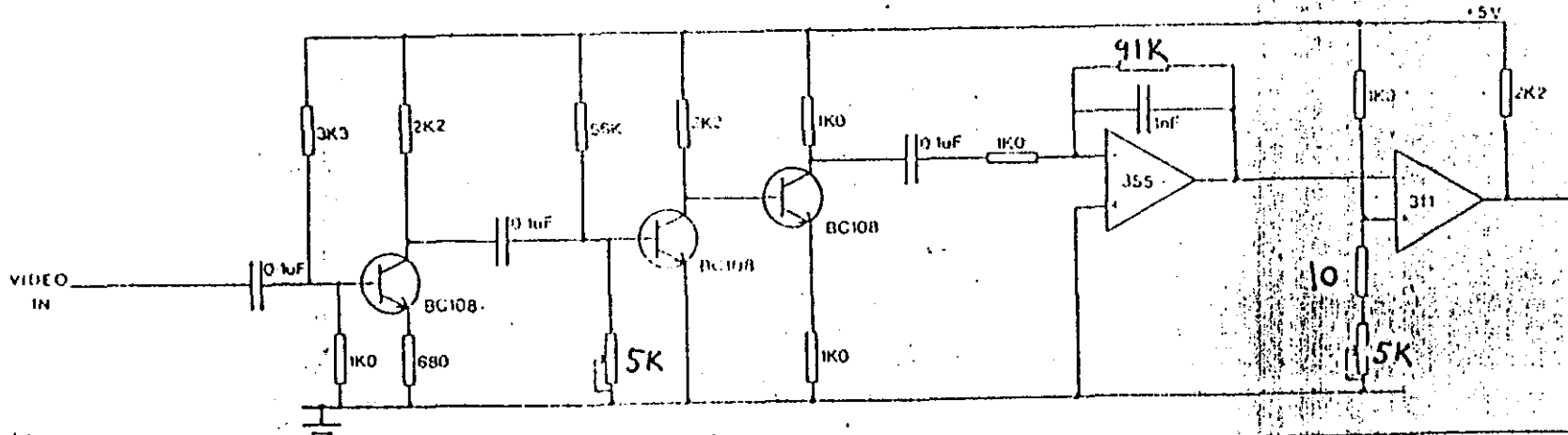
Part A of the circuit serves to separate the data and sync pulses, and then to separate the line and field syncs so that the output of the 311 chip is one pulse every 20ms corresponding in time to the presence of a field sync in the video signal. Closing switch s1 activates the 7474 so that on the arrival of the next and only the next field sync pulse, a pulse is provided to the 74121 that will initiate a snatch by the framsetore and, trigger the laser at the correct time depending upon the type of camera being used. Switch s2 is included to allow a choice between LTD and FTD cameras. Figs 2.3b and c show the timing sequences used for the two type of camera.

2.5.7 Video Copy Processor

To obtain a hardcopy of the images displayed by the framestore a Mitsubishi P60B video copy processor was purchased. This device provides hard copy prints from standard video signals with an intensity resolution of 16 levels. Unfortunately since the copier only has 16 levels but the framestore has 64 there is some loss of image detail in printing which occasionally means that details easily visible on the tv monitor are much harder to see on the prints. Printed images are reduced by a factor of 3.2 compared to the image on the monitor. Unless specified otherwise figures quoted for magnification are for viewing on the monitor and therefore prints produced by this machine, in particular all images in chapter 2 and chapter 4 have magnifications reduced by 3.2 from the values quoted in the text.

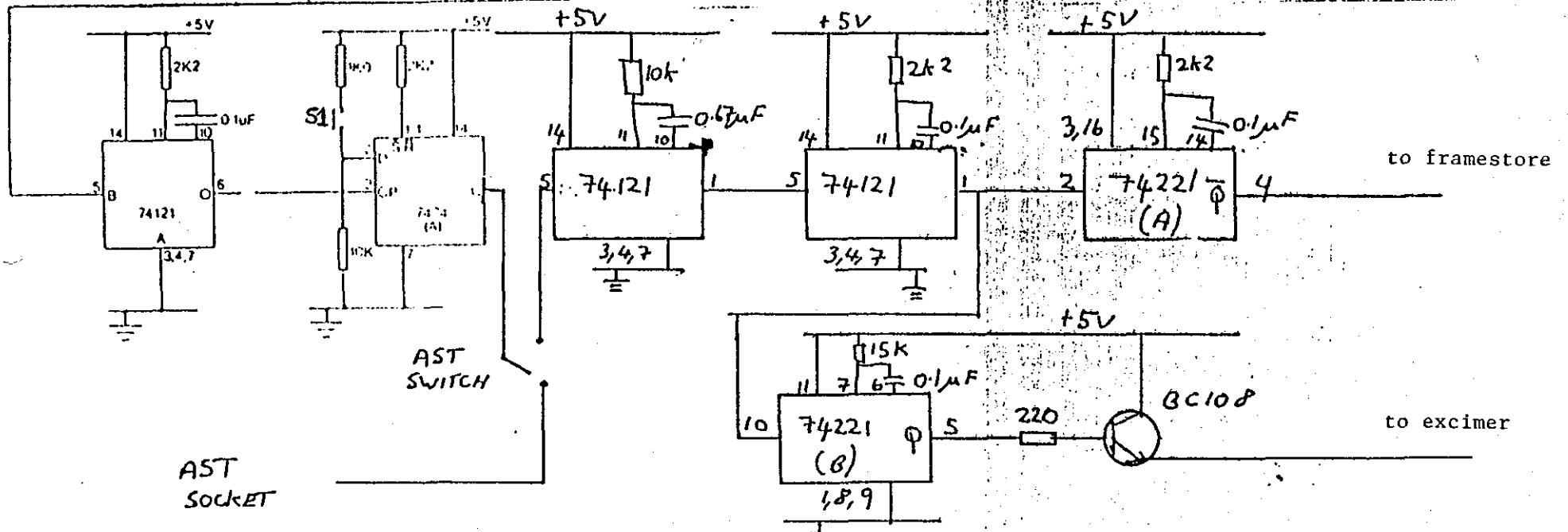
TRIGGER BOX MAIN CIRCUIT

Fig 2.3a



part A

38



to framestore

to excimer

AST
SOCKET

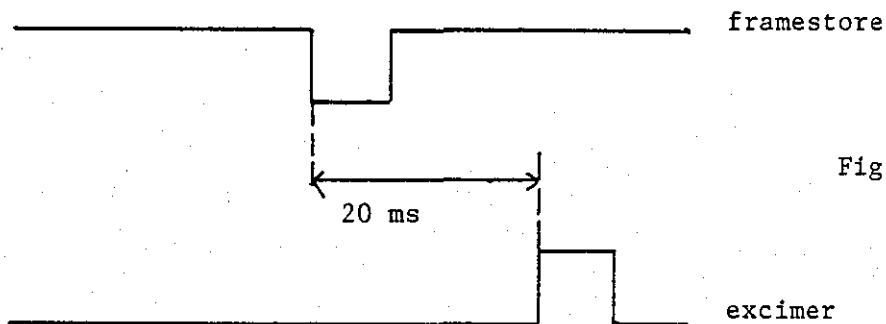


Fig 2.3b

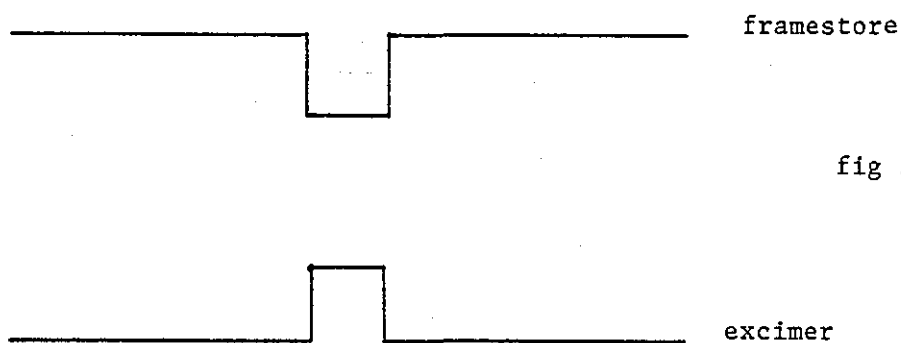


fig 2.3c

Timing sequences for beam profiling

2.3b line type cameras

2.3c CCD type cameras

Note that framestore triggers on a falling edge whereas the excimer triggers on a rising edge.

2.5.8 Noise Considerations

There are 2 sources of noise in the production of the viewed image, the noise associated with the camera and that with the digitisation in the framestore. Manufacturers data shows that the noise in the framestore is not more than 1 greylevel.

There are 3 sources of camera noise, shot noise associated with the incident light, variations in pixel sensitivity across the chip due to variations in the pixel size across the chip and the so called fixed pattern noise which arises from the accumulation of thermally generated electrons on each pixel. Manufacturers data states that the noise in converting the acquired charge to a video signal is much less than any of the above. Again from manufacturers data, the shot noise is at most much less 1% of the peak signal and can be ignored, whilst across the chip the other 2 sources of noise are each at most 2% of the peak signal.

The fixed pattern noise can be removed by subtracting 2 fields as this noise is constant for a given pixel provided the temperature remains the same and simply adds to the signal due to incident light. The noise due to variation in sensitivity can only be removed by division of fields after the fixed pattern noise has been removed.

Hence in an unprocessed image the noise is not more than 4% of the peak signal which once digitised corresponds to ≈ 2 levels. The subtraction routines used to process the data effectively half the noise and mean that there is a maximum of 2 greylevels of noise across the processed image, 1 level from the camera and the other from the digitisation noise. In a comparison of the same point between images, the only noise is the digitisation noise of around 1 level since the variation in pixel sensitivity effects only comparisons between different points. Hence if the level of a given point changes by more than 1 level then this is data and not noise but to compare different points across the image the change must be more than 2 levels to guarantee that it is due to data and not noise.

2.5.9 Fluorescence and Fluorescers

To convert the UV light to a form that the camera can detect a fluorescing medium is used. Fluorescence can be defined as the emission of radiation (usually optical) due to an electronic transition between two energy levels (states) as a result of the excitation of a solid by absorption of photons where, the time duration of the emission is approximately equal to the lifetime of the transition between the levels concerned. It should be noted that often fluorescence results from so called "forbidden transitions" i.e. non electric dipole transitions and therefore it's lifetime can be much

longer (of order 1-10 μs) than the $\approx 10\text{ns}$ associated with dipole transitions.

Early work performed at the Rutherford Appleton Laboratory¹⁵ showed that the most suitable fluorescing mediums for use with a KrF laser were glass based. In particular the work concentrated on commercial Borosilicate glass which has significant Fe^{2+} impurities. It is these Fe^{2+} ions that act as extremely good absorbers¹⁶ of the UV light thus making the penetration depth of the light very small ($\approx 1\mu\text{m}$) and ensuring that a good two dimensional fluorescent image of the beam is produced. In this early work¹⁵ it was shown that borosilicate glass was linear in its conversion of UV light to visible fluorescence over the entire range of excimer fluences used at the time i.e. $10 \mu\text{J cm}^{-2}$ to 10mJ cm^{-2} . Later work performed by the author (see 2.5.12b) showed that when the excimer fluence exceeded approximately 50mJ cm^{-2} the fluorescent image of the excimer beam showed saturation effects, and therefore in practice, profiles were recorded using excimer fluences of $\approx 40\text{-}50 \text{mJ cm}^{-2}$. Further work at the Rutherford Laboratory¹⁶ showed that for this type of glass the fluorescence had a lifetime of $\approx 1\mu\text{s}$.

For use in the system the borosilicate glass was in the form of a thin ($\approx 100 \mu\text{m}$) microscope cover slip approximately 2 cm square. By using the glass in the form of a thin cover slip a good image of the beam was ensured as problems such as distortion due to propagation in the glass were essentially eliminated.

2.5.10 Operation of the system

To obtain the profile of the excimer laser in any plane the front surface of the fluorescer was first accurately positioned at the required point in space. This positioning was accomplished by measuring the distance from the test plane to a convenient fixed point e.g. the holder of the excimer focussing lens by using a micrometer type depth gauge. The fluorescer was then positioned so that the distance from its front surface to the fixed point was the same as that of the test plane. Experience showed that with care a reproducibility of better than $30 \mu\text{m}$ could be achieved.

By using a HeNe laser for illumination and working in transmission (see fig 2.4), the video camera was focussed onto the front surface of the fluorescer. Typically, due to the limited fluorescence of the glass, the system was operated with an optical magnification of only 1-2 as higher values resulted in very weak images at non saturating excimer fluences.

The setting of the excimer fluence each time a beam profile was required was done by trial and error coupled with past experience in the following way. Experiments (2.5.11b) on the variation of the peak grey level in the fluorescent image as a function of excimer fluence showed that saturation of

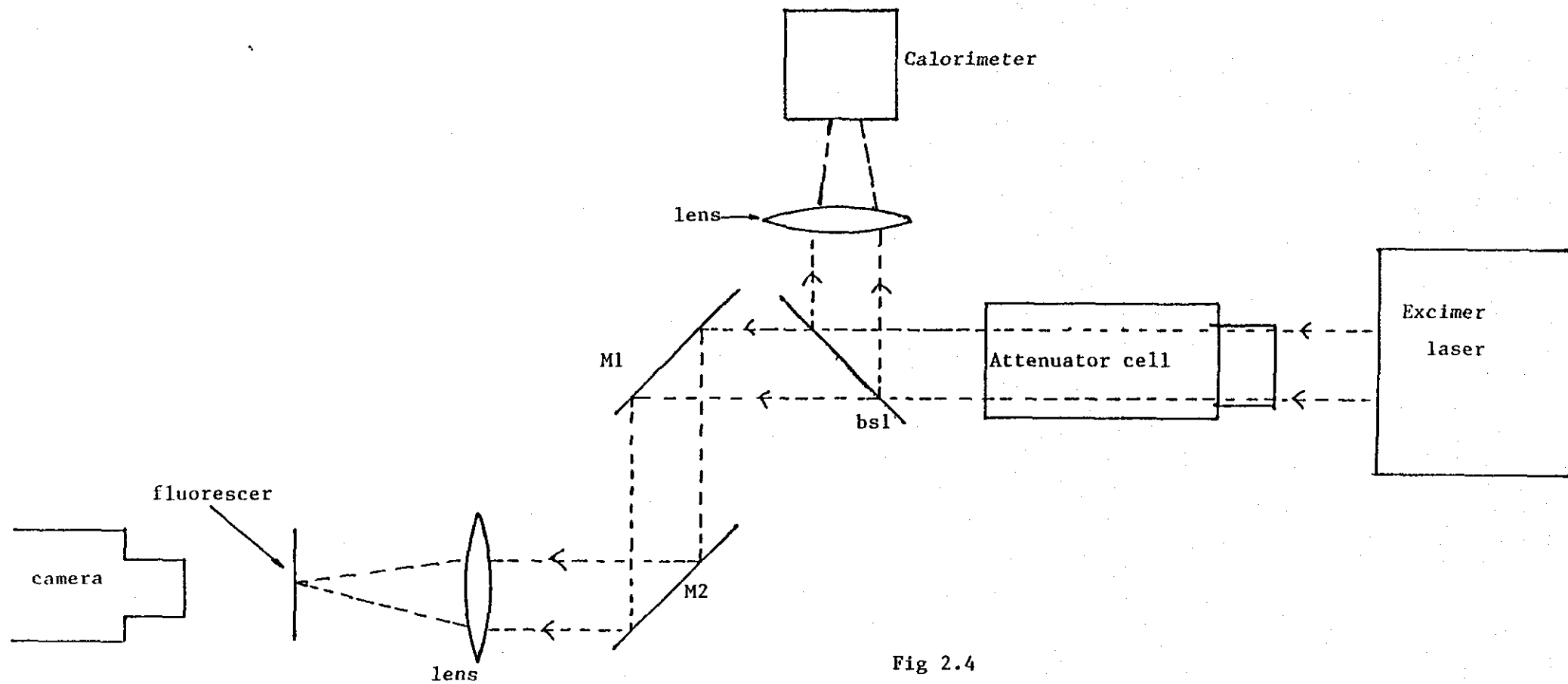


Fig 2.4

Experimental layout for beam profiling

M1, M2 = steering mirrors

bsl = beamsplitter

of the fluorescence started to occur above level 30. Therefore the excimer fluence was adjusted so that the peak in the image was between 20 and 28 so that not only was the image a linear representation of the excimer beam but also to ensure that there were enough greylevels in the image to enable the $1/e^2$ points to be found accurately and for there to be an adequate signal to noise ratio. This was achieved by repeatedly capturing and profiling images until the peak level was correct (see 2.5.12b).

Once the excimer energy was set the trigger box was adjusted to give the correct timing sequence and the fluorescent image was then captured by the framestore. Once in the framestore image processing could begin.

Prior to obtaining any profile the system had to be calibrated in terms of the number of pixels per cm in the horizontal and vertical directions which, due to the curious way in which the tv and frame store combine to present an image are not the same, with the ratio between them being typically $3/2$. To do this the fluorescer was replaced by a piece of graph paper which was positioned (using the depth gauge) so as to be in the same plane as the front surface of the fluorescer. By using HeNe illumination the camera was focussed onto the graph paper and an image of it captured on the framestore. By use of the horizontal and vertical profiling programs it was possible to move cursors around so that the number of pixels corresponding to a mm in the image could be found.

2.5.11 Computer Routines

2.5.11a Computer Processing on the Apple IIe

Two programming languages were available on the Apple, USCD Pascal and Apple BASIC. It is important to note that there was no "stand alone" assembler available for the Apple, the assembly programs that were written were done so from within either with the BASIC or Pascal operating systems. This meant that assembly level programs could not be run by themselves but had to be linked into the appropriate high level program.

Since operation of the framestore/computer interface requires reading from and writing to some of the input/output (I/O) ports of the Apple it was decided to write the programs in a combination of USCD Pascal linked to appropriate 6502A assembly language routines for 2 reasons.

1) In Apple BASIC there is no direct access to the I/O ports and therefore an assembly level program would have to be linked to the BASIC program. To use linked programs in this way from BASIC requires knowing the absolute memory location of the start of the low level program. This is essentially

impossible to determine since factors such as the number of program variables, the size of the program etc all have an effect on where the start of the program can be placed. In Pascal the operating system automatically sorts out the location of programs within memory, the programmer does not have to know absolute locations.

2) The file handling capabilities of BASIC are very limited and use of files from BASIC is very slow since only a single number can be written at once, whereas in Pascal numbers can be transferred in blocks of 256, greatly reducing the time involved. Initial tests showed that it took approximately 25 minutes to save a 512 by 256 pixel image to floppy disc whereas from Pascal saving the same image took 63 seconds.

Assembly language was used for simple input/output routines, basic image processing such as subtraction and enhancement and obtaining the data for graphical display/statistical analysis. Pascal was used for all disc input/output, hardcopy printout routines, performing statistical analysis of the image and for generating the graphics from the data obtained using the assembly level programs.

Although faster in execution, assembly language programs take far longer to write and in particular to debug than high level programs. This is because not only does any individual statement do less than a high level statement but the programmer directly has to set and be careful of the "flags" and "status registers" associated with the processor.

2.5.11b Data Acquisition/Processing Routines

These consisted of routines to ...

1) capture and subsequently display an image upon pressing a key on the keyboard or, to display an image resulting from a trigger box initiated capture.

2) subtract 2 images on the basis that if image 1 (store a) has level A at a point x,y and image 2 (store b) has level B at point x,y then, written to store a is the subtracted image with level B-A at point x,y.

3) read images from the framestore to disc and vice versa.

2.5.11c Profiling Routines

In order to actually profile a beam, the intensity along a given axis is required. The routines allowed profiling in 4 ways, horizontal

,vertical,radial and angular.

The horizontal and vertical modes allow profiling along horizontal/vertical lines respectively whilst the angular mode allows profiling along a line at any angle whose center can be anywhere in the image. The radial mode determines the average intensity value around a circle as a function of the radius of that circle with the centre being any chosen point in the image.

In the horizontal mode the user would move a cursor in the form of a horizontal line up or down the screen until the desired position was reached. Upon pressing "p" a graph representing the intensity variation along the chosen line was superimposed upon the image as in fig 2.5b. If the user was not satisfied with the position chosen, the cursor could be further moved but, not before the areas overwritten by the graph had been restored to their original values which, had been temporarily stored in the computer's memory.

Once the user was satisfied with the position of the cursor pressing "g" would transfer the data to the computer for further processing and graphical display. With the data now in the computer's memory processing was carried out to find 1) the maximum greylevel (intensity) and its position, 2) the position and levels of the $1/e^2$ points and 3) the total number of greylevels contained in the selected profile.

Using the built in Apple graphics, a graph of the profile with the above parameters superimposed on it was generated and displayed on the computer monitor. A printout routine allowed this data to be "dumped" to an Epson FX 80 printer. Fig 2.7a shows such a graph for the excimer beam profile shown in figs 2.5a-c. On each graph the upper number labelled 1 is the total number of greylevels in the profile whilst the lower numbers labelled 2 refer to the maximum and $1/e^2$ points. This latter data is in the form of position/level and from left to right corresponds to the left hand $1/e^2$, maximum and right hand $1/e^2$ points.

All of the profiling modes work in similar ways and perform similar analysis on the selected data. In practice horizontal and vertical profiling is more than sufficient to characterize the excimer here at Loughborough. The angular mode was written at the request of the Rutherford Appleton Laboratory to whom a version of the system was given and the radial profile at the request of the Los Alamos National Laboratory who bought an extended version of the system.

2.5.11d Statistical Analysis Routines

The purpose of these routines was essentially to convert the recorded greylevel at a point to a fluence i.e. $J\text{ cm}^{-2}$ and to provide such

information as peak and average fluences, fluence per greylevel.

The collection of the "raw" data contained in the image was performed using an assembly level program and is complicated by the fact that since the 6502 is only an 8 bit processor it can only handle directly numbers ≤ 255 and therefore special routines to emulate upto 24 bit addition and subtraction were written to cope with the size of number involved in the statistical routines.

The routines herein are complicated by the fact that it is not always possible to have all the light in the image corresponding to fluorescence induced by the excimer beam and therefore some form of allowance for this must be made. If for instance background light is allowed to fall on the camera or the camera were to pick up extraneous light such as that induced in the dye cell attenuator by the absorption of the excimer beam then the background level ("dark level") in the recorded image would not be zero. In the case of a lasers such as the one here at Loughborough which can be pulsed many times a minute it is very easy to remove this unwanted signal (noise) by simple subtraction of two frames, one that has the fluorescence and noise on and a second of just the noise. However, this method of removal is not practical on the big experimental lasers such as "SPRITE" at the Rutherford and "AURORA" at Los Alamos where the laser is working very well if it can be fired once ever 30 minutes.

In order to convert greylevels to fluences the user must supply several parameters, specifically the number of pixels per centimetre in the horizontal and vertical directions, n_x and n_y respectively and the excimer energy (E) that was used to generate the fluorescence. From earlier profiling measurements the user will have been able to determine the background level g_b which, will usually be near if not actually zero.

A routine was written to allow the user if so desired to specify the area of the image from which the data for analysis was to be obtained from rather than using the entire screen which is the default mode.

The routines first determine the number of pixels n that have a level g_i that exceeds the background level g_b and thus represent actual data rather than noise. The total number of data greylevels g_t and the sum of the squares of the greylevels g_{ts} are then determined as follows...

$$g_t = \sum_i n_i * (g_i - g_b)$$

$$g_{ts} = \sum_i n_i * (g_i - g_b)^2$$

where n_i is the number of pixels that have greylevel g_i and such that

$$n = \sum_i n_i$$

with the summation being carried out with i varying from g_b+1 to 63. Hence assuming that at any point the induced fluorescence is directly proportional to the incident excimer fluence it is possible to convert greylevels to fluences as shown below.

The energy per greylevel e_g is given by $e_g = E/g_t$ and therefore the energy (due to the excimer laser rather than noise) e_p falling on a pixel of level g_i is given by

$$e_p = E*(g_i - g_b)/g_t$$

The area in cm^2 of a pixel is $1/(n_x*n_y)$ and therefore the fluence at any pixel f_p is given by

$$f_p = E*(n_x*n_y)*(g_i - g_b)/g_t$$

The peak fluence is then found by finding the peak greylevel in the signal, whilst the average fluence f_{av} is found from the expression...

$$f_{av} = E*(n_x*n_y)/n$$

The standard deviation of beam fluence f_{sd} is also determined by use of the following expression.....

$$f_{sd} = f_{av}*((n*g_t)/g_t^2 - 1)^{0.5}$$

Once determined, all these values are then displayed and printed if required by the user. Fig 2.7b shows the output of the statistical analysis corresponding to the excimer profile in fig 2.5a-c.

2.5.12 Results

2.5.12a Results 1

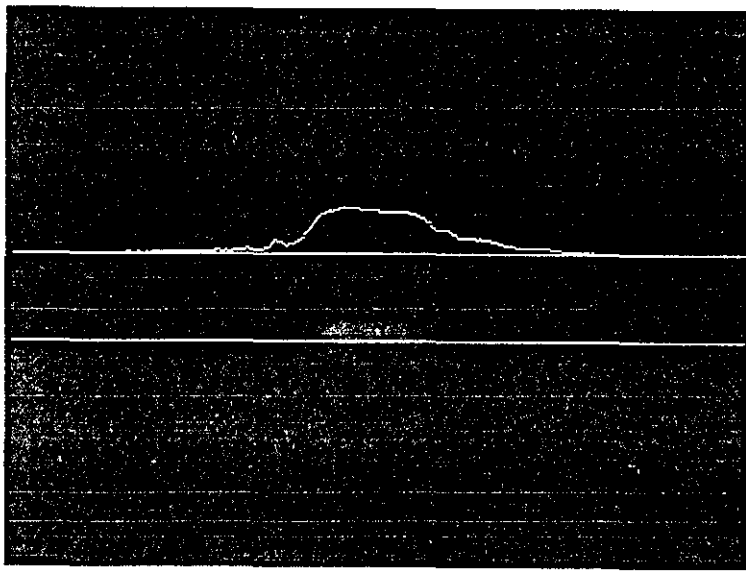
Fig 2.5a is a recorded image of the excimer with the fluence low enough that the fluorescence is linear whilst figs 2.5b and c are respectively horizontal and vertical profiles of the excimer beam. Fig 2.6a is the image of a 5 mW cw HeNe laser with fig 2.6b showing an angular profile of the beam. The graph on the vertical line on the left hand side is a vertical profile of the beam along the angled line whilst that on the horizontal line is a horizontal profile of the beam along the angled line. The 2 profiles

Fig 2.5

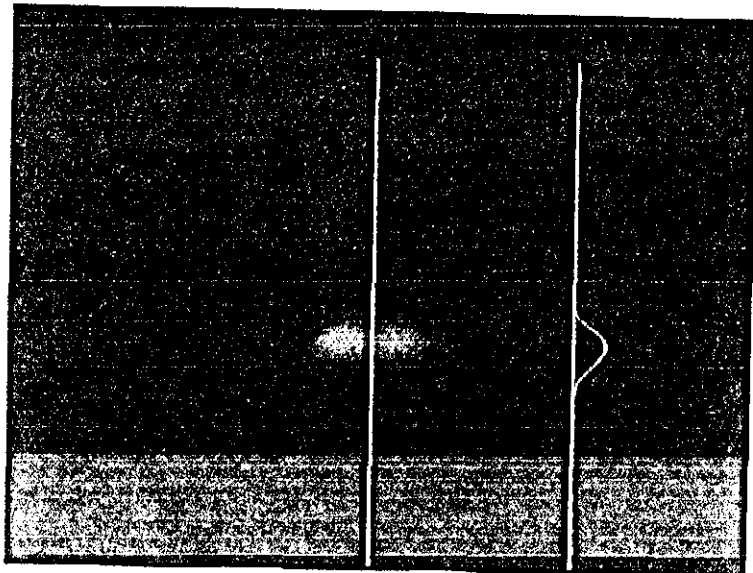
Unsaturated excimer
profile



A



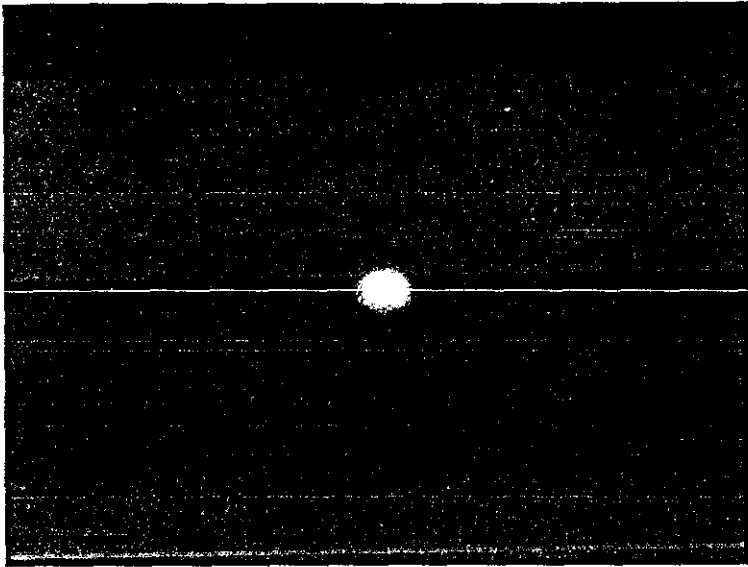
B



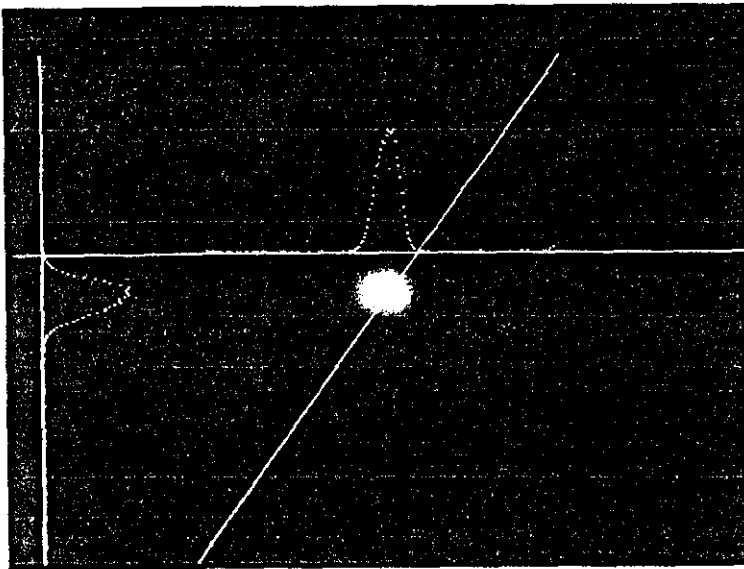
C

Fig 2.6

Profile of HeNe
laser



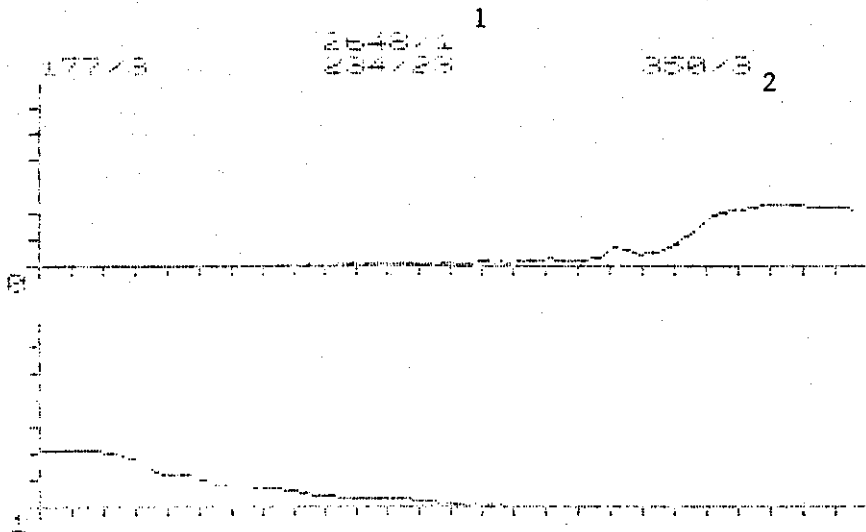
A



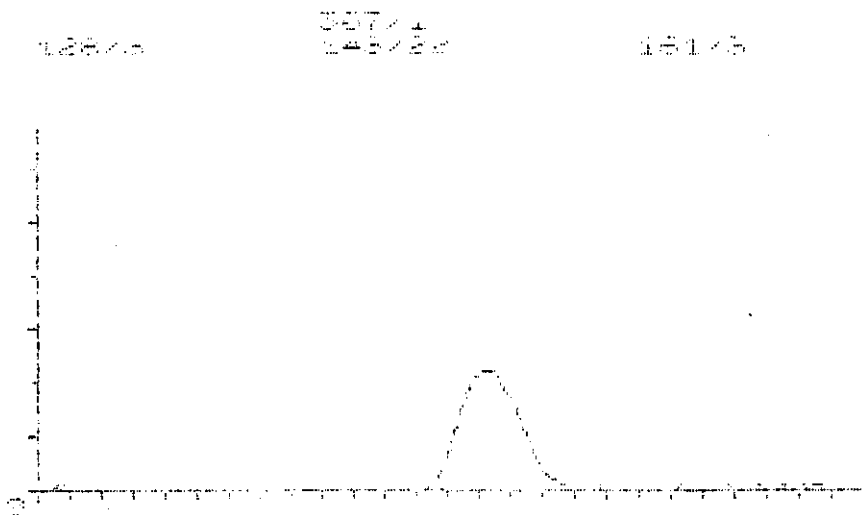
B

Fig 2.7a
Profile of excimer beam

CURSOR NOW AT X=114 Horizontal
 SCALES ARE DIVIDED INTO UNITS OF TEN.....



PRIF
 CURSOR NOW AT X=249 Vertical
 SCALES ARE DIVIDED INTO INTERVALS OF TEN UNITS....



STAT

STARTING LEVEL=1

NUMBER OF LEVELS=1000

FLUENCE FOR START LEVEL= 1.149608E1 J/CM²

AVERAGE FLUENCE= 7.774245E1 J/CM²

STANDARD DEVIATION OF FLUENCE= 7.810355E1 J/CM²

HIGHEST FLUENCE= 2.644555E2 J/CM²

MOST FREQUENT FLUENCE= 3.175645E1 J/CM²

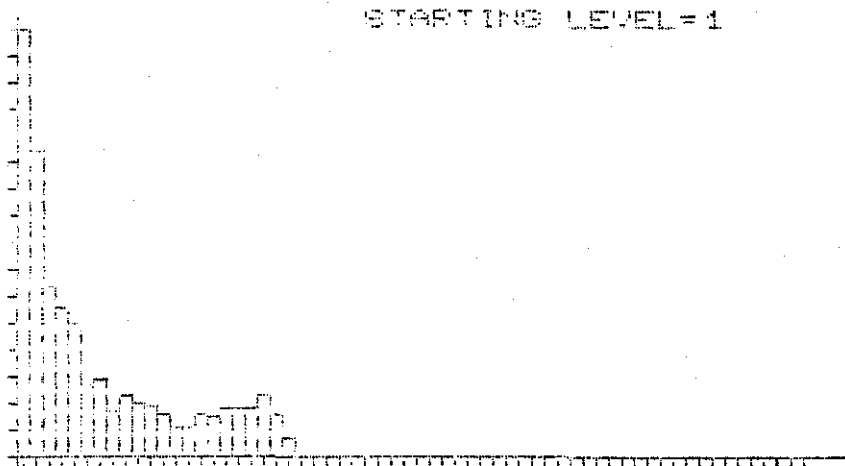


Fig 2.7b

Statistical analysis of excimer beam profile

would, once transferred to the computer be suitably combined to yield the profile of the laser along the angled line.

2.5.12b Results 2

The variation of the peak greylevel in the image and the width at the $1/e^2$ points at a constant distance from the focussing lens as a function of excimer energy on the fluorescer was investigated by capturing and profiling the fluorescence for a number of laser energies, yielding the results shown in table 2.1 and graphically in fig 2.8.

Table 2.1

Energy*	Peak greylevel	Width at $1/e^2$ points	
		H	V pixels**
μJ			
67	12	168	30
135	23	170	32
176	30	170	32
280	35	203	34
418	37	266	38
627	40	290	42
978	42	307	43

*this is the energy that is incident on the fluorescer.

**number are for the horizontal (H) then vertical (V) widths in pixels at the $1/e^2$ points.

It is immediately apparent from table 2.1 and fig 2.8 that the fluorescence shows saturation effects i.e. the relationship between fluorescence and excimer energy becomes non linear at greylevels above ≈ 30 corresponding to an energy of 0.176 mJ. Given that in this data was recorded such that there were 950 pixels per cm horizontally and 650 per cm vertically, this means that the peak greylevel corresponds to a fluence of $\approx 50 \text{ mJ cm}^{-2}$.

Figs 2.9a-c are a set of profiles for the excimer beam recorded at an excimer fluence well in excess of 50 mJ cm^{-2} and the effects of saturation i.e. the "flattening out" of the top of the profiles is obvious.

If saturation is allowed to occur then it has two related effects, firstly the width at the $1/e^2$ points increases and secondly, the peak energy as determined by the statistical routines decreases markedly.

Hence in operation when using the 20 cm focussing lens the excimer energy was always kept well below 176 μJ . Some of the work performed by colleagues requires greater fluences than are possible with a 20 cm focal length lens

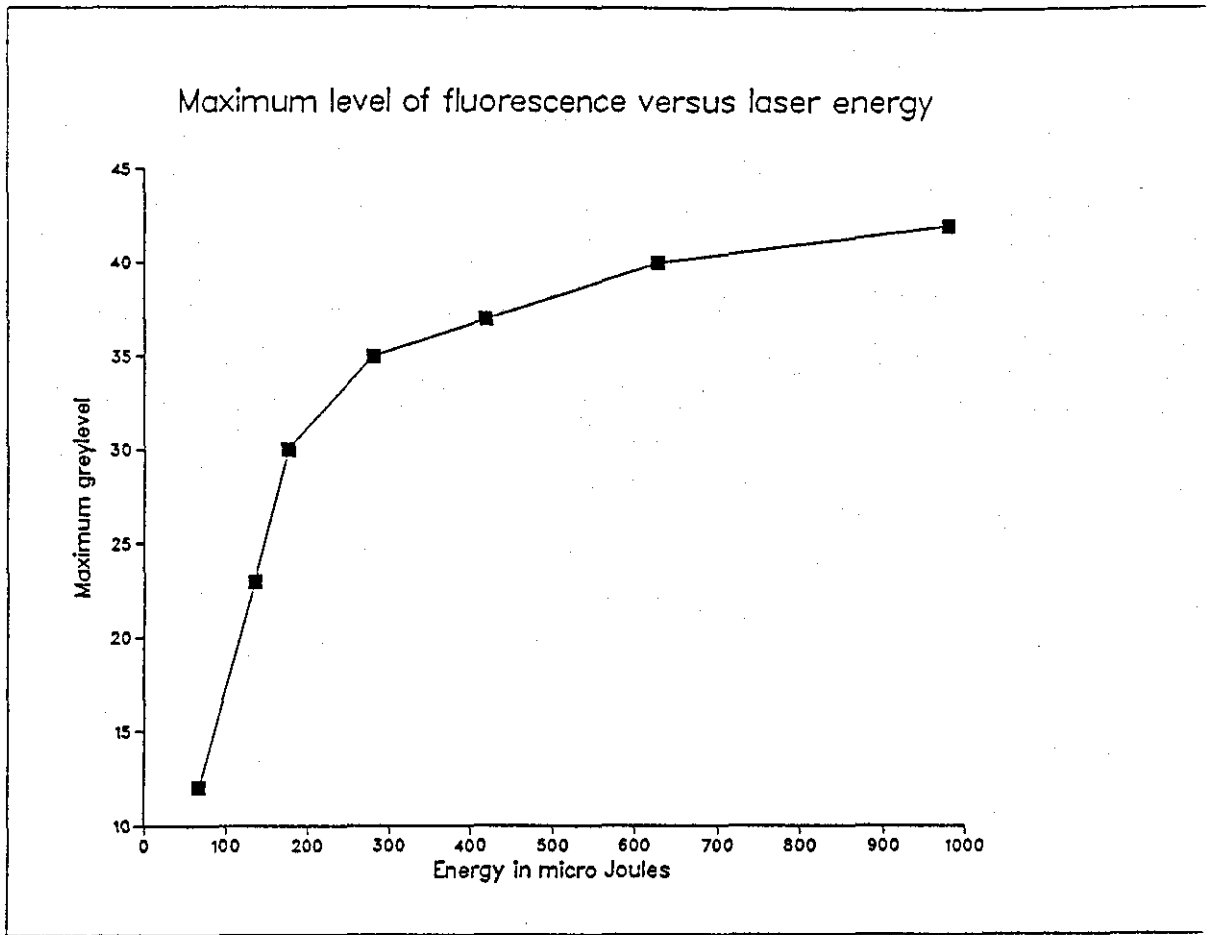
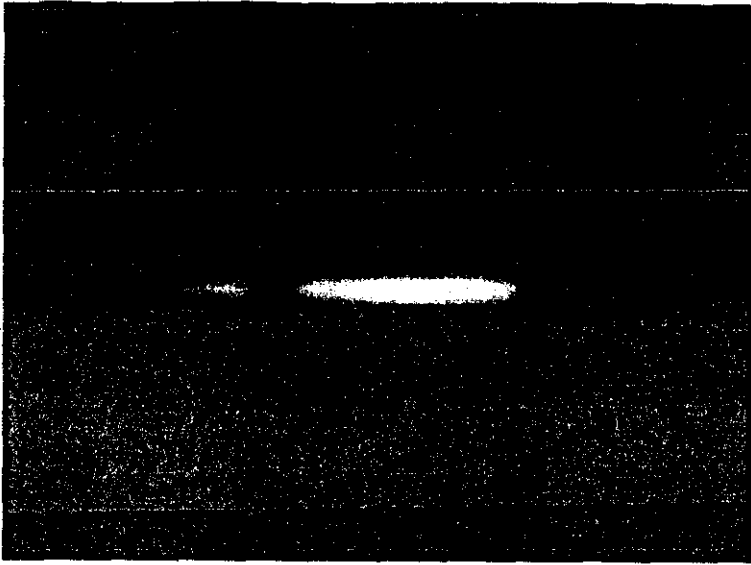


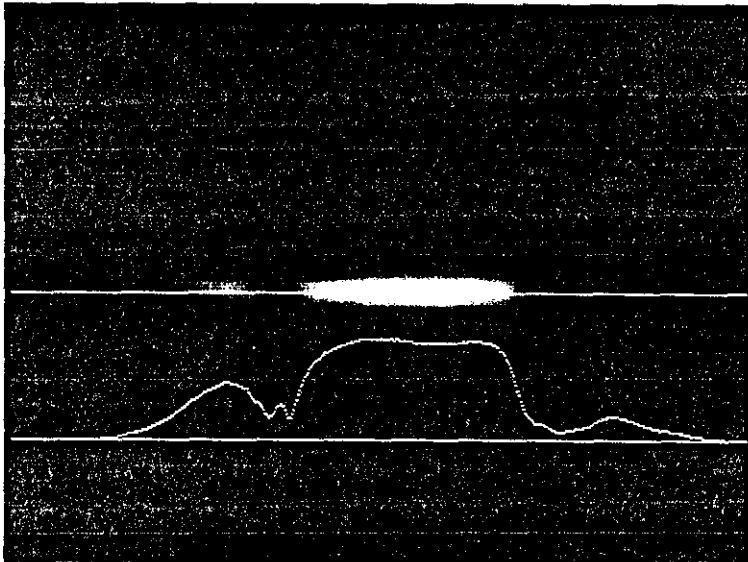
Fig 2.8

Note the saturation of the fluorescence at level 30 and above

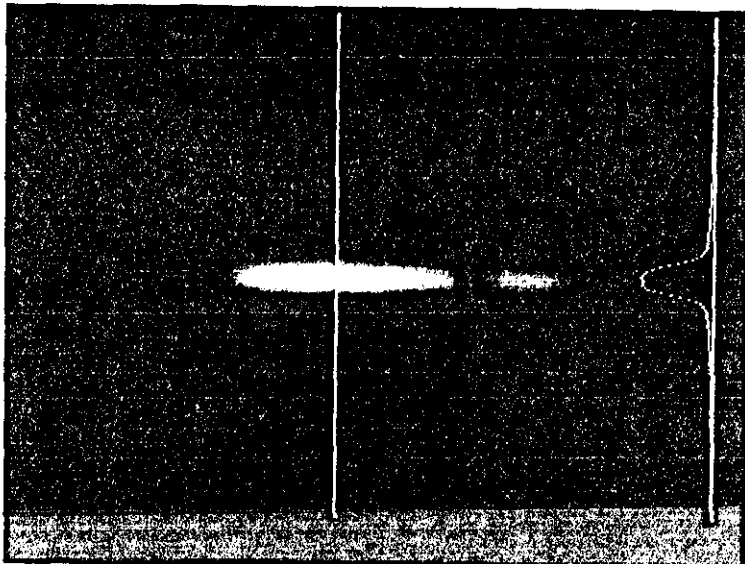
Fig 2.9
Saturated excimer
profile



A



B



C

and therefore require a shorter focal length lens. Since the area of the laser spot at or near focus scales approximately as f^2 , the maximum allowable excimer fluence if saturation is not to occur also scales as f^2 .

2.5.12c Results 3

In order to actually determine the location of the focus, an approximate position was found by viewing the fluorescence induced on a white card placed in the beam and finding the position where the beam appeared smallest. The profiling system was used to record profiles typically every 1 mm for 1 cm either side of this point thus obtaining the variation of spot size around the focus. For the 20 cm focussing lens used throughout, the spot size at focus is ≈ 0.9 mm. Table 2.2 shows the variation in spot size and peak fluence with position from a convenient fixed point.

Table 2.2

Position mm	Spot Area mm ²	Peak Fluence J cm ⁻²
-4	1.30	213
-2	1.24	226
0	1.10	261
1	0.95	272
2	0.94	286
3	0.95	277
4	0.90	288
6	0.89	267
8	0.91	261
10	1.06	215

It is clear from table 2.2 that for several mm around the focal point the spot area and peak fluence do not vary appreciably (i.e. by more than 5%). This is of particular significance for the work in chapter 3 where the sample was tilted at large angles to the beam and therefore interacted with the beam over a distance of several mm around the focal plane and means that the fluence over the whole of the interaction area is very nearly constant.

It is interesting to note that the peak fluence and smallest spot area do not occur at the same position. This is probably the result of both the beam being intrinsically non Gaussian and aberrations/diffraction due to the focussing lens.

2.5.12d Results 4

To illustrate the versatility of the basic ideas involved in the beam profiling system fig 2.10a is the image of a gain switched TEA CO₂ laser obtained by imaging the "burn off" of a graphite block upon which the laser pulse was incident. Figs 2.10b and c are the profiles for this laser and show that the laser was not ideally setup as it appears to be running in the TEM₀₁ mode.

2.5.13 Profiling and Damaging Simultaneously

The operation of the system as described in 2.5.9 is such that profiling and damage cannot be performed simultaneously since the fluorescer is in place of the sample. For small lasers that generate highly reproducible pulses this does not matter, however in the case of large experiment lasers the rep rate is so low and pulse to pulse variations so great that a profile has to be recorded every shot whilst a sample is simultaneously being tested. This requires the beam to be split into two. The splitting can be done at one of two points, either before or after the focussing lens.

Splitting the beam before the lens requires using a second lens to focus the beam which has the problem of trying to match 2 lenses so that both produce identical beams which, in practice is very difficult to do with sufficient accuracy and therefore rarely done.

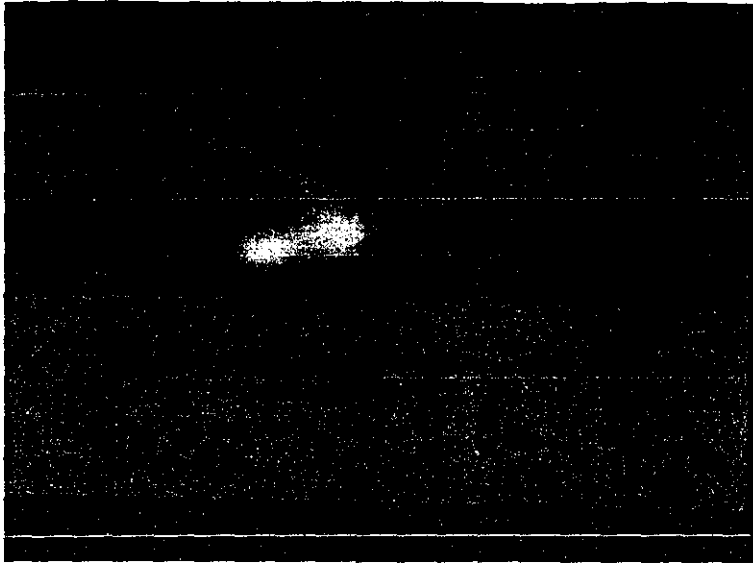
Splitting the beam after the focussing lens requires the insertion of a beamsplitter into a non parallel light beam which means that the transmitted and reflected beams will not have the same profile. On passage through the beamsplitter the transmitted beam will be distorted whilst the reflected beam will not and represents a true image of the beam. The amount of distortion whilst depending in a complicated way upon such things as the thickness of the beamsplitter, width of the beam, focal length of the lens, angle of incidence on the beamsplitter etc is, in general minimised by using wide beams, near normal incidence on the splitter and long focal length lenses. In the case of the large experimental lasers which have large diameter beams and use long ($\approx 2\text{m}$) focal length lenses, the distortion is sufficiently small that profiling and damage can be performed together.

2.5.14 Comments

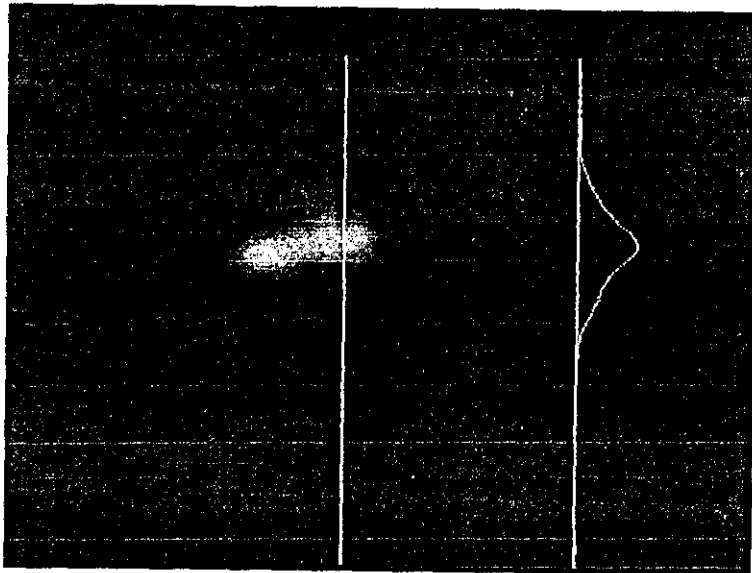
Once everything is setup it takes about 10 minutes from capturing the image to obtaining the profiles and beam parameters, a significant improvement on the 2-3 hours that it required to obtain a less well defined profile by the scanning pinhole method.

Fig 2.10

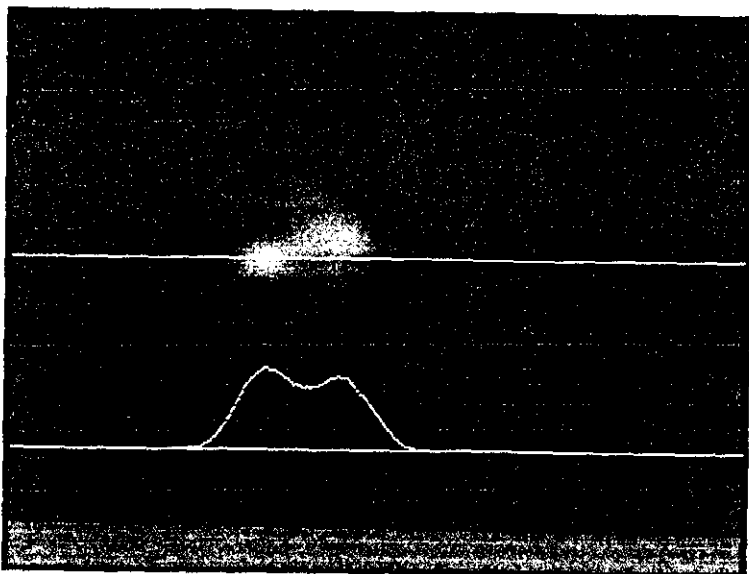
Profile of CO₂
laser



A



B



C

2.6 Damage Detection

2.6.1 Introduction

Initially when the damage facility was setup at Loughborough, damage detection was performed by direct visual observation of the surface and or by the use of a photodiode together with Schlieren techniques to detect changes in the scattered light coming from a test surface.

2.6.2 Schlieren Techniques

The presence of LID on the surface of a sample will result in a changes in the amplitude and or phase of the parts of an illuminating probe beam that reflect off of the damaged areas. At or near the threshold of damage these changes are very small and difficult to detect, in particular the problem of detecting phase changes is notoriously difficult.

One method for detecting these changes is the so called Schlieren technique that was originally developed and used in Germany for the detection of inhomogeneous regions in optical glass which are often in the form of streaks (The German for streaks is schlieren). A good general introduction to Schlieren techniques can be found in the book by Holder and North¹⁷, whilst the review paper by Davies¹⁸ has a substantial bibliography containing references to nearly all the various types of Schlieren spanning a wide range of applications. Essentially, Schlieren involves altering the spatial frequency content of an image so that phase changes and refractive index gradients are made visible. This is achieved by using suitable apertures in the focal plane of the imaging lens to alter the intensity distribution (spatial frequency content) of this plane and hence the image and as such, Schlieren can be viewed as an optical processing system. A good introduction to spatial filtering and optical processing is to be found in the book by Yu¹⁹.

When a uniform surface is imaged by a plane wave, the light in the focal plane of the imaging lens is all contained in the central maxima (zeroth order spatial frequency) of the pattern. The presence of non uniform areas will result in some of the light being shifted out of the central maxima to outer regions of the focal plane (i.e. to higher orders/spatial frequencies). Possible non uniformities are such things as defects, cracks, scratches, scattering sites, phase changes due to having coating layers removed or, areas of anomalous absorption which when heated by the incident laser beam cause a refractive index gradient to appear due to the temperature dependance of the refractive index of the surrounding air. In general the amount of light shifted out to higher orders is a small fraction

of the amount in the centre maxima and therefore these non uniformities when viewed appear as a small ripple on a large dc background. If the zeroth order is attenuated then the higher orders will become a larger fraction of the total observed signal and it will therefore be much easier to detect these non uniformities.

Standard Schlieren (see fig 2.11) involves using a knife edge inserted in the focal plane so that half the zeroth order and all the spatial frequencies on one side of the pattern are blocked i.e. removed from the image. This then means that if the object is a uniform surface the image has had its intensity reduced by a factor of 2. Should the surface now develop areas of non uniformity, light will be shifted out of the zeroth order into higher orders and thus these sites will show up in the image as areas of light or dark depending upon whether the shift was to orders that are passed or blocked by the knife edge. Only the areas of change (e.g. where there are refractive index gradients, phase gradients or regions separating damaged from undamaged, or areas of anomalous absorption) will show up since areas of uniformity will not shift light into the higher orders.

This form of Schlieren is the most sensitive in that it can detect both increases and decreases of light whereas, in the type of Schlieren that totally blocks the zeroth order, the contrast of the light passing the stop is maximised but, the system can then only detect areas that shift light into orders that the knife edge passes i.e. it cannot detect decreases in light coming from an area.

2.6.3 Detection using the Framestore and Schlieren Imaging

The purchase of a video framestore and related video equipment e.g. cameras, monitors etc allowed for the first time single picture capture of damage related events and their subsequent image processing.

Modifications, based upon using 74121 TTL chips to provide delays between input and output pulses were made to the video trigger box to enable it to provide 3 pulses. One to trigger the laser and 2 others 40 ms apart to trigger the framestore. The framestore was initially operated in the 512 by 256 pixel mode giving 2 stores but, this was later changed to the 256 by 256 mode as this gave 4 stores. The framestore set to store 1, was triggered first and then, 20 ms later the excimer fired and a further 20 ms after that the framestore set to store 2 was triggered again (see fig 2.12). This would result in the image in store 1 (image 1) being the "before" image i.e. before the excimer pulse and that in store 2 (image 2) the "after" image. The difference in the images if any would thus represent the induced damage.

The delay of 40 ms between images was chosen for 3 reasons...

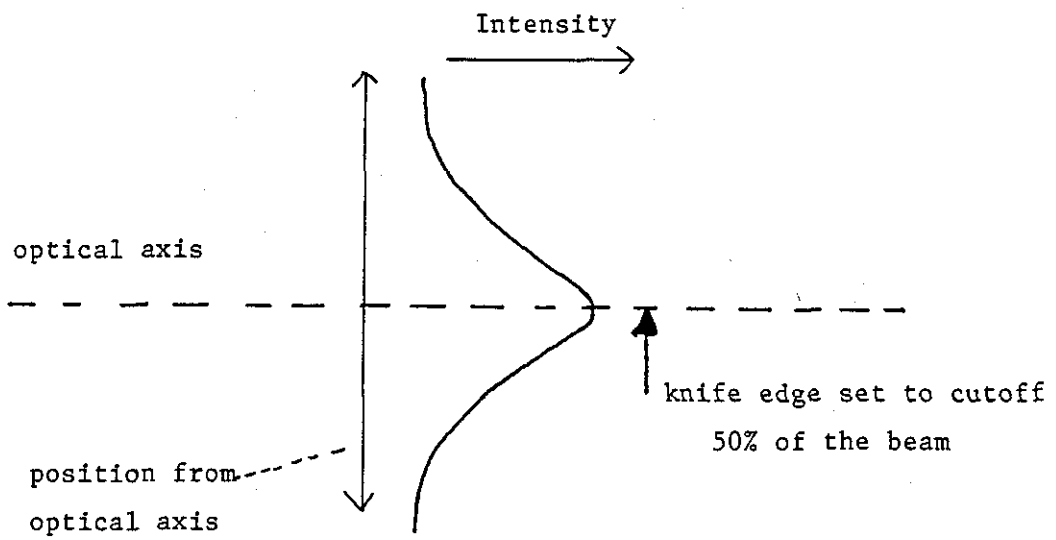
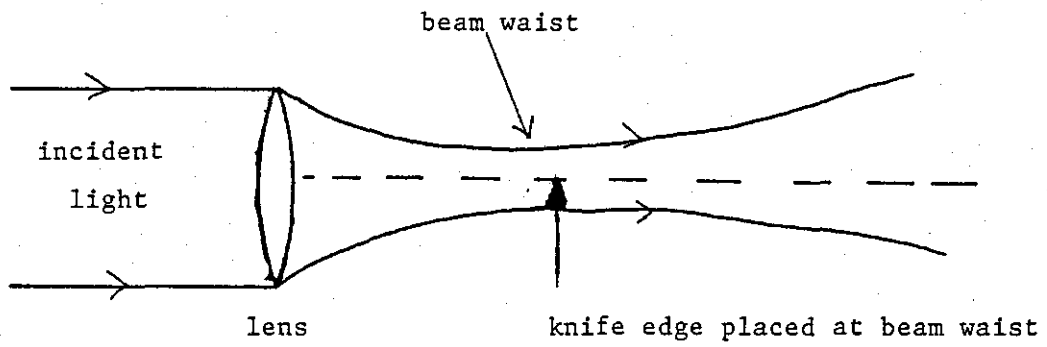


Fig 2.11
Schematic of Schlieren technique

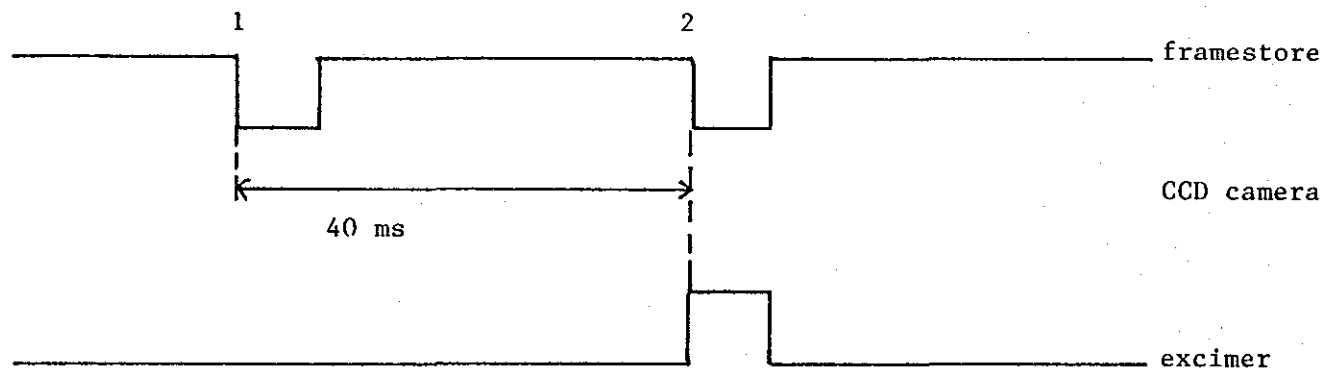
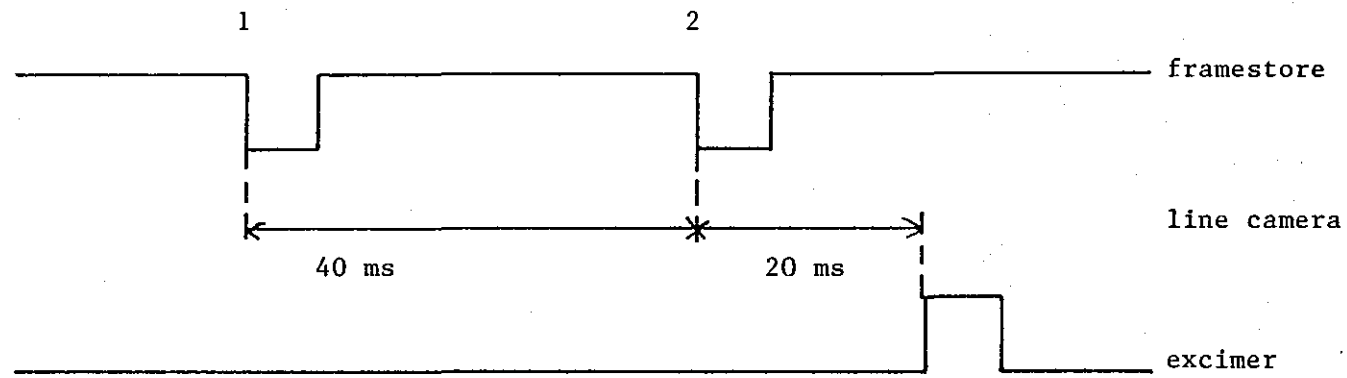


Fig 2.12 Timing sequence for damage experiments

1 = "before image"

2 = "after image"

1) images taken multiples of 40 ms apart correspond to the same video field and so eliminate the half pixel shift associated with CCD devices

2) transient effects would have decayed away

3) 40 ms rather than other multiples of 40 ms e.g. 80 ms etc was chosen as it would reduce the chance of room vibrations and variations in illuminating intensity causing differences in the images when in fact the images should be identical as no damage had occurred.

The switching of the framestore between stores 1 and 2 was done by software routines as it had to be performed faster than was possible manually.

The surface of the sample under test was illuminated by a 5 mW HeNe cw laser and imaged by a video camera. In the focal plane of the imaging lens a knife edge mounted on a micrometer driven positioner was set up so as to remove half the zeroth order of the pattern and thus produce a Schlieren imaging system.

Once everything was setup the two images were captured and stored on floppy disc using the same routines as were originally written for the beam profiling system pending processing.

2.6.4 Damage Processing routines

Since, as a result of damage it is quite possible to have areas darker than they were before damage as well as areas that are lighter, the subtraction routine written for beam profiling which assumed that image 2 was always brighter than image 1 was unusable. A separate subtraction routine was written which allowed the user to subtract image 1 from 2 or vice versa and worked in the following way, at point x, y let image 1 have level g_a and image 2 level g_b then the routine would evaluate the processed level g_p for point x, y as

$$g_p = 31 \pm (g_b - g_a) / 2$$

and write this back to store 1 point x, y . Processing in this way enables positive and negative changes i.e. increases and decreases of light to be displayed whilst ensuring that the data always remains within the 0 to 63 level range of the framestore.

Note that since this was written in assembly level and therefore the numbers are in binary form, the divide by 2 was performed by a 1 place logical shift right of the binary representation of the data. This means that although the absolute value of the maximum value of $g_b - g_a$ is 63 that of

$(g_b - g_a)$ is 31 as 0.5 part is ignored in the shifting operation. Hence g_p varies from 0 to 62 depending upon the data involved. Although this divide by 2 means that the minimum change that can be detected is 2 levels this gives essentially no loss of resolution in practice as the noise level is of the order of 1-2 greylevels.

An assembly language contrast enhancement routine was written that increased (enhanced) the difference of the processed level g_p from 31. This was done by determining the maximum deviation, d_{max} in an image from level 31 (either above or below). Then if the level at point x, y in an image is g , then written back to point x, y in the same store is the enhanced value g_e given by

$$g_e = 31 * (1 + (g - 31) / g_{max})$$

This routine thus makes full use of the available range of the framestore and since the subtraction was done first any deviation from level 31 represents data and not noise.

Figs 2.13a-d show the results of processing a test image. Figs 2.13a and b represent the original images with a slight difference between them. Fig 2.13c is the result of subtracting the images whilst fig 2.13d shows the effect of the enhancement routine on the subtracted image.

2.6.5 Imaging and Alignment of the Sample

Typically the surface of a sample was imaged at approximately 300 times magnification as viewed on the monitor and given that the camera provides times 32 an optical magnification of around times 10 was used. This was achieved by using 2 simple single element lenses, the first one giving times 2 magnification with the knife edge in its focal plane and the second times 5. Two lenses were used because of problems of 1) available working space and 2) aberrations associated with using single element short focal length lenses. With the apertures involved resolution was limited to $\approx 4 \mu m$.

Depending upon whether the sample was a reflector or transmitter of the UV beam it was placed a convenient distance e.g. 1 to 3 mm respectively inside or outside of the focus. The reflecting type samples e.g. metals and high reflectance (HR) dielectric stacks were placed inside the focus so that there is no possibility of the beam being reflected from the sample and subsequently focussed by the lens onto the output windows of the laser which can, and indeed did on one occasion cause damage to the output windows.

The positioning of the sample was achieved by using a micrometer depth gauge to measure the distance from a convenient fixed point. For alignment purposes the sample was then irradiated with a very high fluence excimer

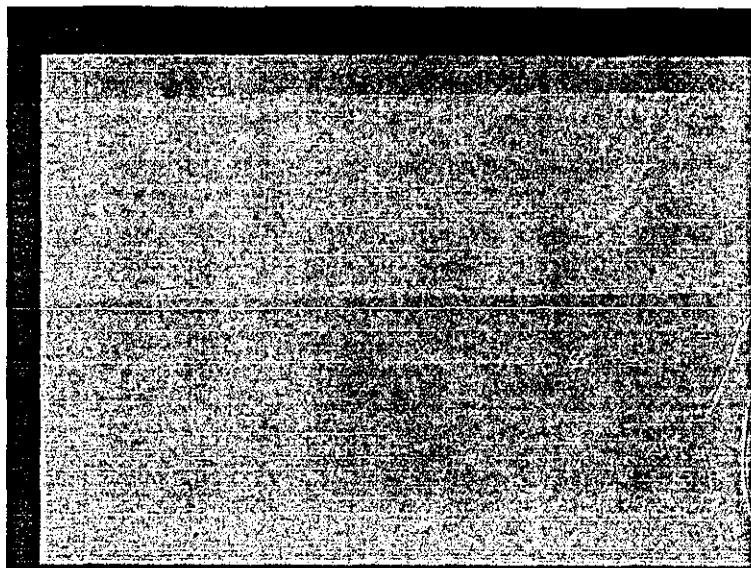
Fig 2.13
Effect of contrast
enhancement



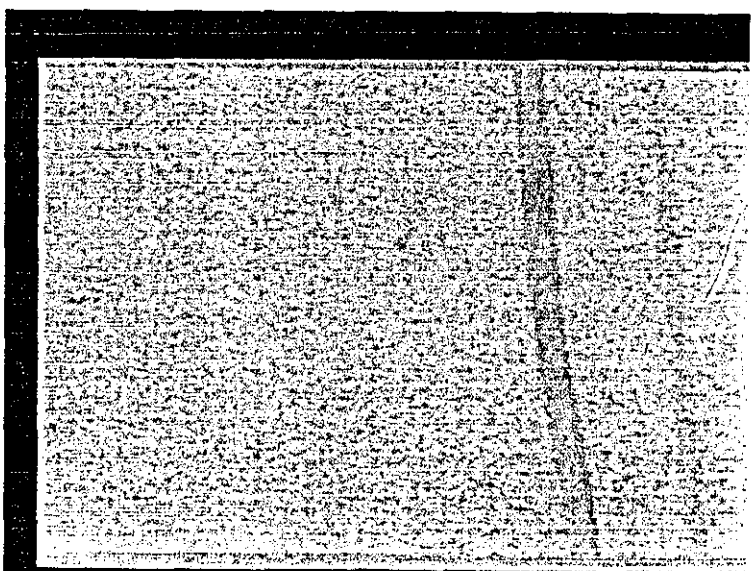
A



B



C



D

pulse, high enough to cause large scale damage e.g. a damage site of around 1mm in size. The illuminating HeNe laser was then centred on this damage site and the video system set to detect this light and so image the surface. The knife edge was then appropriately set.

2.6.6 Damage Testing

The criteria for determining the damage threshold of a sample adopted here at Loughborough is the fluence at which there is a 0% chance of damage occurring i.e. no site tested at the quoted fluence damaged.

A test site was chosen and irradiated with the required number of pulses e.g. 1 for single shot and it was noted if damage occurred. A new site was then chosen and the process repeated. Initially the laser energy was set high and gradually reduced to the point where some test areas did not damage. During this part of the test, damage detection was carried out by viewing the surface in real time via the video system. Only when the fluence was such that damage occurred at some but not all sites was the framestore used as it greatly increases the time required to evaluate the result on a given site but, it is a far more accurate and reliable method especially at or near the damage threshold. The excimer fluence was varied until none of the sites (typically 10) tested at a given fluence damaged.

The use of a beam of spot size much greater than the defect spacing together with multiple sites reduces the defect problem since under these conditions a defect is almost certain to be irradiated at or very close to the peak fluence in the beam, thus justifying the use of the peak beam fluence as the damage threshold rather than one averaged over the $1/e^2$ area.

In performing the tests a distinction was made between changes arising from the "blow off" of dust etc from the surface and damage that degraded the optical properties. Real time observation of the surface via the video system enabled the difference between "blow off" and damage to be determined.

2.6.7 Laser Pre-cleaning

Depending upon the type of sample under test, the variation in damage threshold across the surface could be quite small (e.g. 20%) or in some cases much more than an order of magnitude. In nearly all of the samples with a large variation, this variation was drastically reduced by irradiating the surface with 2-6 very low non damaging pulses. It is the author's contention that these pulses remove surface contaminants such as polishing compound organic molecules, dust etc that were either left on the surface after production or picked up during transit and handling. These pulses thus produce a more uniform surface where the surface properties are primarily

those of the coating and or substrate rather than being mainly due to contaminants.

In practice if a sample was found to have a large threshold variation, each site was pre-cleaned before irradiation by the damaging pulse. This resulted in not only a more reproducible threshold but a significantly higher one due to the elimination of easily damaged contaminants.

2.6.8 Results

Fig 2.14a-f show a typical damage sequence. They are for a dielectric antireflection (AR) coating on a fused silica substrate. Fig 2.14a is the initial surface after 4 low energy cleaning shots. Fig 2.14b is the surface after 1 damaging pulse and fig 2.14c after a second damaging pulse. Fig 2.14d-f are the processed (subtracted) images of respectively 2.14a and b (2.14d), a and c (2.14e) and b and c (2.14f). The difference between 2.14a and b is fairly obvious from the actual pictures but that between b and c is very hard to tell directly and reference to 2.14f shows the advantages that image processing offers in the detection of small changes i.e. damage between images.

2.6.9 Comments

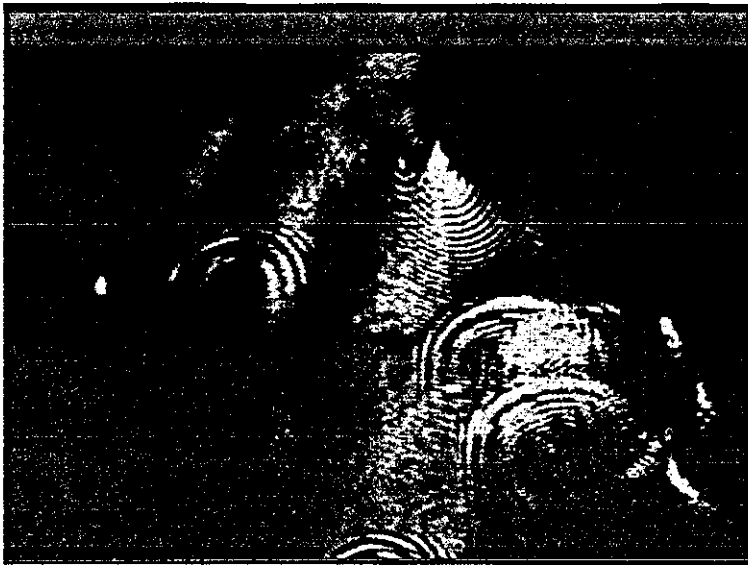
Since this technique is based on imaging it allows the user to determine where a sample has damaged and allows for the possibility of correlating damage to visible defects such as scattering sites and as is shown in chapter 4 an extension of these ideas can be used to correlate damage to non visible thermal defects.

2.7 Conclusions

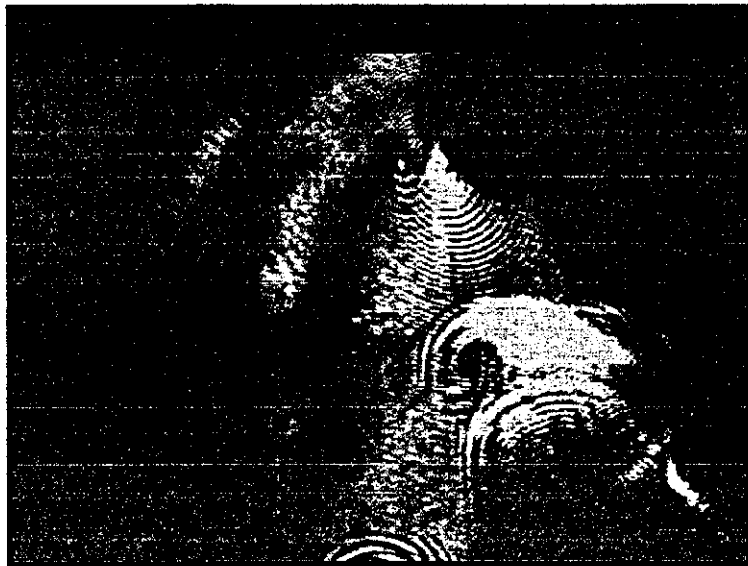
A new technique for beam profiling that enables complete spatial characterization of the beam to be obtained in a single laser pulse has been developed. The technique developed for the detection of LID is not only very sensitive to small changes but because it does not require the user to make a real time judgement is very reliable and allows the spatial location of these changes to be determined.

Fig 2.14

Damage Sequence



A



B

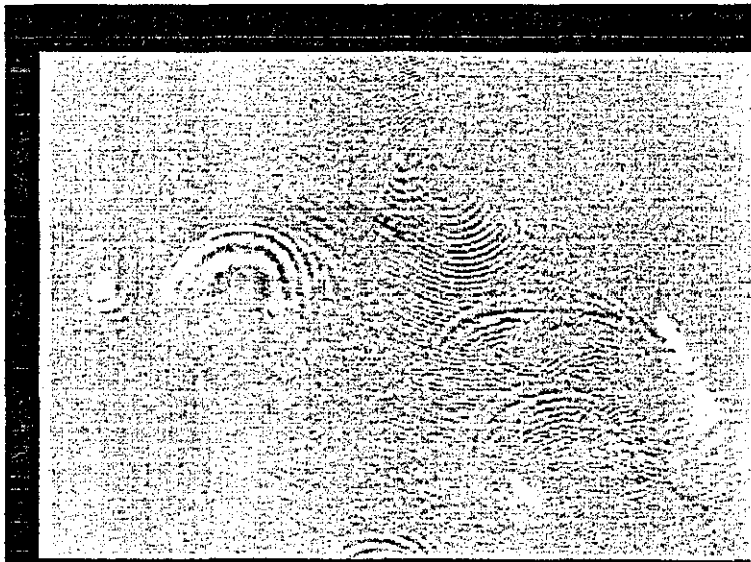


C

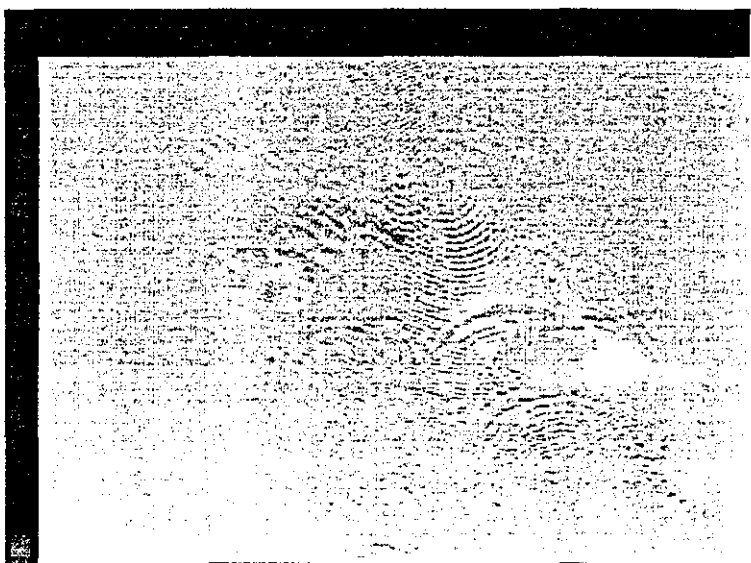
Fig 2.14



D



E



F

References

- 1) I.M. Winer Appl. Opt 5 1437 (1966)
- 2) D.R. MacQuigg Appl. Opt 16 2020 (1977)
- 3) D. Milam Appl. Opt 20 169 (1981)
- 4) F. Hurter and V.C. Driffield J. Soc Chem. Ind 9 445 (1890)
- 5) ref 4-12 in the review paper by D. Hull and A.F. Stewart Lasers and Applications November 1985 p71
- 6) C.B. Roundy Proc SPIE 499 95 (1984)
- 7) S.E. Clark et al published in proceedings of the 1985 conference on Laser Induced Damage in Optical Materials, Boulder, Colorado, USA in press.
- 8) D. Milam Proc. SPIE 140 52 (1978)
- 9) S.C. Seitel and M.T. Babb published in proceedings of the 1986 conference on Laser Induced Damage in Optical Materials, Boulder, Colorado, USA in press.
- 10) A. Rosencwaig et al Appl. Opt. 19 4133 (1980)
- 11) P.A. Temple NBS Spec. Pub. 568 p333
- 12) N. Thomasl et al NBS Spec. Pub. 620 p159
- 13) I. Laidler, final year project, Dept. of Physics, Loughborough University, 1983.
- 14) A.J. Maloney, final year project, Dept. of Physics, Loughborough University, 1985.
- 15) B.A. Omar and M.J. Shaw, Rutherford Appleton Laboratory internal report 1985 unpublished.
- 16) B.A. Omar, final year project, Dept. of Applied Physics and Physical Electronics, Portsmouth Polytechnic, 1986.
- 17) "Schlieren Methods" by D.W. Holder and R.J. North, published by the National Physical Laboratory, ,1963.
- 18) T.P. Davies Optics and Laser Tech. February 1981 p37
- 19) "Optical Information Processing" by F.T.S. Yu, published by Wiley & sons 1983.

Chapter 3

Laser Induced Ripple Structures

3.1 Introduction

Over the last 25 years or so researchers working with high power laser beams have noted many different types of damage and damage morphology depending upon the material and laser involved however, there is one type of widely reported surface damage known as Laser Induced Ripple Structures (LIRS) which, appears to be almost independent of the material and type of laser used. These structures appear as periodic undulations of the surface, akin to the ripples that appear on the surface of a river or pond if a stone is dropped into it.

The structures observed can be divided into 3 types dependent upon the mechanism of their formation as detailed below....

a) Laser Induced Thermomechanical Structures (LITS) formed by the buckling of an interface as a result of laser heating.

b) Laser Induced Capillary Waves (LICW) formed by the generation and subsequent "freezing in" of capillary (surface tension) waves on a fully molten surface.

c) Laser Induced Periodic Surface Structures (LIPSS) formed as a result of the interference between the incident light and a induced surface electric field.

The aim of this work was initially to investigate the possibility of generating UV LIPSS and if this was possible, to attempt a detailed investigation (experimental and theoretical) of their properties. Following early work by Wiseall and Emmony¹ samples of Al overcoated with silicon oxide (in the form of SiO_x where x is ≈ 2) were initially used. These samples yielded only LITS and were soon replaced by samples of Al and Ge with no overcoats which yielded LICW and the desired LIPSS.

The term LIRS will be used as a general term when reference to all 3 types of ripple structures is being made. The rest of this chapter will be divided into 3 separate sections dealing in turn with each of the above types of LIRS.

Note that all fluences quoted in this chapter (except the "normalised

formation fluence" in 3.4.6d) are based on the area of the beam when normally incident on a sample.

3.2 LITS

3.2.1 Introduction and Review

LITS were initially reported in connection with laser annealing of semiconductors, in particular silicon²⁻⁴. The annealing process involves the rapid heating and cooling of semiconductor surfaces which allows such things as recrystallization of the semiconductor, redistribution of dopants and most importantly provides a mechanism to removing implantation damage. A comprehensive introduction to laser annealing of semiconductors can be found in the book edited by Poate and Mayer².

The great majority of silicon based devices require their surface to be covered with a thin insulating layer. Silicon dioxide is used for this layer as it has exceptional masking, insulating and passivating capabilities. It is usually desirable to have the silicon dioxide layer on the silicon for most of the processing operations.

Laser annealing of silicon covered with its oxide results in ripple patterns (as in fig 3.1a) being formed on the surface. The spacing of these ripples (typically ≈ 2 to $10 \mu\text{m}$) as shown in the work by Stephen³ depends largely on the thickness of the oxide coating, increasing with thickness of oxide and is unrelated to the wavelength of the annealing laser.

Hill and Godfrey⁴ proposed the generally accepted explanation of this effect in terms of induced stress caused by the different expansion of each material as a result of the laser heating. In this model the underlying silicon melts first as it is heated directly by the laser beam whereas the oxide is heated only by conduction from the silicon. This leaves the heated oxide layer (within the beam target interaction area) floating on a liquid surface but, it is still constrained at the edges where the silicon has not melted. Since the heated area will attempt to expand laterally but cannot as it is pinned by the areas that have not melted it will, in order to relieve the stresses created buckle upwards as it is not constrained (except by its own strength) in this direction, thereby suffering permanent non elastic deformation.

The thicker the oxide layer the stiffer it is and therefore for the same induced stress i.e. laser heating, the less the surface will buckle and hence the spacing of the ripples formed will increase.

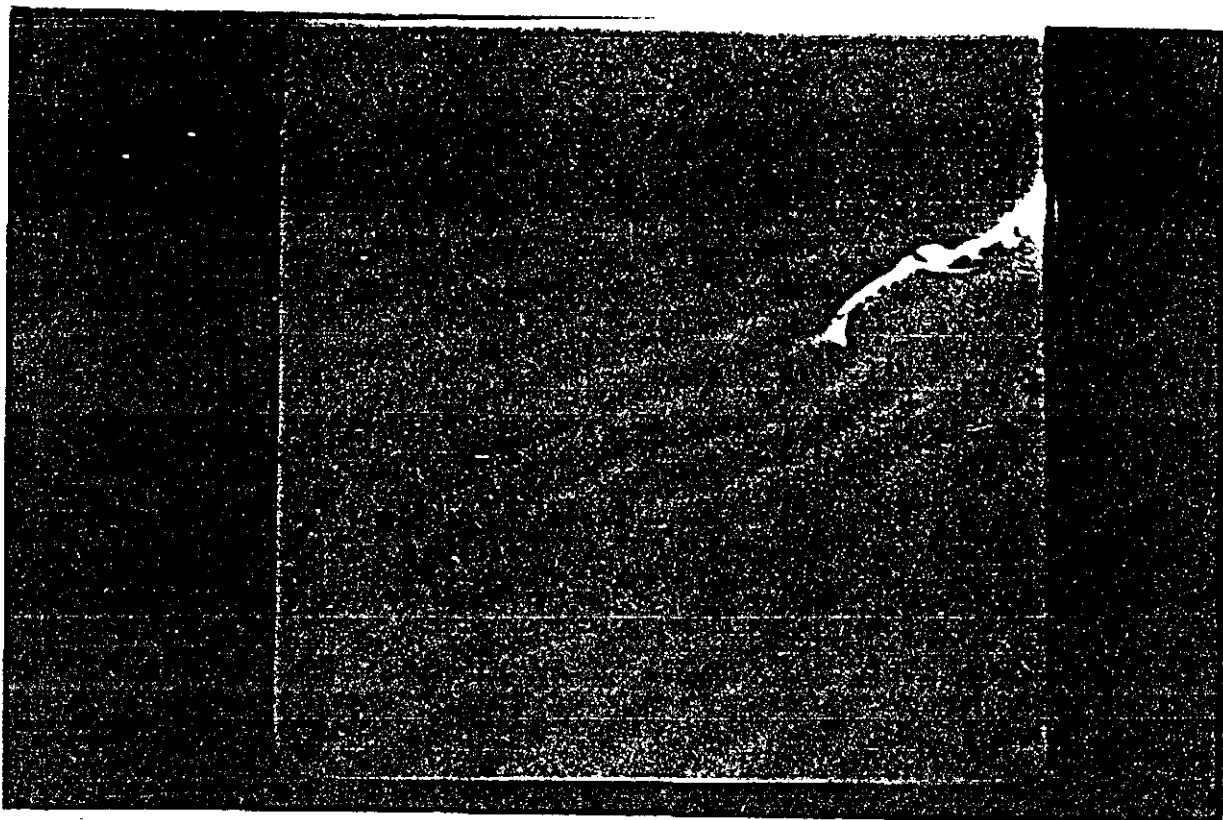


Fig 3.1

SEM pictures of ripples on overcoated Al deposited on glass

Above a) magnification X5,000

Overleaf b) magnification 5,500 The top part of the picture shows ripples in the underlying Al where the overcoating has been "blown away". The lower part still has the overcoating layer and the continuation of the ripples into this area can with careful examination be seen.



3.2.2 Experimental Configuration

The same experimental arrangement was used for all LIRS work and is shown in fig 3.2

As will be discussed in 3.4.2 previous work had shown that the polarization of the light used has a significant effect on the generated structures and thus some form of UV polarizer was required. Since no purpose built polarizer was available, the light reflected from a single beamsplitter orientated at Brewster's angle was used.

Bs2 (fig 3.2) was set to Brewster's angle for fused silica in the UV ($\approx 56^\circ$). By using 2 photodiodes the fraction of light reflected off Bs2 was measured and found to be 15%, in excellent agreement with that predicted by using the Fresnel reflection equations for an unpolarized incident beam. This showed as expected that the laser output is essentially unpolarized. As only metal and semiconductor samples were used, having to work with at most only 15% of the beam was not a significant problem.

The reflected light at Brewster's angle is S polarized which given the orientation of the components used results in the E field of the light being in the vertical plane. This then allowed the variation of LIRS with angle to be determined for S polarized light by simply rotating the sample in a horizontal plane but, in order to repeat this for P polarized light a special mount had to be made which allowed rotation in the vertical plane as well. This worked because by aligning the sample in the horizontal plane to be normal to the beam, and then rotating it in the vertical plane, the plane of incidence is rotated by 90° which, effectively "converts" the S polarized light to P.

Since the experiments involved rotation of the sample, it's surface was not imaged during the experiments, instead it was simply observed directly by eye. This was adequate for the work herein but for planned future work a proper imaging system which rotates with the sample will be used.

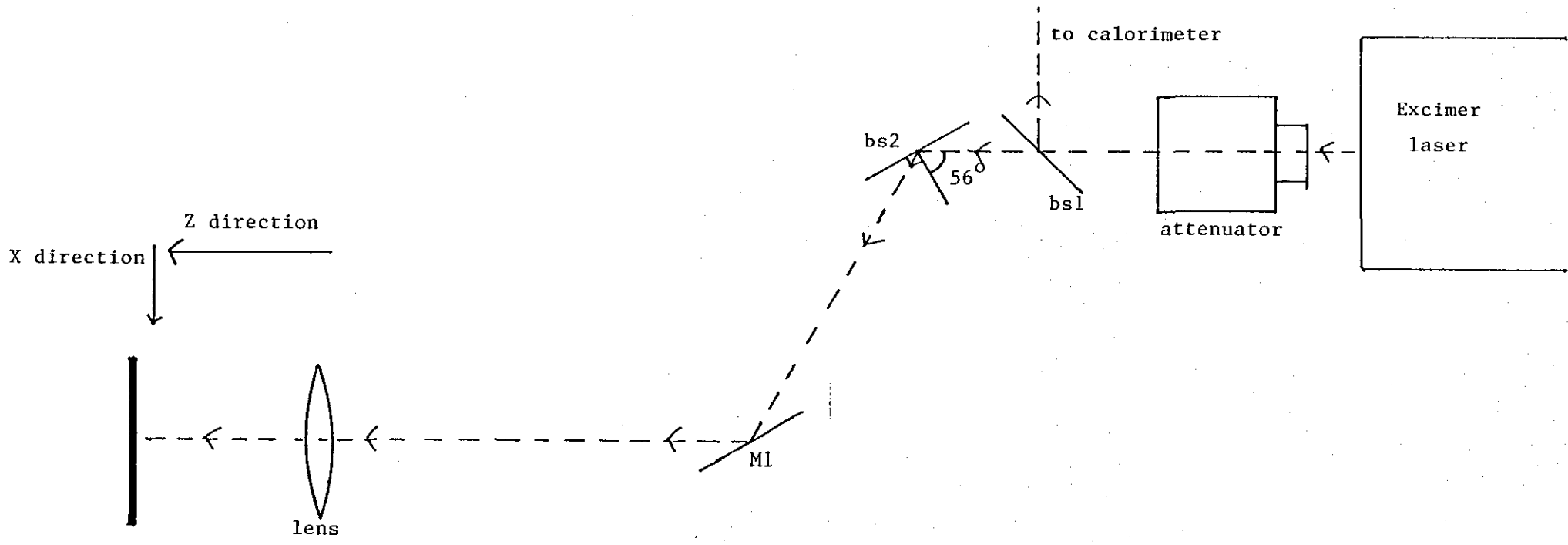
The samples, being high reflectors were placed inside the focus of the lens by some 3-4 mm. Beam profiling measurements showed that for a distance of approximately 4 mm either side of the focus the beam profile and peak fluence were very nearly constant.

3.2.3 Procedures for Analysis of Results

The same analysis procedures were used for all LIRS work and consisted of optical (OM), scanning electron (SEM) and transmission (TEM) microscopy.

The OM work which was based on a Riechart MeF3 with a maximum useful magnification of 1900 was mostly performed by using an interference contrast

Fig 3.2



76

target plane
rotate in XZ plane for S pol.
rotate in YZ plane for P pol.

technique although dark and bright field illuminations were occasionally used.

The SEM work was performed on a Cambridge Steroscan device which in practice was limited to magnifications of $\approx X10,000$ if picture quality was to be maintained. The samples were inserted directly into the SEM and only the overcoated samples (because of their lack of surface conductivity) had problems (albeit small ones) with surface charging due to the electron beam.

The TEM work was performed using a JOEL TEM 1000CVI which was easily capable of $X100,000$ or more. In order to actually perform the TEM work replicas of the surface using a standard acetate/carbon coating technique were made. Since the replication process involves carbon coating an acetate replica of the surface this, at high magnification results in a roughness (i.e. grain structure) on the pictures which is not related at all to the original surface.

3.2.4 Samples Tested

The samples tested were provided by Comar Instruments Ltd. of Cambridge and consisted of an optically thick layer of Aluminium on a glass substrate flat to $\lambda/5$ (at $\lambda \approx 550\text{nm}$) either with or without a half wave thick overcoat layer of silicon oxide in the form SiO_x where x is ≈ 2 .

3.2.5 Experimental Procedure

Once suitably positioned relative to the focal plane of the lens, areas on the samples were irradiated by either single or multiple (2 -40) excimer pulses of varying fluence. The fluence was varied from being so low that there was no visible damage through to fluences so high that they produced a large plasma immediately above the surface of the sample which was accompanied by an audible "cracking" sound. The angle of incidence and the polarization of the incident light were varied by suitably rotating the sample.

3.2.6 Results

The ripples found on the overcoated samples showed no dependence of the spacing on either the angle of incidence or the polarization of the incident light. No ripples were detected on the samples of bare Al i.e. those with no overcoating.

At very low fluences no ripples were ~~observed~~ ^{observed} but on sites exposed to higher fluences e.g. above 0.2 Jcm^{-2} ripples were readily detectable. Fig 3.1a is typical of the ripples produced on the overcoated samples and shows a spacing of $\approx 2\mu\text{m}$.

In this picture the oxide overcoat has been left covering the entire metal surface, however in fig 3.1b observed on a site exposed to ≈ 10 pulses of fluence 0.4 Jcm^{-2} it is clearly evident that there are areas of rippling that have no overcoat on them and indeed the continuation of the ripples into overcoated regions is apparent.

The ripple spacing was not constant and indeed appeared to vary randomly even within a single damage site with values of ≈ 2 to $4 \mu\text{m}$ being observed.

Fig 3.3 is typical of the surface of non overcoated samples and was observed on a site exposed to the excimer beam in the same way as 3.1b. It is obvious that there is no evidence of the rippling that occurs on overcoated samples.

3.2.7 Analysis of Results

The absence of ripples on non overcoated samples which clearly indicates that the rippling is related to heating of the metal oxide interface, together with the lack of dependence of the ripple spacing on the angle of incidence or polarization of the incident light is in the author's view sufficient to label the ripples as LITS. The labelling of these ripples as LITS can be justified since the conditions for LIPSS (which are electric field related) are essentially satisfied for both types of sample whilst those for LICW (which require a surface to be fully molten to a suitable depth) are more likely to be satisfied on the non overcoated samples (as the surface is not constrained by the overcoat) on which no ripples are found.

3.2.8 Conclusions

This work shows that ideally annealing processes should be carried out without any form of overcoating layer and that if the process has to be performed with an overcoat layer than the heating should be minimal and performed extremely slowly to try and maintain equilibrium at the substrate overcoat interface.

The results of this section clearly show that overcoated samples are not suitable for LICW/LIPSS studies and that bare Al coated on glass is not suitable either. This latter conclusion whilst initially somewhat surprising is, as will be shown later, explainable in terms of the Al coating being too thin to support any form of LIRS.

The samples were then changed to solid polished blocks ($\approx 3-6 \text{ mm}$ thick) of Al and Ge upon which it proved possible, as will now be discussed, to produce LICW (particularly on Al) and LIPSS (particularly on Ge).

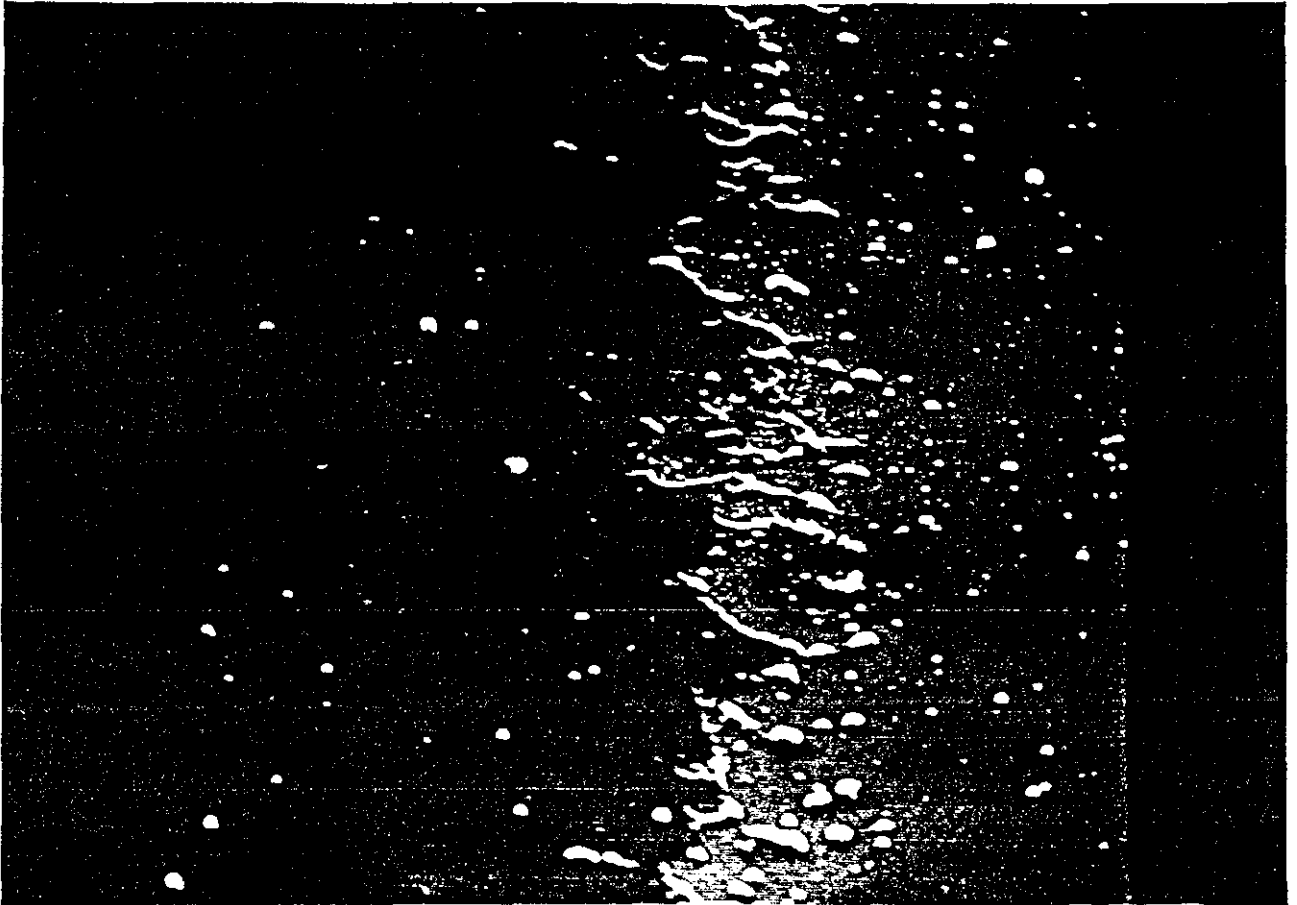


Fig 3.3

SEM picture of the substrate (left side), Al coating (right side) interface on the non overcoated Al on glass samples. Magnification 2,000. Note the absence of any ripples.

3.3 LICW

3.3.1 Introduction and Review

Capillary (surface tension) waves are a well known and documented phenomena on liquid surfaces, in particular on water surfaces. Excellent introductions to this effect can be found in the book by Landau and Lifshitz⁵ or Main⁶.

Essentially the waves occur as a result of the effect that surface tension forces have on the equilibrium shape of a fluid surface. From ref 6 it can be shown that if a fluid has depth h , density ρ and surface tension σ then the dispersion relation of its surface waves can be written as..

$$\omega^2 = (gk + \sigma k^3/\rho) \tanh(kh) \quad 1)$$

where g is the acceleration due to gravity, ω angular frequency of the wave and k it's wave number.

On the assumption that the liquid depth is much greater than the wavelength of the oscillation and that $k^2 \gg \rho g/\sigma$ then surface tension is the dominant restoring force and the dispersion relation can be written as..

$$\omega^2 = \sigma k^3/\rho \quad 2)$$

For metals and semiconductors in the liquid state σ is typically 10 times that of water i.e. $\approx 0.8 \text{ Nm}^{-1}$ and with $\rho \approx 5000 \text{ kgm}^{-3}$ requires the wavelength of the oscillation to be much less than 2.6 cm for surface tension to be the dominant restoring force which, since typically the wavelength of observed structures is less than $2\mu\text{m}$ is well satisfied in practice by the samples used in these experiments.

Viscous forces within the liquid will act to damp the oscillations with time t , causing the amplitude A of the capillary waves to decay as...

$$A \propto \exp(-\gamma t) \quad 3)$$

with

$$\gamma = (2\eta/\rho)k^2 = 2vk^2 \quad 4)$$

where η is the shear viscosity and v the kinematic viscosity of the liquid.

Both the phase and group velocities of capillary waves are proportional to the square root of k and are given by...

$$v_p = w/k = (\sigma k/\rho)^{1/2} \quad \text{and} \quad v_g = dw/dk = 1.5v_p \quad 5)$$

where v_p and v_g are respectively the phase and group velocities.

Capillary waves can be initiated on a liquid surface by various "driving forces" such as the impact of particles onto the surface, ablation of particles off the surface, localised anomalous heating which sets up stress waves in the sample or by the propagation of shock waves into the molten regions either up through the bulk of the material or as a result of a pressure wave in the air above caused by such things as the presence of a small volume of plasma.

The first report of LICW appears to be that by Rozniakowski⁷ where the pattern generated on an Al surface exposed to high laser fluences (in the infrared) was, with little detailed justification attributed to laser induced capillary waves.

A much more detailed analysis was performed by Keilmann⁸ where the dispersion of the waves induced (by a pulsed CO₂ laser) on the surface of liquid metals such as Hg was determined and shown to be that of capillary motion. To obtain the required information, principally^{at} the oscillation period and the decay time, the diffraction caused by interaction of the capillary waves with a cw laser beam incident within the area irradiated by the high energy pulsed laser was monitored as a function of time.

Young et al⁹ (hereafter known as III) performed a detailed analysis of the structures generated on Ge irradiated by intense 1.06 μm light and concluded that for fluences sufficient to cause uniform melting of the surface capillary waves were induced on the liquid surface. They further showed that if the decay of the capillary waves was not complete by the time that the surface resolidifies then a permanent record of the ripples is "frozen" into the surface forming what are termed LICW.

Assuming that the surface still has ripples on it at the time of resolidification then, exactly what is permanently recorded on the surface will depend upon the relative magnitudes of the oscillation period of the wave and the time it takes to solidify the surface by removing the latent heat of fusion once the surface has cooled to its melting temperature (known as the "freezing time"). It is the author's contention that to record well defined ripples requires the freezing time to be much less than the period of oscillation otherwise the ripple will have changed significantly during the freezing time and thus be recorded as a somewhat blurred out structure.

The mechanism of initiating these waves ("seeding") will depend upon the laser fluences being used. The fluences used in ref 8 and 9 were sufficiently

low that no plasma was produced above the surface and in these cases the seeding of the waves is attributed to scattering of the incident laser light by surface roughness. The fluences used by the author always produced a small volume of plasma above the sample's surface and in this case, it is the author's contention that in addition to the previous mechanism, the forces induced on the surface by the presence of the plasma provide a possible second seeding mechanism.

Given that these waves require a liquid surface and as this takes a finite time to produce then, the seeding of the waves can only occur for the time (seeding time) that the molten surface is subject to the driving force i.e. to the incident laser pulse and or the induced plasma. In order to relate accurately the frequency of the generated waves to the seeding time a Fourier transform would have to be performed however, a measure of the generated frequency f maybe obtained by simply taking $f \approx 1/(2t_i)$ where t_i is the seeding time.

Table 3.3 lists the wavelength ,decay time and velocity that capillary waves generated on Al with the stated interaction times would have. The material parameters for Al are taken from table 3.1.

3.3.2 Samples Tested

The same types of samples were used for LICW and LIPSS work and consisted of 50 mm diameter 6 mm thick discs of Al and 35mm diameter 3 mm thick discs of Ge. There was no overcoat layer on any of these samples.

The Al samples were of a commercially pure grade (99% Al) and polished by Struers Ltd of Birmingham to a low quality optical finish. On these samples it was possible to see the so called "orange peel" effect characteristic of grain boundaries.

The Ge samples were polycrystalline high purity CO₂ laser mirrors polished on both faces to a good optical finish. The surfaces were flat to around $\lambda/5$.

3.3.3 Experimental Procedure

Initially the samples were irradiated by single excimer pulses at normal incidence with the fluence range as discussed in 3.2.5. Subsequent experiments involved varying the number of excimer pulses and both the angle of incidence and polarization of the incident light by suitably rotating the sample.

Attempts were made to monitor the diffraction of a probe beam by the capillary waves by using a cw laser linked to a fast photomultiplier. Unfortunately the largest available cw laser was only a 10 mW HeNe and as a result these experiments failed due to insufficient light

being incident on the detector for the timescales (of order 100 ns) involved.

3.3.4 Comment

In the course of performing these experiments both LICW and LIPSS were observed, often existing in neighbouring regions of a damage area. The results in this section will concentrate exclusively on LICW, the results etc for LIPSS can be found in 3.4.

3.3.5 Results

The ripple spacing produced showed no dependence on the angle of incidence and polarization of the incident light. The fluence required to form ripples was dependent on the number of shots used, the angle of incidence and polarization of the incident light.

Single shot normal incidence type experiments showed that at fluences below $\approx 2 \text{ Jcm}^{-2}$ no ripples were produced although the sample would badly damage. Between 2 and 3 Jcm^{-2} the ripples were found to form near the centre of the damage site as in figs 3.4a-b. As the fluence was further increased, the ripples formed progressively further out from the centre in an annulus surrounding the badly damaged centre of the site (fig 3.6b).

Multiple shot experiments showed the same trend except that the formation fluence was significantly lower e.g. 2-4 times lower depending upon the number of shots used.

Experiments involving varying the angle of incidence and polarization of the incident light showed the same trends as those at normal incidence with the formation threshold being higher (up to a factor of 4) than the normal incidence case for the same number of laser pulses.

The uniformity and definition of the ripples improved with increasing numbers of shots. Fig 3.6a shows the pattern generated after 20 moderate fluence pulses where, the generated ripples are quite uniform and well defined with a spacing of around $1.4 \mu\text{m}$.

In general the spacing of the ripples as determined from OM work was found to be in the range ≈ 0.8 to $2 \mu\text{m}$, with the spacing decreasing towards the outer edge of the damage site.

SEM work was performed in an attempt to determine the lower limit on the size of the ripples. Figs 3.5a-b show typical SEM pictures of the induced ripples where, the ripple spacing is found to vary from a few hundred nanometres upwards. In areas that had ripples visible in the OM it was, in general found that smaller and smaller ripple patterns were found as the magnification was increased, there appeared to be no lower limit on the size of the ripple spacing.

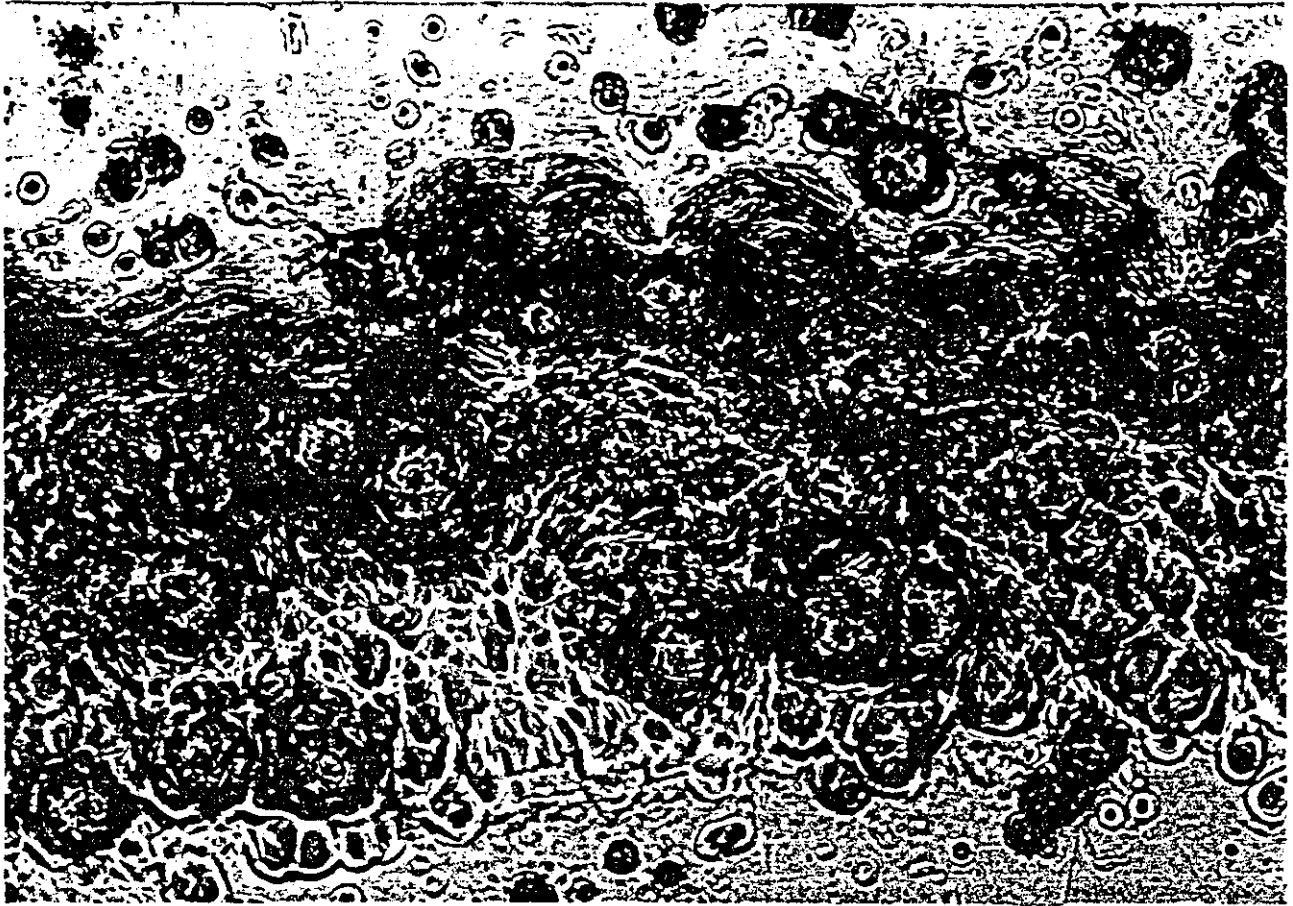


Fig 3.4

All samples irradiated by a single excimer pulse of fluence 2 Jcm^{-2} at normal incidence.

Above a) Magnification X500 Observe how the capillary waves have formed in the centre of the damage site,

Overleaf b) Magnification X3,000 The variation in wavelength of the capillary waves is evident and in particular the very fine structure at the edge of the damage area (top right hand corner).

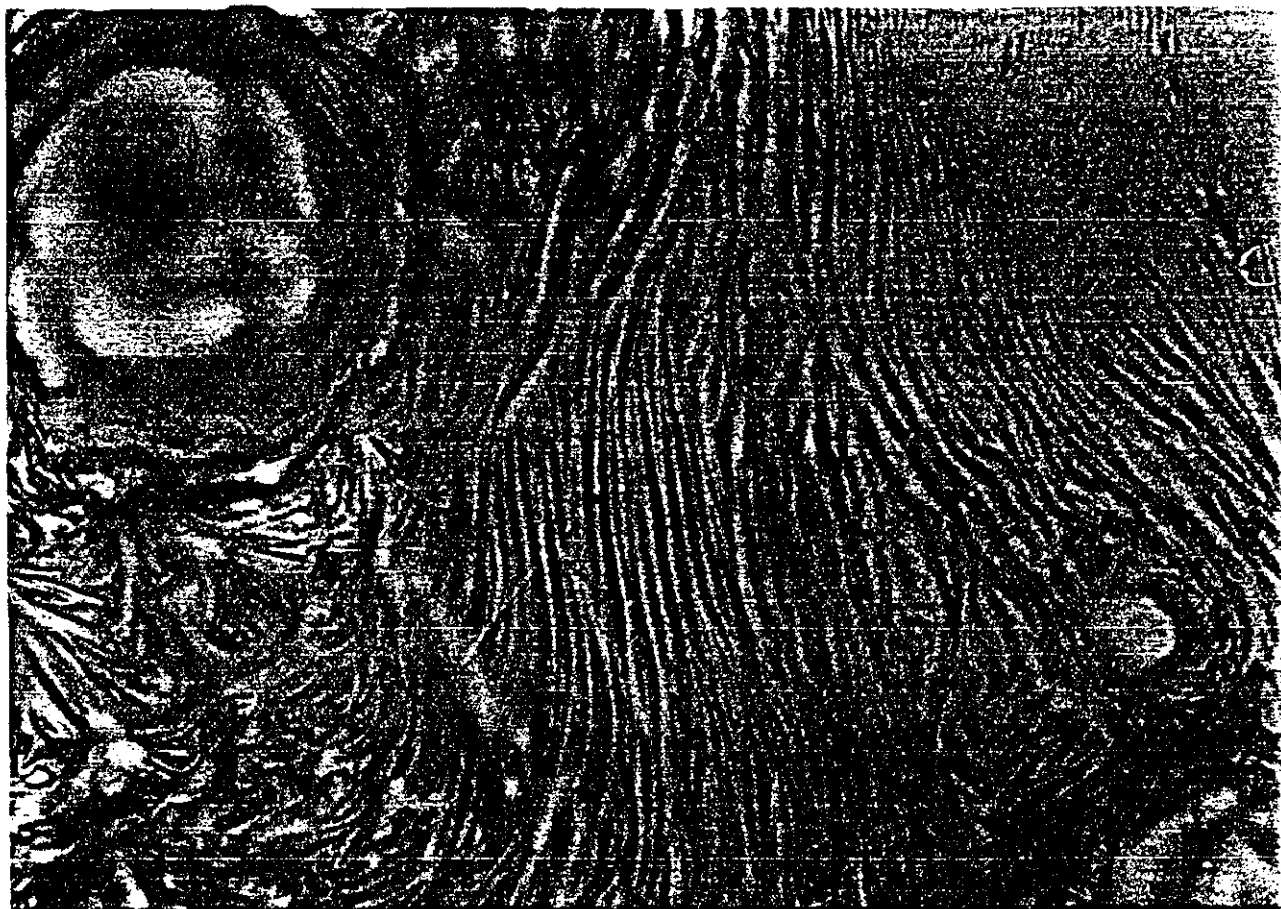




Fig 3.5

SEM pictures of capillary waves on Al samples.

Above a) Single excimer pulse incident normally magnification 5,000

Overleaf b) Multiple P polarised excimer pulses incident at 60° magnification 13,000. The capillary waves have been generated from the vortex at the top of the picture and subsequently frozen in. Note how the waves have begun to break over and also the variation in the period of the waves.



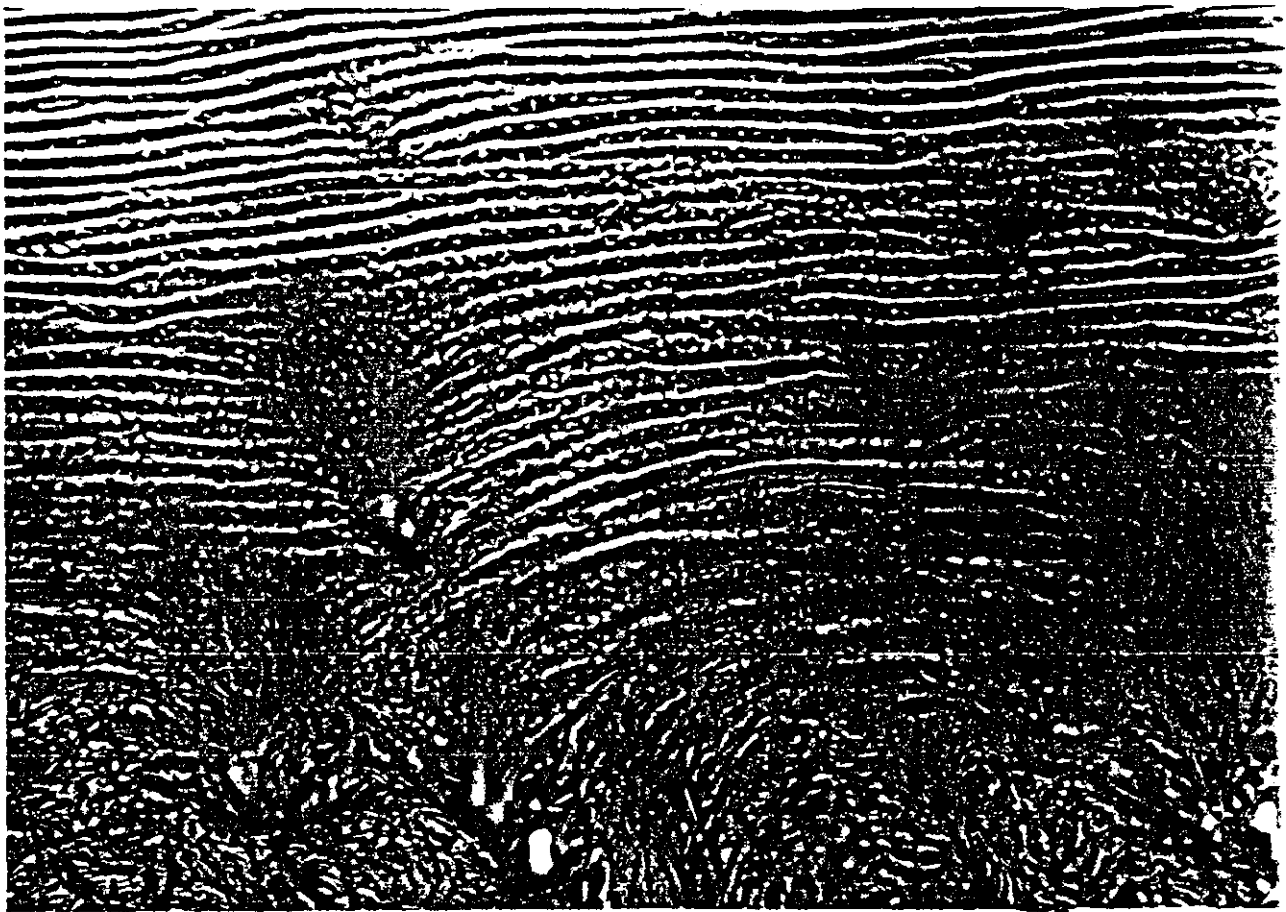
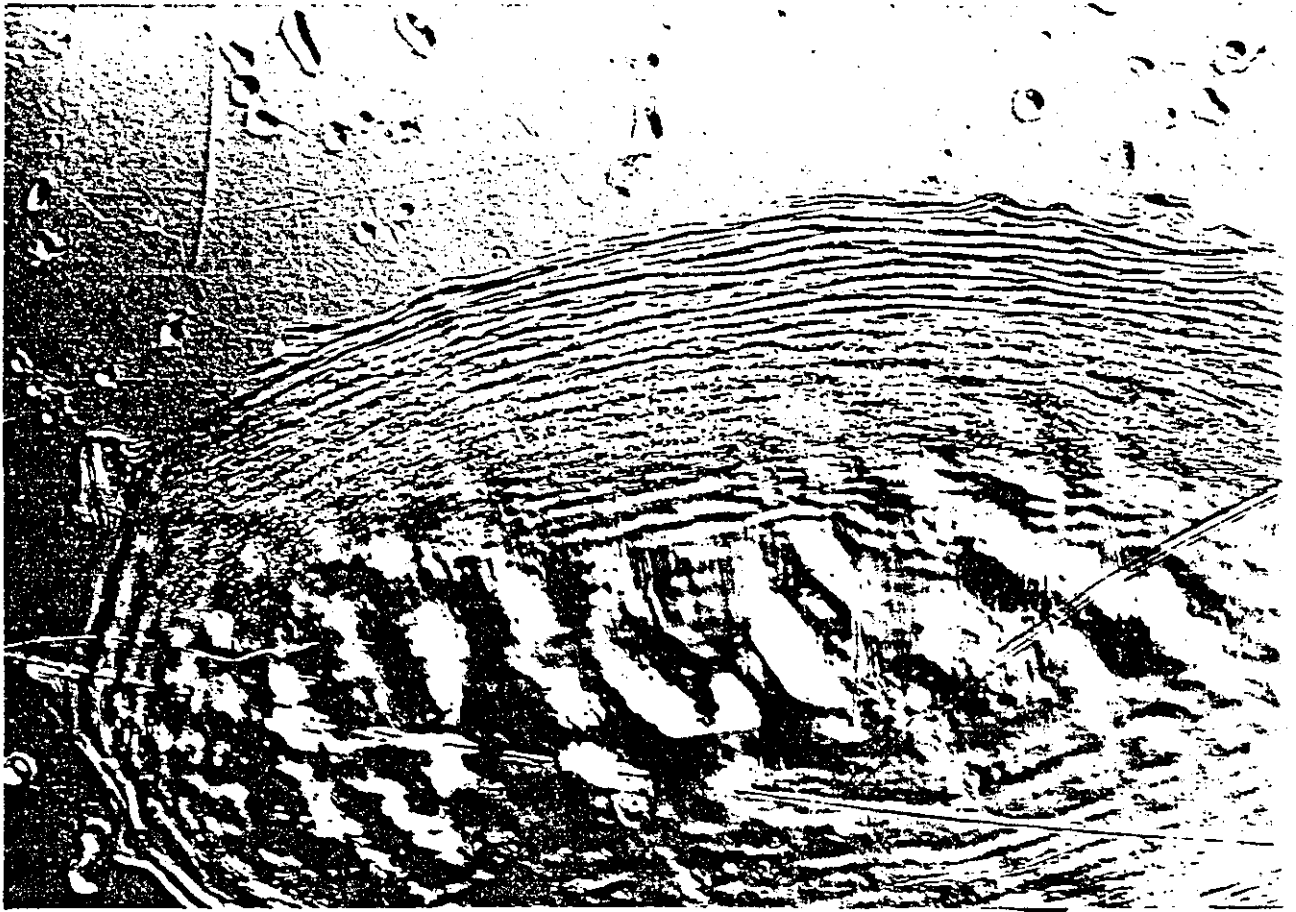


Fig 3.6

Multiple pulse induced capillary waves on Al by P polarised light.

Above a) 60° incidence magnification 1,500 Note the uniformity of the capillary waves and the increase in their spacing as the centre of the damage site (centre of picture) is approached.

Overleaf b) 75° incidence magnification X750 The capillary waves in this case have formed in an annulus around the centre of the damage site.



3.3.6 Laser Heating - Theory

Before these results can be analysed and as it is the author's contention that the ripples are LICW which require a molten surface some indication of the heating of the surface by the laser beam must be obtained.

The theory used herein and indeed by many workers is that developed by Ready¹⁰ which uses as a starting point the generalised equations developed by Carslaw and Jaeger¹¹.

In the model all material parameters such as thermal conductivity, specific heat capacity and density are assumed to be independent of temperature. This assumption is not too unreasonable for Al but, Ge being a semiconductor will have a temperature variation of its material parameters so large that it cannot be ignored and for this reason coupled with the fact that the LICW were far more prevalent on Al, the heat conduction calculations were performed for Al only.

The conduction of heat is assumed to occur only in the metal, any loss to the surrounding air is ignored. In view of the relative diffusivity of metals and air this is not an unreasonable approximation. Hence the conduction is assumed to occur in a semi infinite medium with the laser pulse incident on the $z = 0$ plane and the sample extending along the positive z axis.

The model further assumes that the absorption of the incident light occurs in a distance small compared to the distance that the heat will diffuse into the bulk of the sample in the duration of the laser pulse i.e. it is a surface heating model. This assumption can be justified for metals by comparison of the optical penetration depth and the thermal diffusion distance.

Consider a light wave propagating in a medium of refractive index n_t then, the amplitude of the electric field can be written in the form...

$$E \propto \exp(i(\omega t - n_t k x)) \quad 6)$$

where ω is the angular frequency, k the vacuum wavenumber, n_t the refractive index, t time and x distance. If the refractive index is complex as it is in metals and semiconductors then writing

$$n_t = n_r + i n_i \quad 7)$$

where n_r and n_i are respectively the real and imaginary parts of the refractive index yields by 6) an exponential decay in the E field given by..

$$E \propto \exp(-n_i k x) \quad 8)$$

Typically n_i is of order 3 for metals and semiconductors at 249nm hence

giving as the skin depth at 249nm $\approx 10\text{nm}$ and given that the intensity is $\propto E^2$, the intensity is down by a factor of more than 10^3 in a distance of 50nm.

In general solutions to the heat conduction equation for laser heating of a solid surface have a term(s) of the general form

$$T \propto \exp(-x^2/(4kt)) \quad 9)$$

where T is temperature and k the thermal diffusivity. The thermal diffusivity is defined as

$$k = K/(pc) \quad 10)$$

where K is the thermal conductivity, p the density and c the specific heat capacity.

Hence a measure of the thermal diffusion of heat with time is..

$$x = (4kt)^{1/2} \quad 11)$$

Typically for a metal k is $\approx 10^{-4}$ giving a diffusion distance for a 30 ns laser pulse of $\approx 3\mu\text{m}$ i.e. much greater than the optical skin depth and thus justifying the surface heating model.

The presence of a plasma above the sample which ideally should be included in any model would at first sight appear to render the situation intractable. Since the spatial extent of the plasma is closely related to the spatial profile of the beam, it is the author's contention that to first order the plasma will simply scale uniformly the heating effects due directly to the laser beam i.e. change the temperature of any point by the same fraction as any other point. Thus for the purposes of these calculations there is no direct allowance for the presence of plasma. However, indirectly, through the use of a coupling factor an attempt is made to allow for the presence of a plasma as now discussed.

Based upon the optical values in table 3.2 and using the Fresnel equations, at 249 nm the reflectivity of Al at normal incidence is 92%, thus 8% of the incident light is absorbed (coupled) as heat into the sample. Woodroffe et al¹² have shown that at the fluences used in this work the coupling of light as heat into Al is significantly greater than that theoretically predicted (i.e. 8%). This increase they attribute to a plasma surface interaction and at the fluences used in this work approximately 40% of the incident light was found to couple into Al as heat. Thus for these calculations a 40% conversion of incident light energy into heat (thermal energy) is used.

Carslaw and Jaegar¹¹ give the following expression for the temperature at a point r, z and time t in a semi infinite solid due to heating of the surface ($z = 0$ plane) by an instantaneous ring source of radius r' and total absorbed energy $q(r')$

$$T(r, z, t) = \frac{1}{4(\pi kt)^{3/2} pc} \int_0^\infty q(r') e^{-\left(\frac{r^2 + r'^2 + z^2}{4kt}\right)} I_0\left(\frac{rr'}{2kt}\right) 2\pi r' dr' \quad (12)$$

where r is the distance from the centre of the beam, z the depth below the Al surface and I_0 is a modified Bessel function of order zero.

Temporally for these calculations the beam is modelled as a square pulse of 25 ns ($=t_0$) duration. Reference to fig 2.1 will show that a 25 ns square pulse is not an unreasonable representation of the beam.

Again from Carslaw and Jaegar¹¹, if the instantaneous temperature at a point r, z, t is $g(r, z, t)$ and the heat is provided at a rate $a(t)pc$ per unit time from $t= 0$ to $t= t$ then the temperature at r, z, t is given by..

$$T(r, z, t) = \int_0^t \phi(t') g(r, z, t-t') dt' \quad (13)$$

For the purposes of these calculations the excimer beam is assumed to be a perfectly symmetric Gaussian beam of the form...

$$E(r') = E_0 e^{-\left(\frac{2r'^2}{w^2}\right)}$$

and temporally a rectangular pulse of duration t_0 . The radius of the assumed perfect Gaussian beam is chosen so that it's area at the $1/e^2$ points is the same as that of the excimer beam in reality. Thus for such a distribution $q(r')$ is given by

$$q(r') = \frac{2aE}{\pi w^2 t_0} e^{-\left(\frac{2r'^2}{w^2}\right)} \quad (15)$$

where E is the total pulse energy and a the coupling factor representing the fraction of incident energy that is coupled into the sample as heat and taken as 0.4 (i.e. 40%).

Substitution of 15) and 12) into 13) and performing the integration by use of suitable Bessel function identities as first performed by Ready¹⁰ yields the temperature at a point r, z at time t as.....

$$T(r,z,t) = \frac{2aE}{t_0 \pi p c \sqrt{\pi k}} \int_0^{\tau'} \frac{1}{(t-t')^{\frac{1}{2}}} \cdot \frac{1}{(8k(t-t')+w^2)} \times e^{-\left[\frac{z^2}{4k(t-t')} + \frac{2r^2}{(8k(t-t')+w^2)} \right]} dt' \quad (16)$$

where $\tau' = t_0$ $t \geq$ beam duration
 $= t$ $t <$ beam duration

A computer program was written to evaluate the above expression as a function of r, z, t and beam fluence. The parameters for Al and the excimer beam that were used in the calculations are listed in table 3.1. The values for Al are all taken from ref 13 except the viscosity which is from ref 14.

Figs 3.7a-c are typical graphs of the temperature distribution on Al calculated using the above theory, figs 3.7a and b are the radial distributions for different beam fluences with 3.7a being for 2 Jcm^{-2} and 3.7b for 3 Jcm^{-2} whilst 3.7c shows the temperature as a function of depth at the centre of the excimer beam for a fluence of 2 Jcm^{-2} . To avoid possible problems with ablation due to the high fluences being used the profiles were evaluated 100 nm below the surface. The melting point (660°C) and the vapourisation point ($\approx 2400^\circ \text{C}$) will have a significant bearing on the interpretations of these temperature profiles.

3.3.7 Analysis of Results

The calculated temperature profiles show that the Al is easily melted at the fluences used and, can stay molten for upto 200ns near the centre of the beam. The neglect of any phase changes (solid to liquid and liquid to vapour) in the calculations is a serious limitation that means truly quantitative data cannot be obtained from these graphs but, will not alter the general trends shown in the graphs which are, that the peak temperature, the fraction of the surface that melts and the time that the surface is molten all increase with increasing fluence.

The lack of any dependence of the ripples on angle of incidence and polarization of the incident light eliminates the possibility that the ripples are generated by E field interference effects.

The formation of ripples near the centre at lower fluence whilst they are formed only near the edge at higher is consistent with the fact that at higher fluences, the centre of the surface is molten long enough that any induced ripples will have decayed away before resolidification occurs.

Table 3.3 shows that in the centre only widely spaced ripples with their

Al heat distribution...r...z....

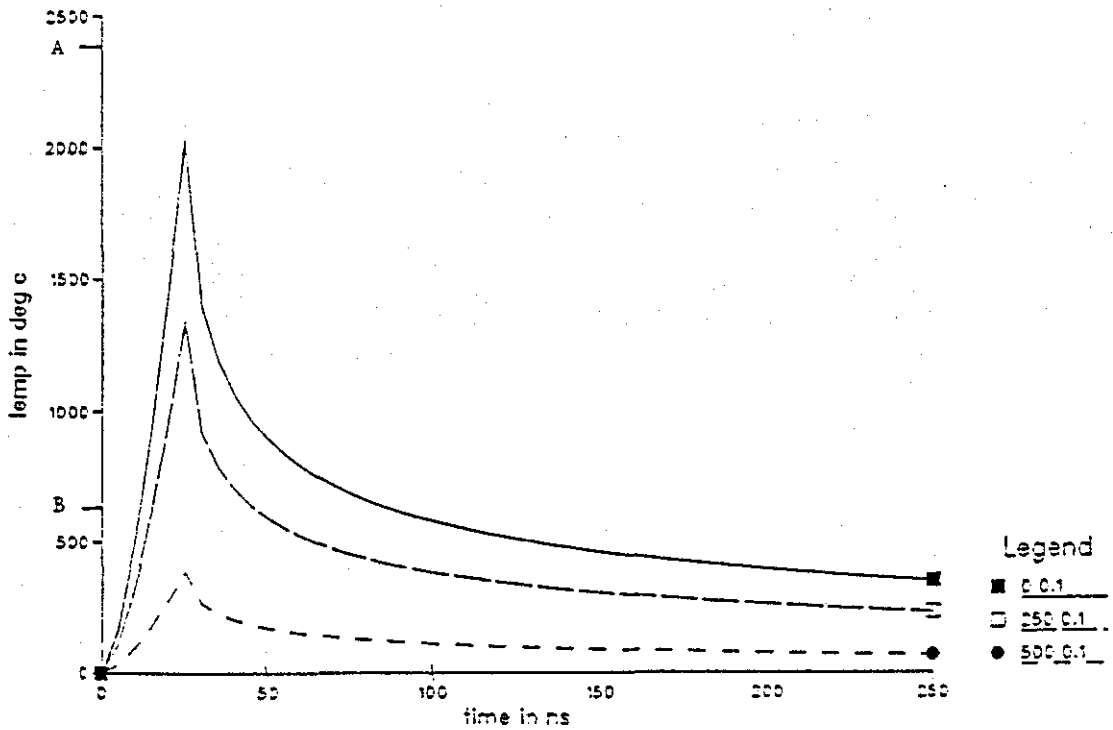


Fig 3.7a

Temperature profile of Al for an incident laser fluence of 2 Jcm^{-2} .

The first number in the legend is the distance from the centre of the beam and the second is the depth below the surface, both in microns.

Label A is the vaporization temperature and label B is the melting point.

Al heat distribution.....z....

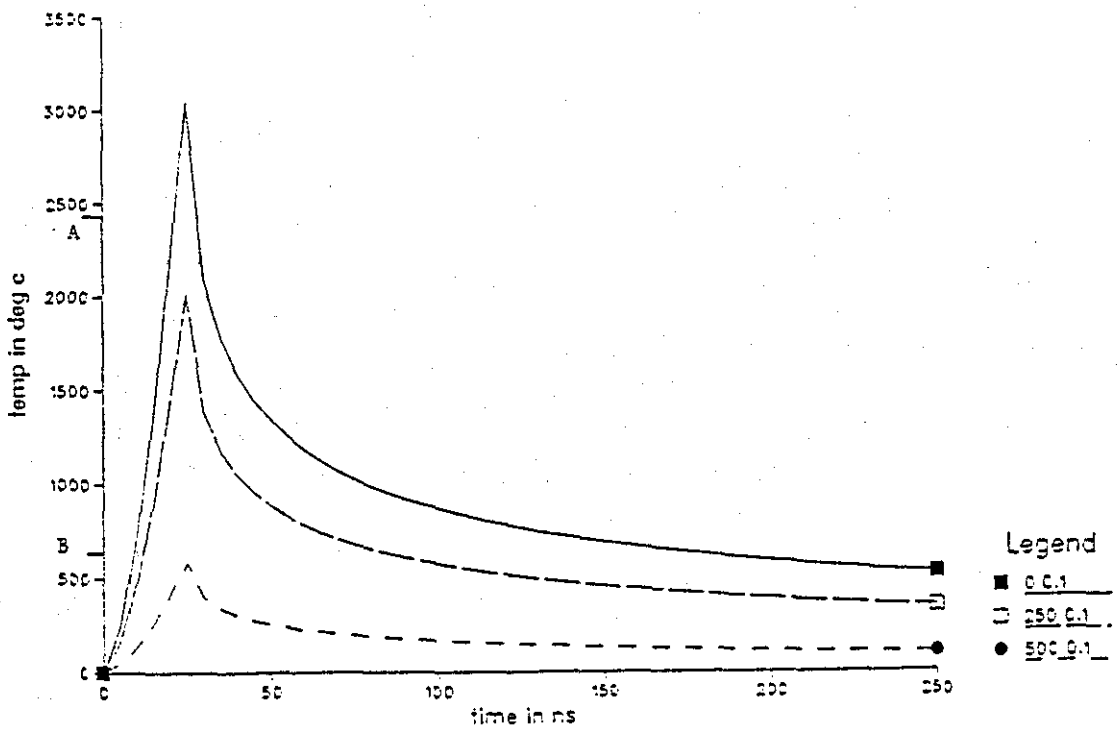


Fig 3.7b

Temperature profile of Al for an incident laser fluence of 3 Jcm^{-2} .

See fig 3.7a for an explanation of the legend and labelling.

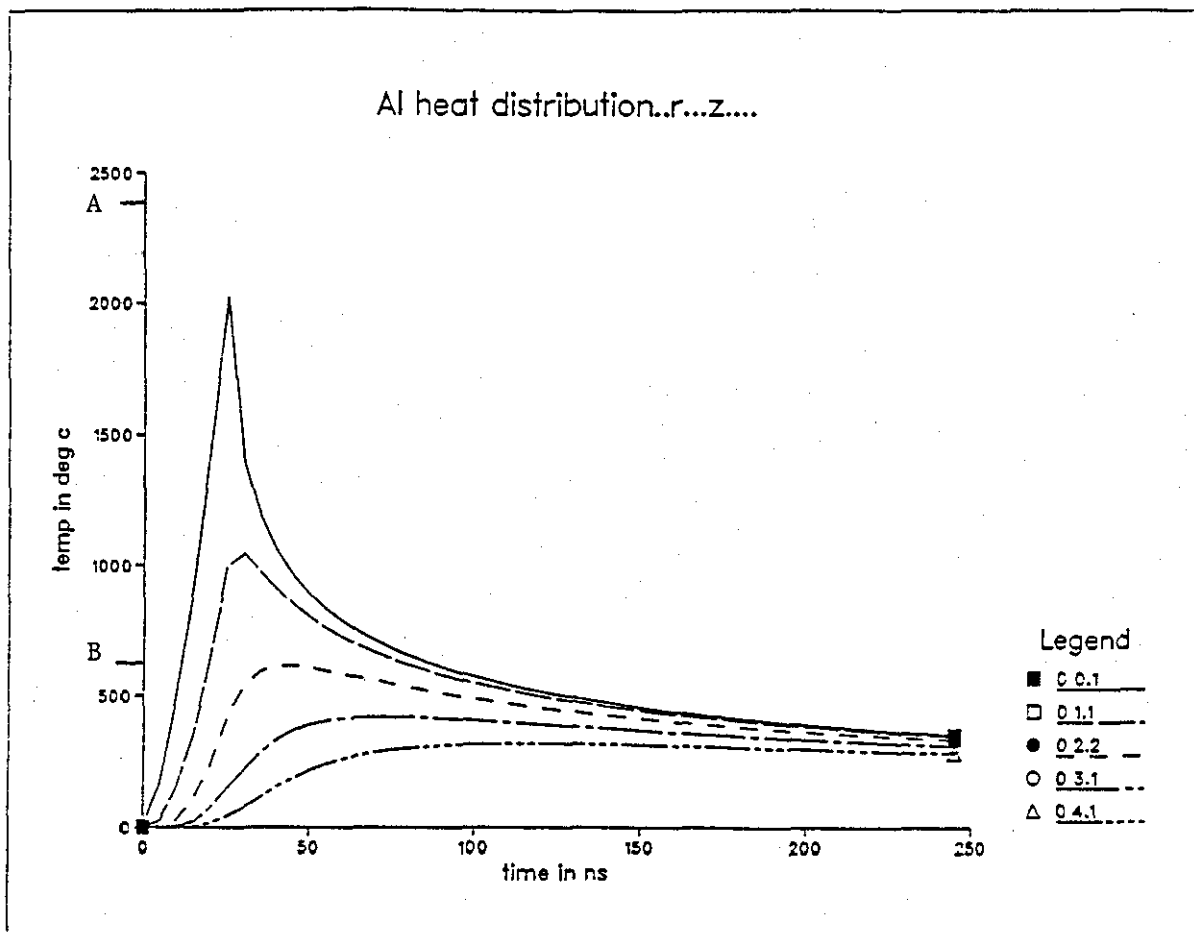


Fig 3.7c

temperature profile as a function of depth below the Al surface.

See fig 3.7a for an explanation of the legend and labelling.

long decay time could be left when the surface resolidifies. The lack of such large ripples the author attributes to the slow velocity of these ripples since, even if such ripples were initiated they would not be able to travel more than about 1 - 2 μ m in the time that the surface is molten which is a small fraction of their wavelength..

The observation that the ripple spacing decreased towards the outer edge of the damage site is consistent with LICW since, at the outer edges if the surface melts at all, it will only be for a short time due to the reduced fluence at the edge which, will enable short wavelength ripples to be frozen in before they have decayed away.

The increase in the fluence required to generate the ripples with increasing angle of incidence can be attributed to two effects, the increase in both the reflectivity and beam target interaction area with increasing angle of incidence. The former results in less light being absorbed and the latter to the incident light being spread over a larger area of the sample.

The reduced formation fluence required with multiple laser pulses can probably be attributed to a preferential coupling mechanism whereby once the ripples have been initiated as very small surface undulations, it is then "easier" for further shots to enhance the pattern to a detectable point, and thus the formation fluence is effectively reduced.

3.3.8 Conclusions

It is the author's contention that even without time dependent studies, that based on the evidence discussed above these ripples can be uniquely attributed to LICW.

Table 3.1

Material Parameters of Al

thermal conductivity	201	$Wm^{-1}K^{-1}$
specific heat capacity	913	$JKg^{-1}K^{-1}$
density of solid	2700	Kgm^{-3}
density of liquid	2400	Kgm^{-3}
surface tension	0.84	Nm^{-1}
viscosity	0.001	Nsm^{-2}

Excimer Beam Parameters

pulse length	25	ns
normal incidence spot size	0.55	mm
coupling factor	0.4	

Table 3.2

Material	Wavelength nm	n_r	n_i
Al	1060	1.8	9.3
	249	0.2	2.9
Ge	1060	4.0	0.1
	249	1.4	3.2

Table 3.3

t_i ns	wavelength μm	decay time ns	velocity ms^{-1}
1	0.2	1.2	105
5	0.6	11.0	61
10	1.0	28.0	47
20	1.5	70.0	38
30	2.0	120.0	33
100	4.5	600.0	22
200	7.1	1500.0	18

3.4 LIPSS

3.4.1 Introduction

Of the three type of LIRS ,LIPSS are not only the most difficult to generate but are also the most difficult to accurately theoretically explain.

LIPSS (fig 3.8) are a widely observed phenomenon and consequently there are many differing explanations as to their cause.Common to all these explanations and the generally excepted cause of LIPSS is interference between the incident E field and that due to a "surface field" caused by scattering of the incident light by surface roughness.This interference produces via the resulting E field a periodically varying surface temperature which, it is argued causes periodic melting, producing on resolidification the observed ripple patterns.

Since interference involving the incident light is involved in the formation of LIPSS, the observed spacings are directly related to the wavelength of the incident light.When in the course of attempting to obtain LIPSS the author used very high fluences, sets of ripples were observed that had the same dependence on the angle of incident and polarization of the incident light as did the LIPSS produced at low fluences but, had spacings consistent with the wavelength of the incident light being many times larger than it really was.These ripples designated as anomalous LIPSS (ALIPSS) by the author are discussed in detail in section 3.5 along with the author's contention that they can be attributed to the interaction of the incident light with the large plasma above the sample's surface that was always associated with their formation.

Most of the reported work has been performed in the infrared using pulsed lasers (at 1.06 or 10.6 μ m) using semiconductor samples¹⁵⁻¹⁷, in particular on Ge which has been shown to be especially suitable for LIPSS formation.Limited work on crystal insulators such as NaCl and on metals ,notably Al has also been performed.

At the time of writing ,to the author's knowledge, the work herein is the first quantitative study of LIPSS at UV wavelengths which, since in the UV Ge behaves optically as a metal (i.e. $\epsilon < -1$ where ϵ is the dielectric constant) is a region of particular significance for two reasons.Firstly, new and previous unreported structures are observed to be generated by S polarized light and secondly, by using information from these new structures, a more testing comparison between experiment and current theory is possible than can be achieved in the infrared.

For reasons to be discussed, the theory used in this work is based upon that developed by Sipe et al¹⁸ (hereafter known as I) with parameters

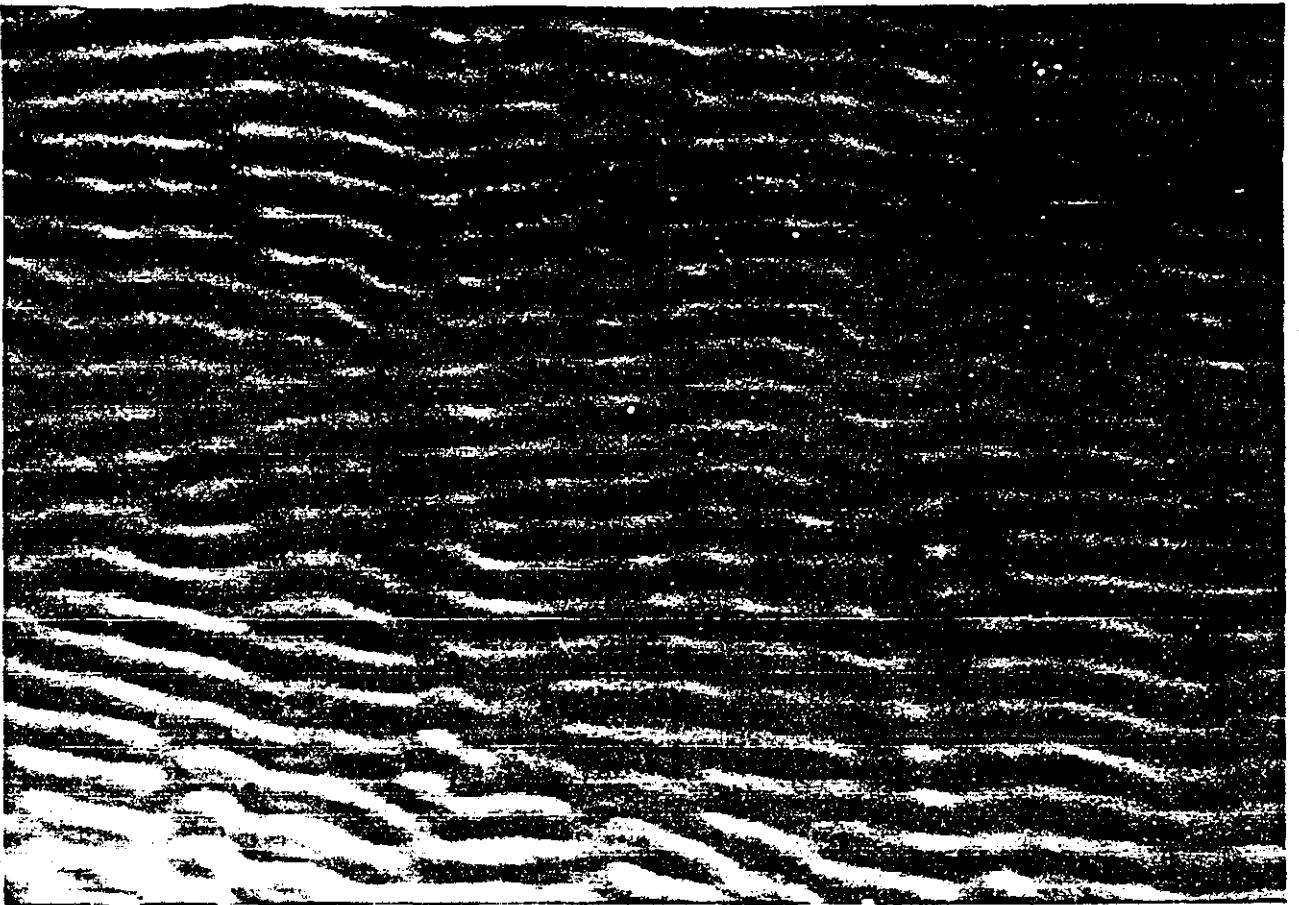


Fig 3.8

Germanium sample irradiated by P polarised light incident at 75°
magnification X2500.

appropriate to 249 nm laser irradiation.

3.4.2 Review

3.4.2a Theories

LIPSS were first reported some 20 years ago by Birnbaum¹⁹ on semiconductors irradiated by a pulsed ruby laser and at the time were attributed to diffraction effects associated with the optical elements used in the experiments. It was not until Emmony et al¹⁵ first reported LIPSS generated on Ge by a pulsed CO₂ laser that the basis of currently accepted theory i.e. interference between the incident light and some form of "surface wave" was first invoked. In their paper these authors assumed that a surface scattered wave of wavelength equal to that of the incident radiation was induced on the sample surface and that this interfered with the incident light yielding regions of large electric field which caused damage and resulted in the observed patterns. From a simple geometric argument they predicted and subsequently experimentally verified that for light incident of wavelength λ at an angle θ there should be two patterns with spacings

$$\lambda/(1 \pm \sin\theta) \quad \quad \quad 17)$$

Whilst this phenomenological model (here after known as the "surface scattered wave" model) is appealing it is

1) physically incorrect since it requires longitudinally polarized waves which would violate Maxwell's equations and

2) is unable to explain the dependence of the ripple spacing on the polarization of the incident light.

Two approaches depending upon whether metals/semiconductors or low index dielectrics are involved were developed in the late seventies/early eighties.

1) Insulators

Temple and Soileau²⁰ initially reported that LIPSS on dielectrics such as NaCl could be explained on the basis of resonant defect enhancement of the electric field in the dielectric. Subsequently²¹ they explained the phenomenon on the basis of the non radiative parts of the electric field associated with induced polarization charges and were able to show that at normal incidence the spacing on dielectrics of refractive index, n , should

be λ/n . The importance of the non radiative parts of the E field was subsequently shown rigourously in I.

2) Metals

Within this area there are two clearly identifiable approaches used to explain LIPSS..

- A) Surface electromagnetic waves (SEW) and
- B) Solutions to Maxwell's equations in the presence of a rough surface.

A) SEW²² are polarized waveforms that form valid interfacial solutions to Maxwell's equations provided that at least one of the media involved has $\epsilon < -1$ (where ϵ is the dielectric constant given by $\epsilon = n_t \cdot n_t^*$) i.e. behaves optically like a metal. Prokhorov et al²³ have solved Maxwell's equations to obtain the amplitude of the SEW as it propagates along a rippled interface. Emel'yanov et al²⁴ have performed a two step analysis whereby solutions are found to the diffraction of an incident electromagnetic wave by an arbitrary Fourier component k of surface roughness and the resulting electric fields used to provide a surface heating term in a non linear heat conduction equation. A temperature profile is then obtained as a function of Fourier component and it is argued that if this is large at a component k then, LIPSS of spacing k are formed.

Both approaches predict ripples of spacing given by 17) for P polarized light and orientated perpendicular to the plane of incidence and spacings of

$$\lambda/\cos\theta$$

18)

for S polarized light orientated parallel to the plane of incidence.

Given that for SEW to exist at least one of the media must have $\epsilon < -1$ and as the experiments are usually carried out in air it is difficult to reconcile this mechanism with the generation of LIPSS in dielectric solids, although the work of Keilmann and Bai²⁵ can be explained by SEW as the material used (quartz) has a dielectric resonance around the laser wavelength used (10.6 μm).

B) In this approach as proposed by Guosheng et al¹⁶ Maxwell's equations are solved in the presence of a corrugated surface of arbitrary period corresponding to a general Fourier component (k) of surface roughness. Expressions are developed for the total diffracted light as a

function of surface roughness. It is then argued that when this is large LIPSS are formed. They predict that for P polarized light the spacing is given by equation 17 and that there are no LIPSS for S polarized light.

Although a formal solution of Maxwell's equations is undertaken in this work there is, a requirement for the surface to melt uniformly which is inconsistent with the work in III where, it is shown that LIPSS can form at sufficiently low fluences that there can only be melting in the immediate vicinity of the maxima in the E field i.e. localised surface melting.

The main problem with the above theories is that none of them can explain all of the following three important experimental observations...

1) The existence of LIPSS of spacing $\lambda/\cos\theta$ orientated parallel to the component of incident wavevector that is parallel to the surface for p polarized light, as first reported in Young et al¹⁷ (hereafter known as II).

2) The existence in the UV of spacings of the order of $\lambda/(1 - \sin\theta)$ for S polarized light and

3) The large deviations at large angles of incidence of the observed spacings from $\lambda/(1 - \sin\theta)$ in the case of P polarized light.

Observations 2 and 3 are those of the author and are discussed in more detail in subsequent sections.

The theory derived in I based on an integral solution method for Maxwell's equations in the presence of a randomly rough surface is able to explain all of the above points and is discussed in detail in section 3.4.6b.

3.4.2b Experimental Work

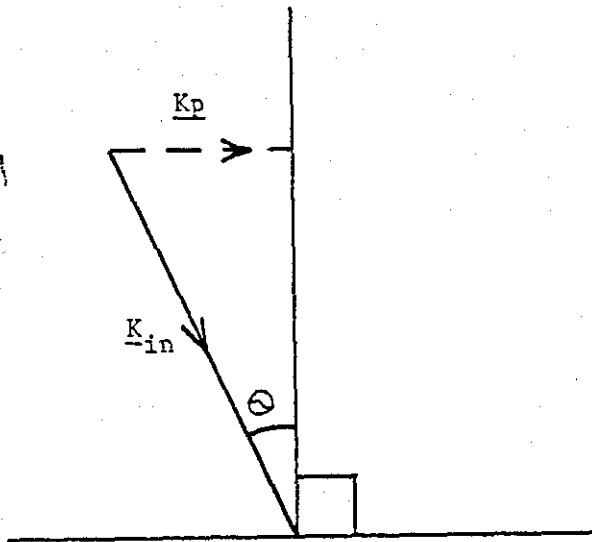
LIPSS have been found on a wide range of materials for lasers ranging from high power picosecond to low power cw beams²⁶⁻²⁷ with most of the work being performed on Ge at 1.06 microns.

LIPSS can be produced by a single pulse as in II but in general around 30-40 shots are required to establish a distinct pattern which, if the fluence is too high or the number of shots used too great, can be destroyed.

As shown in I it is the component of wavevector parallel to the surface that is important and thus the convention that the component of wavevector parallel to the surface is designated by k_p (fig 3.9a) is introduced.

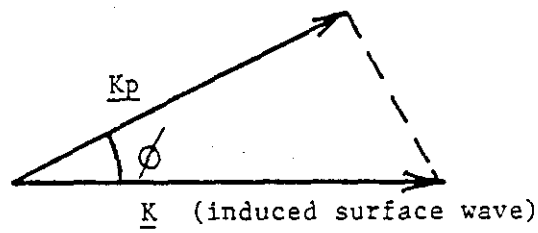
The orientation of the ripples is expressed in terms of the angle between k_p and a "ripple vector" which is defined by consideration of the Fourier transform of the pattern as now detailed.

Fig 3.9a



$$\underline{K}_p = \underline{K}_{in} \sin \theta$$

Fig 3.9b



If the Fraunhofer diffraction (i.e. Fourier transform) pattern of a diffraction grating with vertical lines is observed, dots are seen in the horizontal direction and a grating vector in the horizontal direction can be defined. By analogy to this, the ripple vector is defined to be perpendicular to the lines of ripples in a pattern i.e. horizontal if the ripples are in vertical lines and vertical if they occur as horizontal lines as this definition makes the ripple vector occur in the same direction as the maxima in the Fourier transform pattern of the ripples. Hence for example, the ripple vector in fig 3.8 would be in the vertical direction.

LIPSS can be formed with both P and S polarized light with the former being far more widely reported. For P polarized light the induced fringes are generally parallel to \underline{k}_p . These ripples of spacing given by 17) for small angles are labelled p^\pm type fringes (the + corresponds to the ripple of smaller spacing). Structures perpendicular to \underline{k}_p and of spacing given by 18) have been reported²⁸, particularly in II. For S polarized light essentially the only previous report of fringes is in II where a complex pattern with no simple wavelength or orientational dependence is reported. As will be shown later S polarized LIPSS can, at least in the UV, exist with both simple wavelength and orientational dependences.

Generally speaking for high index dielectrics/metals/semiconductors the normal incidence ripple spacing is given by λ but for low index transparent dielectrics the spacing is λ/n where n is the refractive index. However, recently²⁷ it has been shown that anomalous ripples of spacings $2\lambda/(n+1)$ can exist. Non-linear LIPSS resulting from high order interference terms have been reported by Fauchet and Sigman²⁹ who attributed the effects to interference between induced surface gratings of varying wave vectors.

Time dependent studies of LIPSS have been done, notably by Ehrlich et al³⁰ and in III. In the former, Ge was irradiated at high enough fluences to effectively uniformly melt (i.e. homogeneously melt) the surface, thereby, since liquid Ge is optically metallic, allow SEW's to exist which, it is then argued, cause the observed effects. In III the fluence is kept sufficiently low that from reflectivity measurements the authors deduce that the surface melts only in the vicinity of the peak in the E field (i.e. localised or inhomogeneous melting) and under these conditions SEW's cannot exist. They argue that it is the non radiative parts of the induced E field (see I) that provide the "waves" required for the interference effect that produces LIPSS.

The work in III and results by the author herein show that the spacing and morphology of the induced ripples is dependant upon the degree of surface melting. These differences arise respectively from the different dielectric

constants and formation mechanisms that occur in each phase.

A few workers have used UV light in LIPSS experiments, in particular Osgood and Ehrlich³² and Ursu et al³³. Osgood and Ehrlich used a low power cw beam in laser photodeposition of thin metal films. SEW's are generated on the metal surfaces which interfere with the incident light producing a periodic heating of the surface which dramatically effects the nucleation rate leading to preferential periodic photodeposition. Ursu et al³³ have looked in some detail at the interaction of 308nm XeCl radiation with metallic surfaces and have reported finding a multitude of structures, some non periodic and some periodic but with no obvious dependence on wavelength or angle of incidence of the light. It is the author's suggestion that the observed periodic structures with anomalously large spacings could be due to the interaction of the incident light with the plasma immediately above the sample's surface that was always associated with their formation. This is further discussed in 3.5.

3.4.3 Coordinate System Used

In order to use the results of the theoretical calculations and indeed to appreciate the experimental results it is essential that the reader is fully conversant with the coordinate system used and the orientation of the E field of the incident light in this system.

The system used is shown in fig 3.10 where the excimer beam is taken to be travelling in the positive z direction away from the reader. When viewed in the direction of beam propagation the x direction runs right to left and the y direction vertically up. Under this system the E field of the incident light is always in the y direction irrespective of the polarization incident on the samples.

When S polarized light was used the sample was rotated in the xz plane whereas for P polarized light the sample was tilted in the yz plane and kept normal to the beam in the xz plane. Hence for S polarization the incident light had the component of its wavevector that was parallel to the surface in the xz direction whilst for P polarized light the parallel component was in the yz direction. This meant that effectively the geometries for S and P polarized light were rotated by 90° relative to each other which, as will be shown later is critical to understanding the orientation dependence of the ripples on the polarization of the incident light.

3.4.4 Experimental Procedure

The same experimental procedures and fluence ranges were used in this work as was used in the LITS work. However, the work initially concentrated on

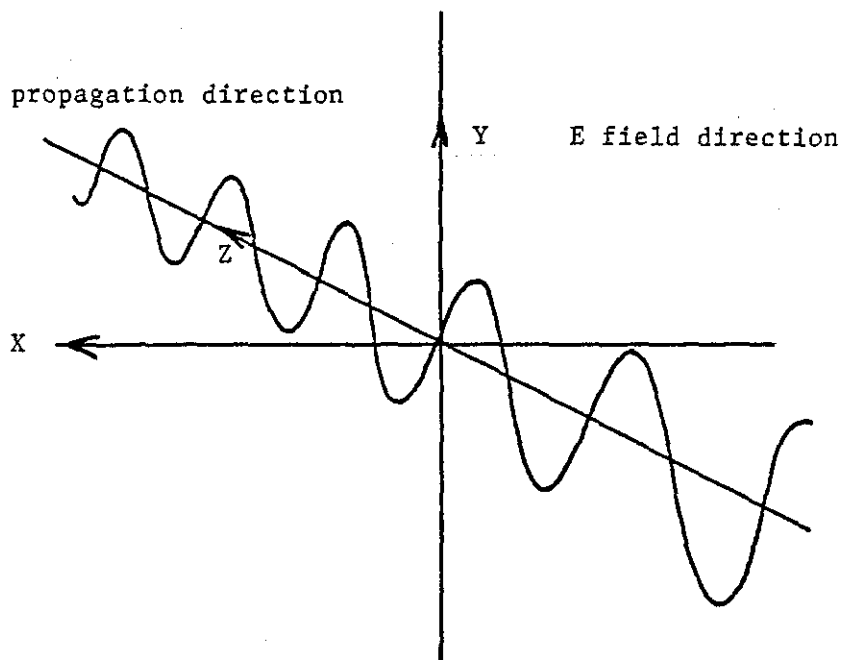


Fig 3.10

angles of incidence of the incident light of at least 30° , although later work was performed at normal incidence.

The reason for this can be seen by considering the spacing of the generated ripples which, was expected to follow 17) and possibly 18). For reasons of convenience, cost and speed of operation OM was, in the first instance the preferred analysis technique. The available microscope had a resolution limit of $\approx 0.5 \mu\text{m}$ which, with a wavelength of 249nm corresponds to ripples that have been generated at an angle of incidence of 30° . Use of angles smaller than this would have meant that even if LIPSS had been produced they could not have been resolved optically, necessitating the use of time consuming and expensive SEM and TEM as primary rather than secondary analysis techniques.

3.4.5 Results

LIPSS formation was found to be a function of on-target fluence, number of shots incident and, the angle of incidence and polarization of the incident light. At low fluence and or number of shots, the ripples tended to form in the centre of the damage site i.e. where the beam is most intense but, at higher fluences the centre part became badly damaged and the ripples formed in an annulus towards the edge of the damage site where the local fluence was lower. If too many shots were used then, in general the pattern was 'washed out' leaving a badly damaged site with a variety of intricate anomalous structures on it. At still higher fluences, LIPSS were obtained with the spacing in the centre of the site different to that found at the outer edges as shown in figs 3.11a and b.

As far as the author was able to observe, even if several hundred shots were used no LIPSS were formed unless there was some form of surface breakdown i.e. damage that was clearly visible, usually in the form of a laser induced plasma. Typically, the most well defined ripples were seen after 40 to 50 shots, with breakdown having occurred after 15 to 20 shots. Even at fluences that would give breakdown after 20 shots unless at least 4 shots during which there was a small plasma (breakdown) had occurred there was no detectable LIPSS although the surface had been modified, typically breaking up into small globules (fig 3.12).

If surface breakdown occurred after less than 10 shots and even if only a few subsequent shots were used, the surface was left badly damaged and generally with no normally spaced LIPSS discernible. Under these conditions of early breakdown periodic structures were observed but with very large spacings e.g. $50 \mu\text{m}$ (fig 3.13) when the observed/predicted spacing for LIPSS was an order of magnitude or more smaller. The author has designated these as ALIPSS and they are discussed in 3.5 but ignored for the rest of this

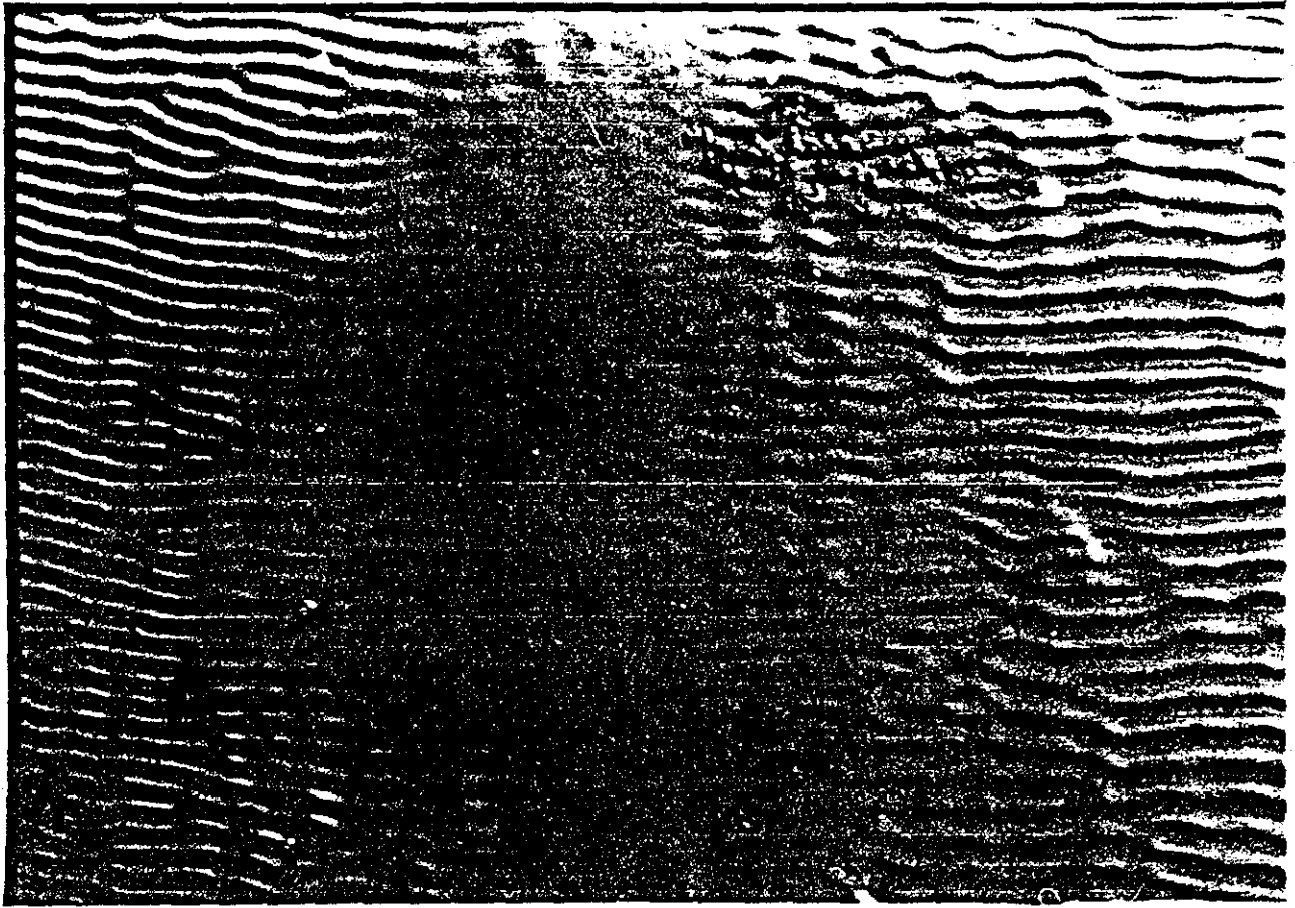
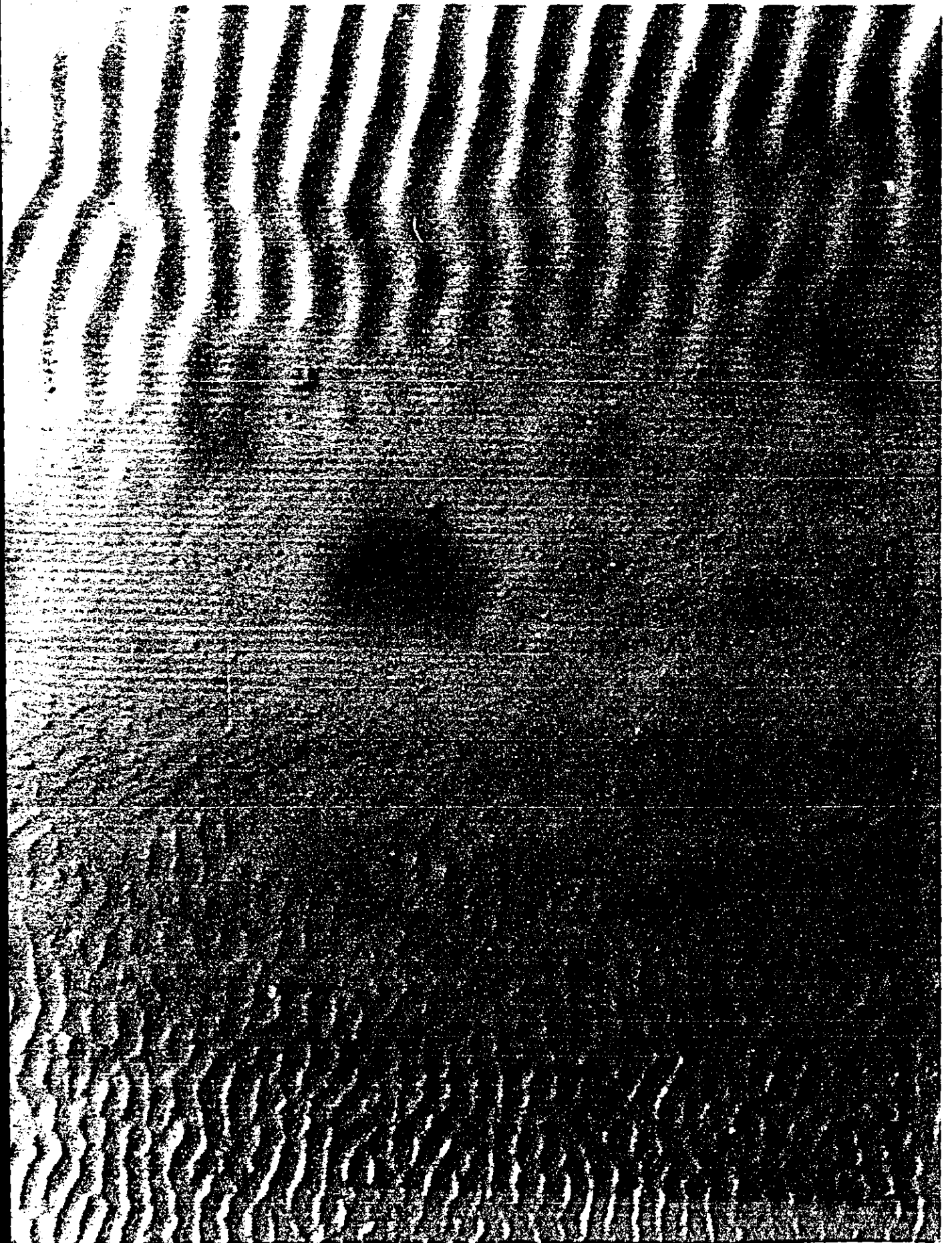


Fig 3.11

Germanium sample P polarised light incident at 75°

Above a) Magnification X1000 Note the transition from ripples formed on a locally melted surface (smaller spacing) to those formed on a uniformly melted surface (larger spacing) and the region between where there is essentially no patterning.

Overleaf b) Magnification X2500. This photo shows both types of p^- fringes together with the extremely uniform c type orientated perpendicularly to the others. Note that the ripples formed on the locally melted surface appear to have a much flatter more sinusoidal profile.



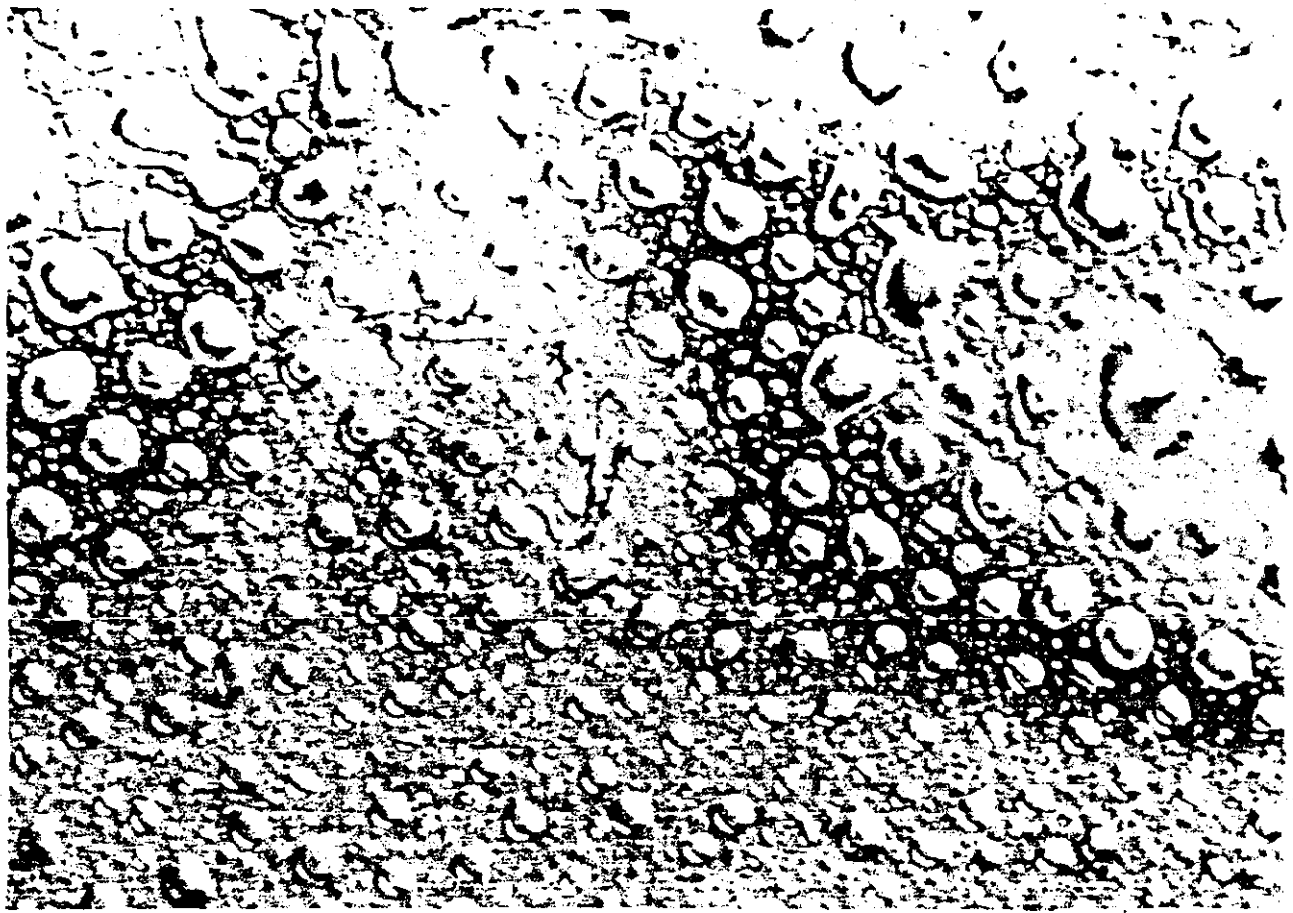


Fig 3.12

Germanium sample P polarised light incident at 75° magnification
X950. Surface has broken up into small globules following the incidence of
one laser pulse.



Fig 3.13

Germanium sample P polarised light normally incident magnification X500. Sample hit by multiple high energy pulses and shows complex structure of very large spacing and depth.

section.

It is not clear as to the role of the plasma in the formation of LIPSS but it is the author's contention that, the presence of a plasma in under these conditions is an indication that the surface has at least in parts reached melting point.

As the number of shots during which a plasma occurred was increased the pattern would build up becoming observable after about 4 plasma shots, reaching its most distinct after some 20 to 40 shots and then being gradually 'washed out' as the number of shots increased.

The fluence required to form the fringes in general increased with the angle of incidence and was higher for S polarized light than for P.

Four types of patterns two each for S and P polarized light were observed. For S polarized light (fig 3.14) the ripples were found to run parallel to k_p and have a spacing given by $\lambda/(1 - \sin\theta)$ even for large angles. Following the nomenclature introduced in II we label these as s^- type fringes. A second well defined pattern was observed to run almost perpendicular to k_p with a spacing that has no simple dependence on either the angle of incidence or polarization of the incident light. The author has, for reasons to be discussed designated these as b type fringes.

For P polarized light (fig 3.11a and b) two types of ripples were found, the more common pattern runs parallel to k_p but with a spacing that shows substantial deviations from $\lambda/(1 - \sin\theta)$ at large angles i.e. angles $>60^\circ$. The author has designated these as p^- type fringes. At large angles of incidence an extremely uniform set of fringes that run perpendicular to k_p i.e. parallel to the electric field are produced with a spacing of the order of $\lambda/\cos\theta$ and again following II these are labelled as c type fringes. It is interesting to note that the c type fringes form at lower energies than the p^- type. These c type fringes were often found to cover areas of 1mm^2 or more.

Within the p type fringes two distinct types of pattern with different spacings were observed particularly at large angles (figs. 11a and b). Of the patterns produced the one with larger wavelength occurred nearer the centre of the damage site (where the fluence is highest) than the other which, occurred towards the edge.

In general the fringes that formed on Al (figs 3.15a and b) showed the same dependence on the angle of incidence and polarization of the incident light as Ge but were less well defined than their counterparts on Ge with, the fringes on Al requiring typically two to three times more fluence to form than those on Ge. A noticeable difference was the complete absence of any of the uniform c type fringes.

TEM studies revealed a variety of sub excimer wavelength (i.e. $<249\text{ nm}$)

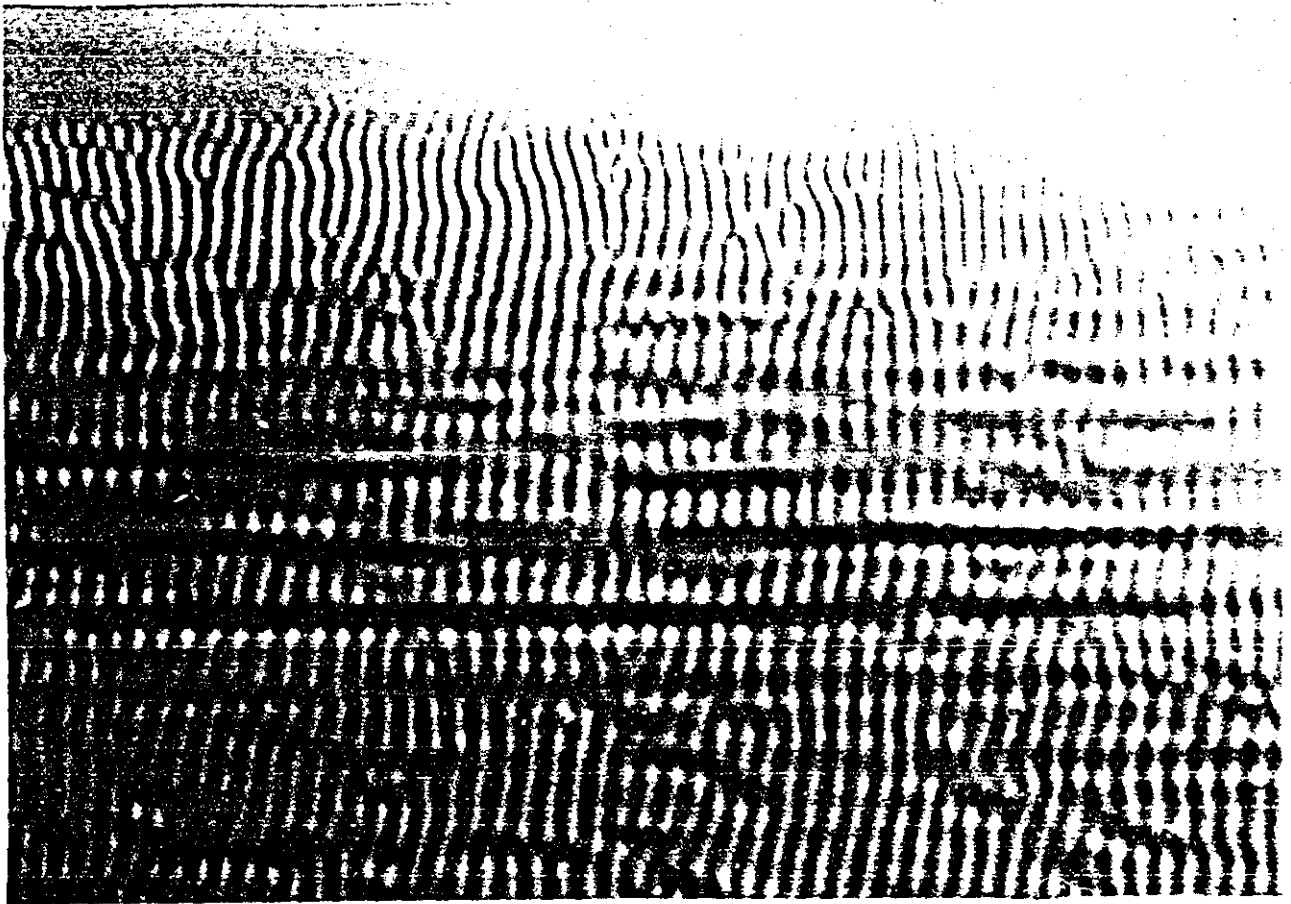


Fig 3.14

Germanium sample S polarised light incident at 70° magnification X1250. Note variation of spacing vertically along the picture which is attributed to a varying dielectric constant due to a temperature gradient. Also note large near horizontal b type fringes of spacing $\approx 8 \mu\text{m}$.

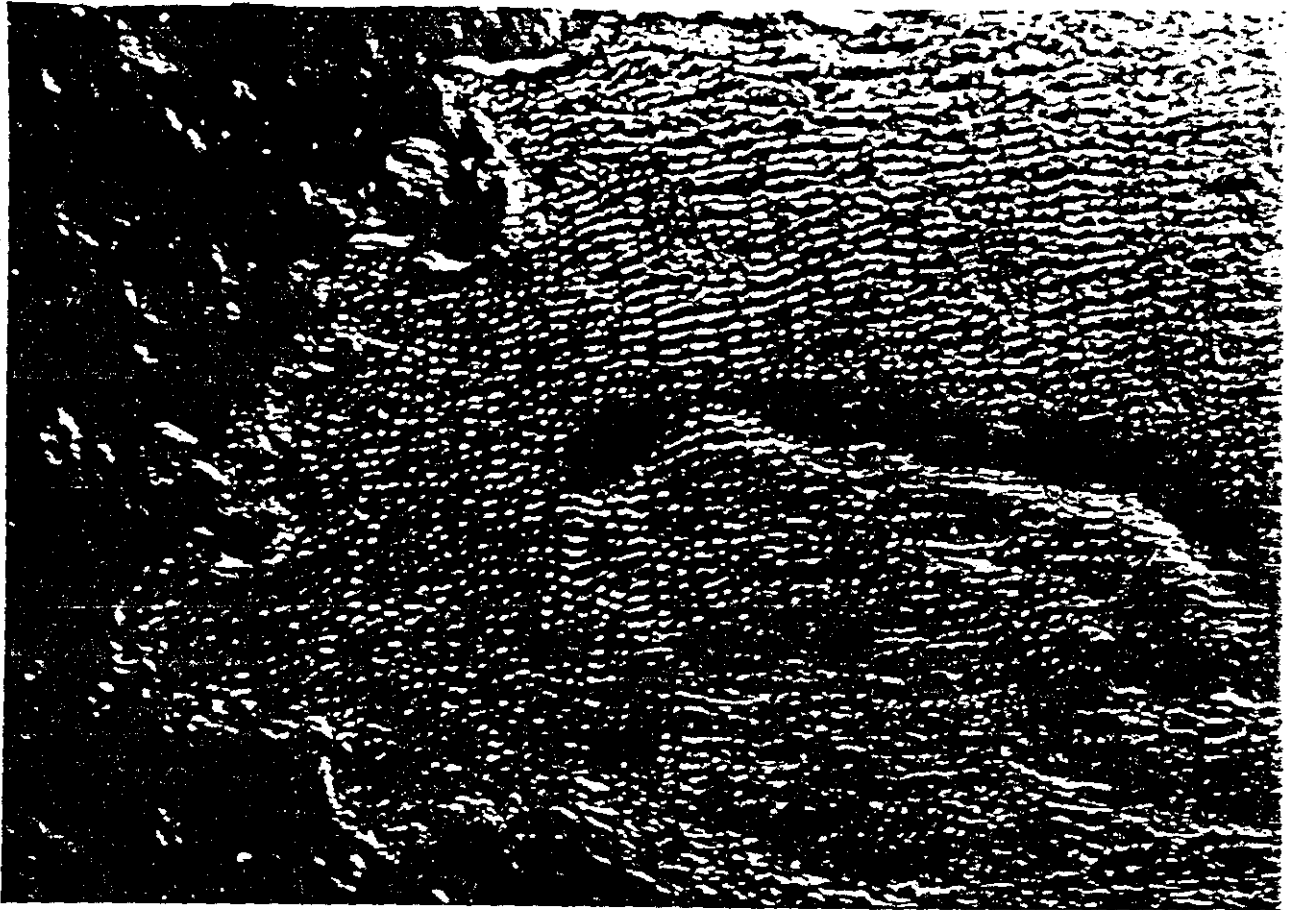


Fig 3.15a

Aluminium sample irradiated by P polarised light incident at 60°
magnification X1250. Note the poor quality of fringes as compared to those on
Germanium.

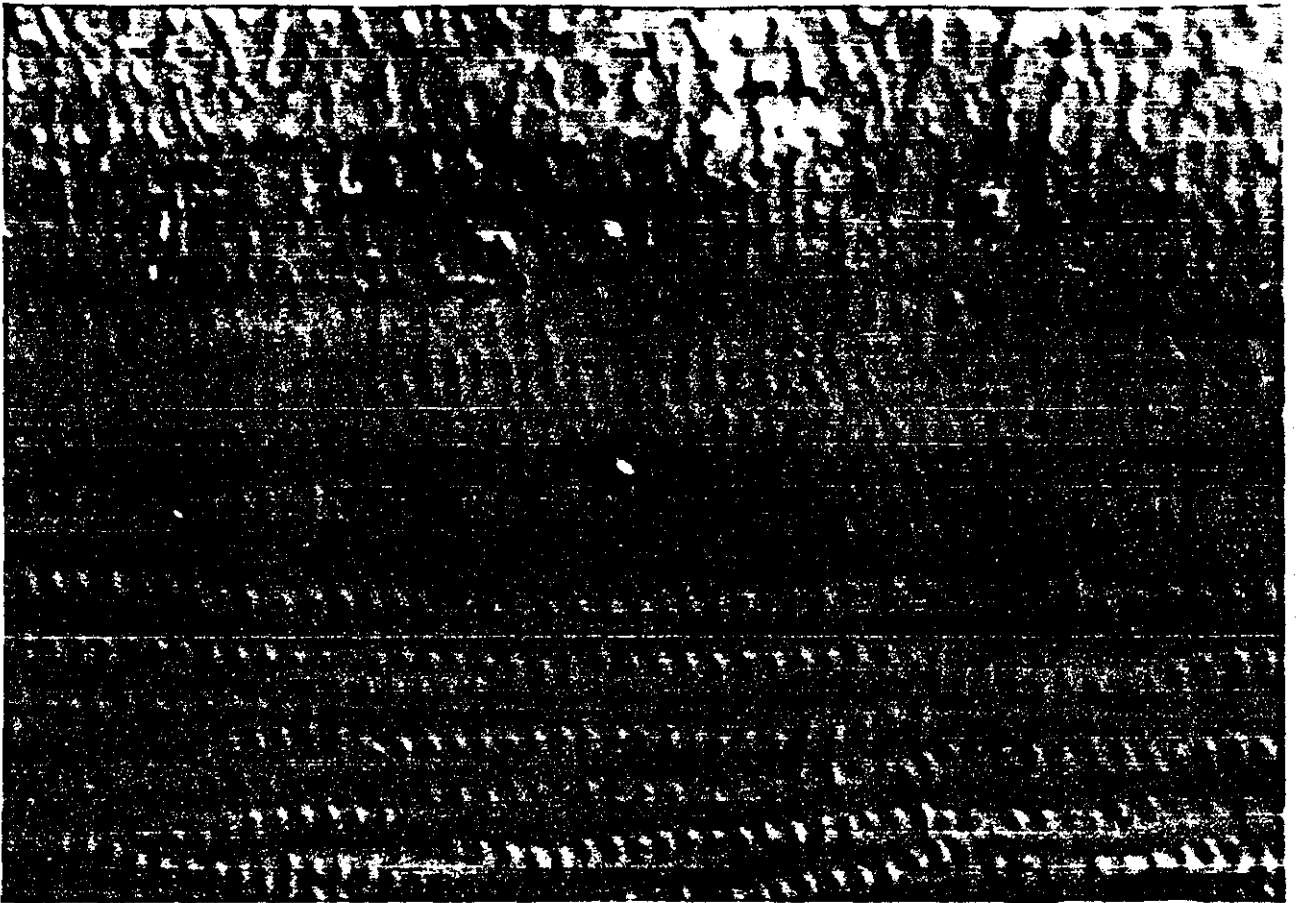


Fig 3.15b

Aluminium sample irradiated by S polarised light incident at 45°
magnification X5000. Fringe quality comparable to that in Germanium. Note the
large ($\approx 2 \mu\text{m}$) near horizontal b type fringes.

structures ranging from approximately 300 nm to 90 nm depending upon angle of incidence, polarization and the type of sample. Unfortunately due to both the replication process and the intrinsic uncontrollable rotation of the image in the TEM itself it was essentially impossible to accurately determine the orientation of the observed fringes, only their spacing could be found.

Results for LIPSS are summarised in table 3.4 with, table 3.2 listing the optical constants³⁷ of Al and Ge (solid phase) at 249nm and also those at 1.06 microns for comparison purposes (see 3.4.7). The figures quoted in table 3.4 for formation fluence are based on using the normal incidence beam target interaction area whereas those quoted for normalised formation fluence have an allowance for the variation of absorption and beam target interaction area with angle of incidence built into them (see 3.4.6d).

3.4.6 Theory

3.4.6a Introduction

Ideally the theory (or theories) used should be able to explain three main effects..

- 1) the spacing and orientation of the ripples
- 2) why the surface does not resolidify to its original profile
- 3) the dependence of the formation energy on angle of incidence and polarization of the incident light.

Three separate pieces of theory are used to explain the above effects, of which that used to explain the last effect was developed by the author.

3.4.6.b Spacing and Orientation Theory

Given that LIPSS are generally accepted to be due to interference between the incident light and a "surface wave" which has been initiated by scattering due to the microscopic roughness of the surface then, the obvious starting point in any theory to explain these effects is Maxwell's equations and ideally these should be solved to yield the electric field on a microscopically randomly rough surface. This to say the least is extremely difficult and not surprising alternative approaches or models were sought.

Various theories have thus been put forward over the last ten years or so, each of which can explain some of the observations, but to date the only one which can account for all previously reported observations is that developed in I based on a formal solution of Maxwell's equations in the presence of a

Table 3.4

Material	Angle of Incidence°	Orientation of Fringes*	Polarisation	Spacing Microns	Fluence of Formation $J\ cm^{-2}$	Normalised Formation Fluence $J\ cm^{-2}$
Ge	75	Perpendicular	p	0.78	0.84	0.37
	"	Parallel	p	3.28	1.28	0.56
	"	"	p	4.11#	-	-
	60	"	p	1.56	0.66	0.50
	45	"	p	0.80	0.60	0.54
	30	"	p	0.50	0.60	0.58
	70	"	s	3.72	2.21	0.29
	60	"	s	1.92	0.82	0.22
Al	75	"	p	2.26	3.09	1.32
	60	"	p	1.38	1.74	1.41
	45	"	s	0.80	2.55	1.26
	30	"	s	0.50	1.61	1.19

* Fringe orientation is measured relative to the direction of k_p for each polarisation.

Fringes formed in the centre of damage site on a liquid pool, see text.

randomly rough surface yielding an expression for the inhomogeneous deposition of energy into the material. As the work in I is to the author's knowledge the most comprehensive and rigorous attempt to model LIPSS it is used as the basic theory in the work herein.

The work in I is a solution to Maxwell's equations, based on an integral transform method developed by Sipe some years earlier and works by decomposing the electric fields into their various Fourier components (k), written in terms of dyadic Green's functions which enables explicit identification of the S and P polarized components.

Solutions are then built up to the electric fields in what is called the selvedge region (i.e. a region of roughness bounded by air/vacuum on one side and the material bulk on the other but where the depth of roughness is restricted to being \ll incident wavelength) by separately treating the longitudinal and transverse components as well as using a variational approach to allow for local field effects with parameters to model the roughness of the surface being incorporated.

Once the selvedge fields have been found, they are added vectorially (to allow for phase variations) to the normally refracted electric field component (which has wavevector k_p parallel to the surface) to yield the total electric field in the selvedge region. By taking the time average of the square of the total field, a measure of the energy deposited (inhomogeneously) into the material as a function of k , angle of incidence and polarization of the incident light is obtained.

The final expressions, reproduced herein, have only two adjustable factors known as the shape (s) and filling factors (f) and as explained in I/II the expression for S polarized light is essentially independent of these factors although the one for P polarized light depends upon their ratio.

The f factor corresponds to the ratio of the volume actually occupied by the material in the selvedge to that available if the region was entirely filled, and following II the value of 0.1 was used in all calculations. The shape factor is derived from the modelling of the microscopic random roughness of the surface which, was performed by assuming a square wave type filling of the surface but of random periodicity and essentially relates the width of any occupied area to the depth of roughness within that area. To this end it should be noted that the value of 0.4 used throughout (as in II) corresponds to spherically shaped surface sites.

In the theory it is shown that the inhomogeneous absorption of energy can be written as the product of two terms, both functions of k , where k is the induced surface wavevector, normalised to the incident wavevector ($2\pi/\lambda$) and given by..

where the second term in 19) is a slowly varying function of \underline{k} and essentially relates to the amplitude of surface roughness at wavevector \underline{k} whereas, the term $\eta(\underline{k}, \underline{k}_p)$, known as the efficacy factor, is a response function describing the efficacy with which surface roughness at \underline{k} leads to inhomogeneous energy deposition in the selvedge. It is shown in I that the efficacy factor exhibits sharp peaks as a function of \underline{k} and is essentially responsible for the observed spectra.

The basic premiss in I is that whenever $\eta(\underline{k}, \underline{k}_p)$ and therefore $I(\underline{k})$ is large i.e. has sharp peaks in it then the surface will damage (melt) and LIPSS formation is likely. Since $b(\underline{k})$ is shown to vary slowly with \underline{k} , it can be ignored and only the efficacy factor need be calculated. Indeed it is because $b(\underline{k})$, which essentially is the term representing the effects of surface roughness varies so slowly with \underline{k} that the formation of LIPSS is as shown in II and herein largely independent of the quality of the surface used provided it is not too rough, which in practice means that a surface must have been finished by polishing with at least a coarse paste.

To facilitate calculation of η and, following II, the convention that the angle between \underline{k}_p and the induced wavevector \underline{k} is designated as angle ϕ is introduced. Fig 3.9b shows the orientations of ϕ , \underline{k}_p and \underline{k} .

The expressions used in the calculations are given below.

$$\omega_0 = (\tilde{\omega}^2 - k^2)^{\frac{1}{2}} \quad (20)$$

where $\tilde{\omega}^2 = \frac{2\pi}{\lambda}$ and $k^2 = \underline{k} \cdot \underline{k}$

$$\omega = (\omega^2 \epsilon - k^2)^{\frac{1}{2}} \quad (21)$$

where $\epsilon = n_t^2$

$$t_p = 2\omega_0 n_t (\omega_0 \epsilon + \omega)^{-1} \quad (22)$$

$$4\pi\gamma_z = (\epsilon - 1) \left\{ \epsilon - (1 - f)(\epsilon - 1) [h(s) + R h_I(s)] \right\}^{-1}$$

$$4\pi\gamma_t = (\epsilon - 1) \left\{ 1 + \frac{1}{2} (1 - f)(\epsilon - 1) [h(s) - R h_I(s)] \right\}^{-1} \quad (23)$$

where $s = \text{shape factor}$, $f = \text{filling factor}$ and $R = (\epsilon - 1)/(\epsilon + 1)$

$$h(s) = (s^2 + 1)^{\frac{1}{2}} - s$$

$$h_I(s) = \frac{1}{2} \left[\left(s^2 + 4 \right)^{\frac{1}{2}} + s \right] - \left(s^2 + 1 \right)^{\frac{1}{2}}$$

$$\eta(\underline{k}, \underline{kp}) = 2\pi \left| v(\underline{k}_+) + v^*(\underline{k}_-) \right| \quad \text{where } \underline{k}_\pm = \underline{kp} \pm \underline{k} \quad 25)$$

$$h_{ss} = 2i\omega \left(\omega_0 t \omega \right)^{-1}$$

$$h_{kk} = 2i\omega \omega_0 \omega^{-1} \left(\omega_0 \epsilon + \omega \right)^{-1}$$

$$h_{zz} = 2ik^2 \omega^{-1} \left(\omega_0 \epsilon + \omega \right)^{-1}$$

26)

$$h_{zk} = 2ik\omega_0 \omega^{-1} \left(\omega_0 \epsilon + \omega \right)^{-1}$$

$$h_{kz} = 2i\omega k \omega^{-1} \left(\omega_0 \epsilon + \omega \right)^{-1}$$

giving as final expressions,

$$v(\underline{k}_\pm) = \left[h_{ss}(k_\pm) \left(\hat{k}_\pm \cdot \hat{x} \right)^2 + h_{kk}(k_\pm) \left(\hat{k}_\pm \cdot \hat{y} \right)^2 \right]$$

$$\times \gamma t \left| t_s(\underline{kp}) \right|^2 \quad \text{for S polarized light} \quad 27)$$

and

$$v(\underline{k}_\pm) = \left[h_{ss}(k_\pm) \left(\hat{k}_\pm \cdot \hat{y} \right)^2 + h_{kk}(k_\pm) \left(\hat{k}_\pm \cdot \hat{x} \right)^2 \right] \times \gamma t \left| t_x \right|^2 + h_{kz}(k_\pm) \times$$

$$\left(\hat{k}_\pm \cdot \hat{x} \right) \gamma t z \epsilon t x^* + h_{zk}(k_\pm) \left(\hat{k}_\pm \cdot \hat{x} \right) \gamma t z t x^* + h_{zz}(k_\pm) \gamma z \epsilon \left(t z \right)^2$$

$$\text{for P polarized light} \quad 28)$$

where

$$t_x = \omega(kp) \left(\tilde{\omega n} \right)^{-1} t_p(kp)$$

29)

$$t_z = kp \left(\tilde{\omega n} \right)^{-1} t_p(kp)$$

and \hat{A} means unit vector A

To greatly simplify the calculations, \underline{kp} is taken as being in the xz plane for both polarizations and hence there is a 90° rotation in the definition of angle ϕ between S and P polarizations. Thus $\phi = 0$ for S polarized light lies along the xz plane but along the yz plane for P polarized light.

The main omission in this theory is the neglect of any feedback mechanism

either inter or intra pulse whereby once preferential coupling at a given k has begun, the process can be enhanced by positive feedback to yield LIPSS (refer to III for a full discussion of this).

An extensive literature search failed to reveal any values for the optical constants of liquid Ge in the UV although in II there are values for the IR, therefore for the purposes of a comparison following the idea of Ehrlich et al³⁴ the author has used a free electron model in an attempt to derive at least very approximate values as described below.

On the free electron model the dielectric constant is given by

$$\epsilon = \left(1 - \frac{wp^2}{w^2} \right) \quad (30)$$

where wp is the plasma frequency and w the angular frequency of incident light. For this calculation the plasma frequency was determined as follows. Essentially no matter what the excitation mechanism of the electrons an upper limit on the electron density is that obtained by having all four valence electrons free for conduction and on this basis the electron density was set equal to four times the atomic density which upon substitution into 30) yields $n_f = 0$ and $n_i = 5$, figures which were duly used in the calculations.

Whilst this model is wholly inadequate the author is only intent upon showing that a refractive index of the order of that given above (which corresponds to a large negative dielectric constant) can be used to explain the existence of the two differently spaced p type patterns.

A computer program was written to evaluate the efficacy factor as a function of induced wavevector k , angle of incidence and polarization of the incident light for Al and Ge at 249 and 1060 nm.

3.4.6c Resolidification Theory

Given that to a greater or lesser degree the surface is melted by the increased E field due to interference effects the question as to why the surface as it solidifies does not return to its original profile must be asked.

Some insight as to the answer to this problem can be gained by reference to the work of Emmony et al³⁵ in which the shape of a surface upon resolidification is determined.

In their work these authors showed that if a section of a metal was melted then, under certain conditions it would not resolidify to its original profile.

The authors used a two dimensional model based ultimately on the fact that the solid and liquid phase of the materials used, occupy different volumes

due to their different densities (ρ_s and ρ_l for the solid and liquid respectively) and that surface tension was the dominant restoring force on the liquid surface.

The assumption of the dominance of surface tension forces over gravity is justified since the coefficient of surface tension for liquid metals is typically ten times that of water. They then argued that for a pool of liquid metal whose cross section forms an arc of a circle in an otherwise solid structure, the solid liquid interface will always be instantaneously semicircular (as the model is two dimensional) so as to minimize the free surface energy. It was further assumed that the main form of heat loss is by conduction to the solid.

An expression was then derived for the deviation (normalised to the original surface) of the surface profile of the melted area (assumed to initially form an arc of a circle of depth h_0) as a function of the fraction of the melt resolidified.

Fig 3.16 shows the main details for this theory. The functions $f_s(t)$ and $f_l(t)$ represent respectively the profile of the solid and liquid parts of the melt as it resolidifies.

Defining $\xi = t/t_0$ where t_0 is the radius of the initially molten pool and t that when some fraction has resolidified and ξ^* to be the position at which the surface makes the transition from solid to liquid where ξ^* is obtained from the relation

$$(\Delta/h_0)(3\eta-1)/2\eta = 1-\xi^*(3\eta-1) \quad 31)$$

where Δ/h_0 is the normalised incremental solidification layer thickness and $\eta = \rho_s/\rho_l$.

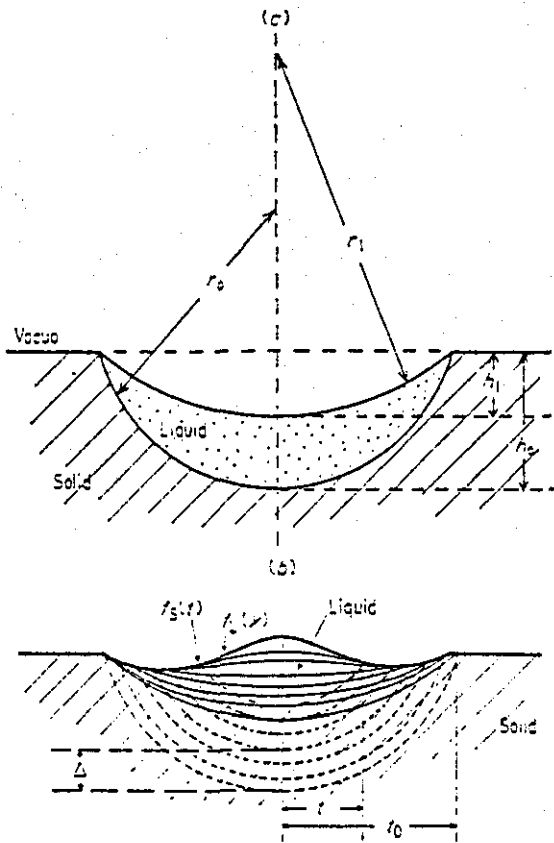
The normalised deviation from the original surface is then given by

$$f/h_0-1 = \begin{array}{ll} f_l(\xi) - 1 & \text{for } \xi \text{ between } 0 \text{ and } \xi^* \\ f_s(\xi) - 1 & \text{for } \xi \text{ between } \xi^* \text{ and } 1 \end{array} \quad 32)$$

where

$$f_l(\xi) = \xi^2 - 2\eta(\xi^*(3\eta-1)-1)/(3\eta-1) + (\xi^2-\xi^{*2})(1-\eta\xi^{3(\eta-1)}) \quad 33)$$

$$f_s(\xi) = \xi^2 - 2\eta(\xi(3\eta-1)-1)/(3\eta-1) \quad 34)$$



A

B

Fig 3.16

Geometry for resolidification calculations from ref 35.

Above top a) Section through the initially molten pool, assumed to form an arc of a circle

Above b) Section through partially resolidified system. The dotted lines represent the solid liquid interface which is always semicircular.

To perform the calculations, the left hand side of 31) i.e. Δ/h_0 was given ten increments between 0 and 1 which then determined ξ^* . This was then used in 32) to determine the new surface profile. A computer program was written that performed this calculation for both Ge ($\rho_s = 5200 \text{ kgm}^{-3}$, $\rho_l = 5600 \text{ kgm}^{-3}$) and Al ($\rho_s = 2700 \text{ kgm}^{-3}$, $\rho_l = 2400 \text{ kgm}^{-3}$).

3.4.6d Energy of Formation Theory

It is generally accepted that the formation of the ripples occurs on resolidification of the surface as was first suggested by Willis and Emmony³⁶ and later, and more formally, in III and by van Driel³¹. It is the contention of van Driel and the author that it is this fact that accounts for the observation of LIPSS on samples irradiated by short (picosecond) or long (nanosecond) laser pulses since, formation of ripples under this assumption, requires only that some degree of surface melting occurs and is independent of the processes that actually bring the surface to melting.

On this premiss, it is the absorbed fluence that is the important parameter as it is this that determines the degree of melting (if any) that is present on the surface and hence, the author has calculated a "normalised fluence of formation" as now detailed.

If the formation of LIPSS is dependent only on reaching some degree of surface melting then when allowance is made for polarization, the variation with angle of incidence of the fraction of the light absorbed and the beam target interaction area then the formation fluence should be constant for a given material.

As LIPSS are formed at lower fluences than the LICW, the reflectivity of a sample and thus the absorption of incident light is assumed to be given by the Fresnel equations with optical constants characteristic of the solid state being used in the calculations.

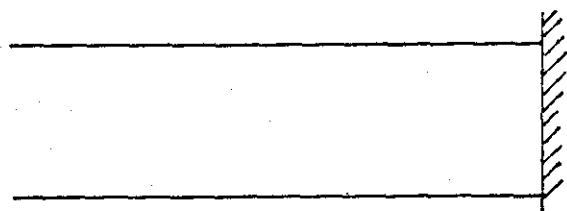
As the angle of incidence increases so does the beam target interaction area as the light is now "spread" over a larger fraction of the surface of the sample and if (fig 3.17) at normal incidence the area is $a(0)$ whilst at angle θ the area is $a(\theta)$ then

$$a(\theta) = a(0)/\cos\theta \quad 35)$$

If the fluence of formation at angle θ is $f(\theta)$ and as the fraction of this absorbed is $1-R(\theta)$ where $R(\theta)$ is the reflectivity at angle θ then the absorbed fluence at angle θ is

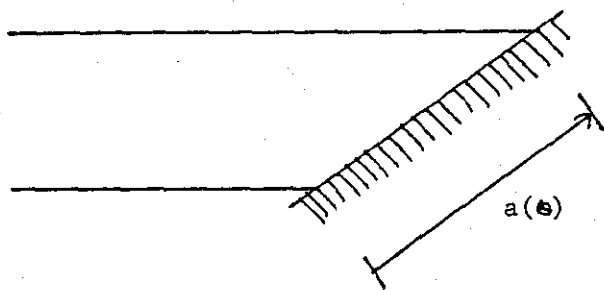
$$f_{ab}(\theta) = f(\theta)(1-R(\theta)) \quad 36)$$

Fig 3.17



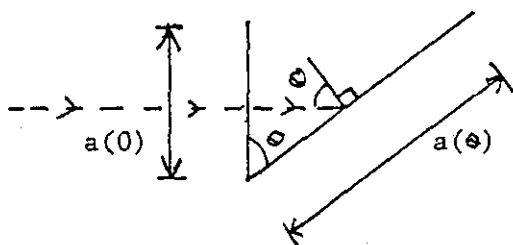
$a(0)$

Normal incidence



$a(\theta)$

Angular incidence



$$a(0) = a(\theta) \cos \theta$$

The fraction absorbed at normal incidence is thus $1-R(0)$ and hence the "expected" fluence of formation at normal incidence is

$$f(0) = f(\theta)\cos\theta(1-R(\theta))/(1-R(0)) \quad 37)$$

Hence if the basic premiss is correct then for a given material when the appropriate values of R and the experimental values of $f(\theta)$ are used, $f(0)$ should be constant and independent of the angle of incidence and polarization of the incident light. A computer program was written to evaluate 37) with the optical constants being taken from table 3.2 and the formation fluences from table 3.4. The results of these calculations are in table 3.4 under the heading of normalised formation fluences.

3.4.7 Comparison of Theory and Experiment

3.4.7a - Spacing and Orientation

The dependence of these ripples on the angle of incidence and polarization of the incident light justifies their LIPSS attribution.

Figs 3.18-20 are the efficacy factors for selected angles of incidence and polarization of the incident light chosen to correspond to the pictures of ripples in figs 3.11 and 3.13-15.

Within each graph, η is plotted for two values of angle ϕ . By definition, the peaks in the graphs of η at $\phi = 0^\circ$ (corresponding to \underline{k} being parallel to \underline{k}_p) represent the \underline{k} values of the s and p type fringes for S and P polarized light respectively. Similarly the peaks in the graphs of η for $\phi = 90^\circ$ (corresponding to \underline{k} being perpendicular to \underline{k}_p) represent the \underline{k} values of the c type fringes.

By taking the value of \underline{k} that produces the peak in η for each value of ϕ in graphs such as these, table 3.5 was constructed and shows the experimental and theoretical fringe spacings and orientations in terms of \underline{k} for all experimentally observed s^- , p^- and c type fringes. Also included in table 3.5 are the values of \underline{k} at which the "surface scattered wave" theory predicts LIPSS.

From table 3.5 it is apparent that the values of \underline{k} found experimentally and those predicted by the efficacy factor are in very good agreement for all the observed fringes. The scattered wave model cannot account for the c type fringes at all and it predicts spacings far too big (i.e. \underline{k} too small) for p polarized light at large ($>60^\circ$) angles of incidence.

From the graphs of η it can be seen that for both s and p type fringes there are two peaks in the graph one at small \underline{k} corresponding to the s^- and

Table 3.5

EXPERIMENTAL

THEORETICAL

<u>Angle</u>	<u>Orientation*</u>	<u>k** measured</u>	<u>Angle ϕ</u>	<u>k*** calculated</u>	<u>k// scattered wave</u>
<u>Ge P polarisation</u>					
75	perpendicular	0.32	90	0.32	--###
	parallel	0.08	0	0.07	0.034
	" ###	0.06	0	0.06	0.034
60	"	0.16	0	0.17	0.14
45	"	0.31	0	0.32	0.29
30	"	0.50	0	0.54	0.50
<u>Ge S polarisation</u>					
70	"	0.07	0	0.06	0.06
60	"	0.13	0	0.14	0.14
<u>Al P polarisation</u>					
75	"	0.11	0	0.10	0.034
60	"	0.18	0	0.20	0.14
<u>Al S polarisation</u>					
45	"	0.31	0	0.31	0.31
30	"	0.51	0	0.50	0.50

- * orientation measured relative to k_p
- ** all k values normalised to incident wavevector
- *** quoted value corresponds to the peak of the $\eta(k)$ v. k graph for the stated ϕ
- // spacings as predicted by the simple "surface scattered wave" model
- ### simple model does not predict fringes in this direction
- #### fringes on liquid surface with liquid dielectric constants

p^- type fringes that are observed experimentally and the other corresponds to s^+ and p^+ type fringes where the spacing is less than the incident wavelength i.e. $k > 1$. It was as discussed in 3.4.5 impossible to determine the orientation of the small ripple structures that the TEM revealed. Typical of these patterns are figs 3.21a and b where the spacing is $\approx 200\text{nm}$ i.e. $k \approx 1.2$. Essentially no fringes with the predicted s^+ and p^+ spacings were found by the author but, many other workers (particularly II) have reported p^+ ripples and thus, it was not possible to compare theory and experiment for this type of ripple. The lack of these types of ripple could be due to one or more of several possible causes such as...

1) The extreme difficulty of observing any subwavelength ripples due to their very shallow profile and narrow spacing with the consequence that the samples have to be tilted at quite large angles in the TEM to observe anything at all, thereby reducing the likelihood of seeing the ripples even if they are present.

2) A different threshold formation fluence for the smaller fringes than for the larger fringes.

3) A tendency during the formation of the larger spaced ripples to destroy (wash out) the smaller structures.

4) The influence of feedback on the size of the formed fringes. It is not inconceivable that once feedback is allowed for in the theory it would show that the larger spaced structures are more likely to occur than the smaller ripples.

Reference to figs 3.18-20 will show that the uniform fine c type fringes are predicted for p polarized light only, as confirmed by experiment. The absence of c type ripples on Al despite an appropriate peak in η is probably explained by feedback which undoubtedly occurs in practice but is lacking in the theory used. The fact that these fringes were only observed at large angles is explained at least in part by their size which, makes detection at 60° or less essentially impossible by OM.

To facilitate a comparison, fig 3.19a is the efficacy factor for Ge at 249 nm whilst fig 3.19b is that in the infra-red at $1.06 \mu\text{m}$. It is apparent that the transition from behaving optically as a dielectric at $1.06 \mu\text{m}$ to a metal at 249 nm allows new surface modes (which are essentially surface plasmons) that have definite peaks in their efficacy factor and result in the readily observed LIPSS with S polarized light.

Though the peaks in the graphs of η for S polarized light incident on Al are only very slight, the fringes are strongly present in reality. This

efficacy factor...Ge...P pol
wavelength...249.0...nm

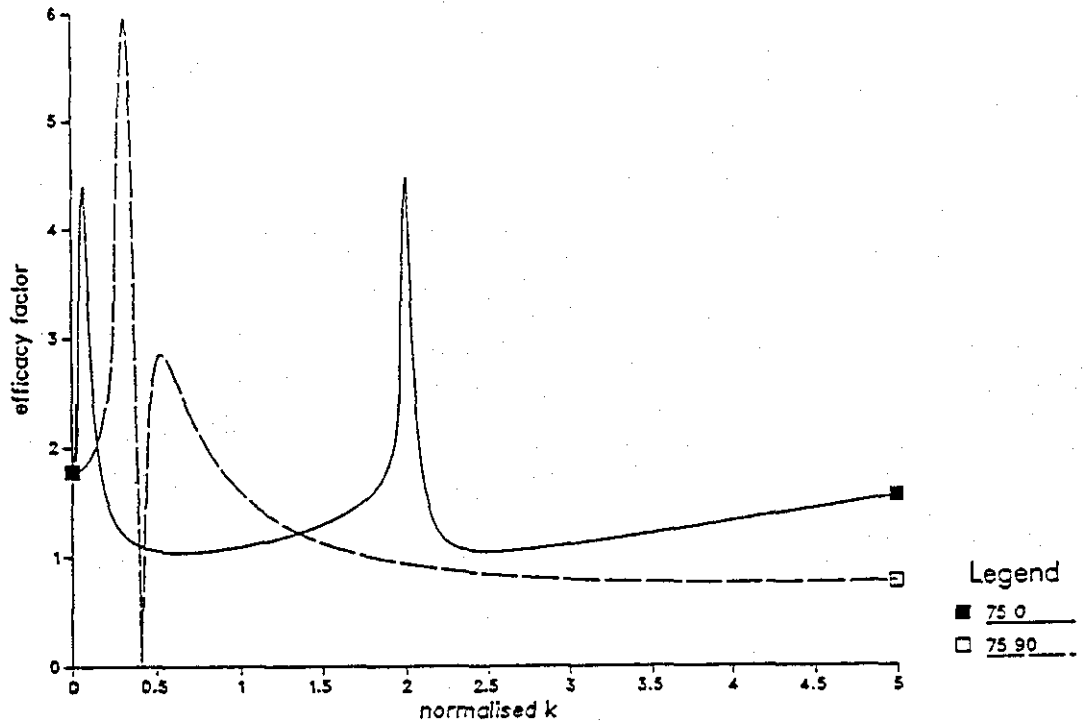


Fig 3.18

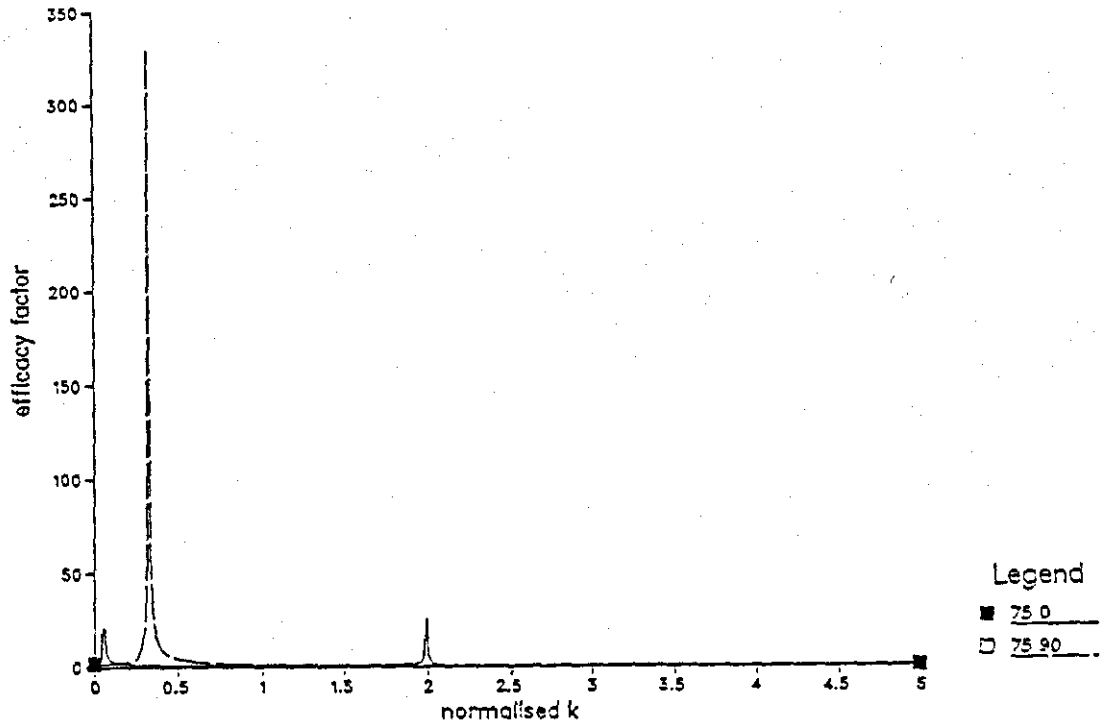
Efficacy factors as a function of normalised k and angle ϕ for Ge under P polarised 249 nm irradiation at 75° incidence.

Above a) Solid Ge with optical properties $n_t = 1.4 + i3.2$ and surface properties $s = 0.4, f = 0.1$

Overleaf b) Liquid Ge with optical properties $n_t = 0 + i5$.

The first number in the legend is the angle of incidence and the second the value of angle ϕ , both in degrees.

efficacy factor...Ge...P pol
wavelength...249.0...nm



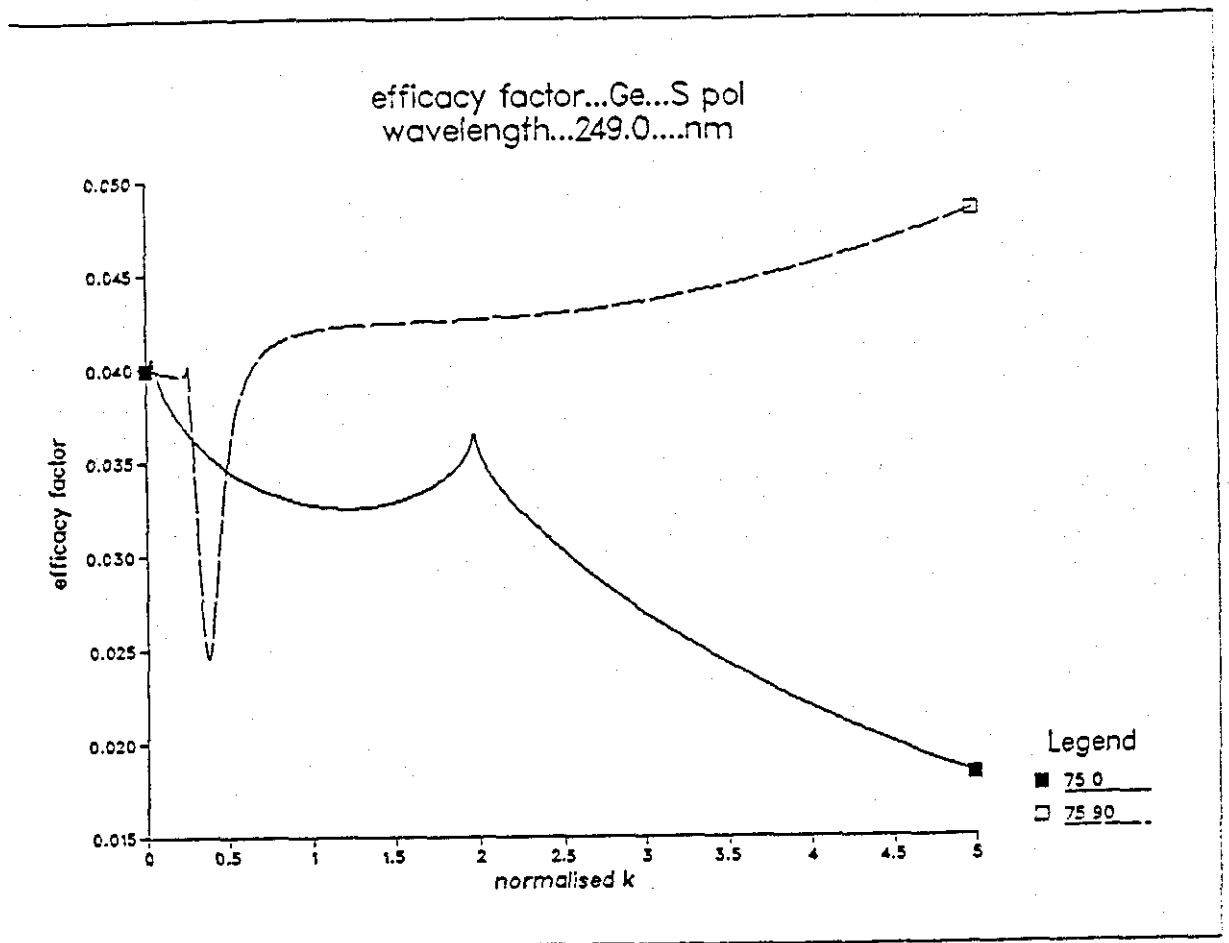


Fig 3.19

Efficacy factors as a function of normalised k and angle ϕ for Ge under S polarised irradiation at 70° incidence with surface properties $s = 0.4$ $f = 0.1$.

Above a) At 249 nm with $n_t = 1.4 + i3.2$

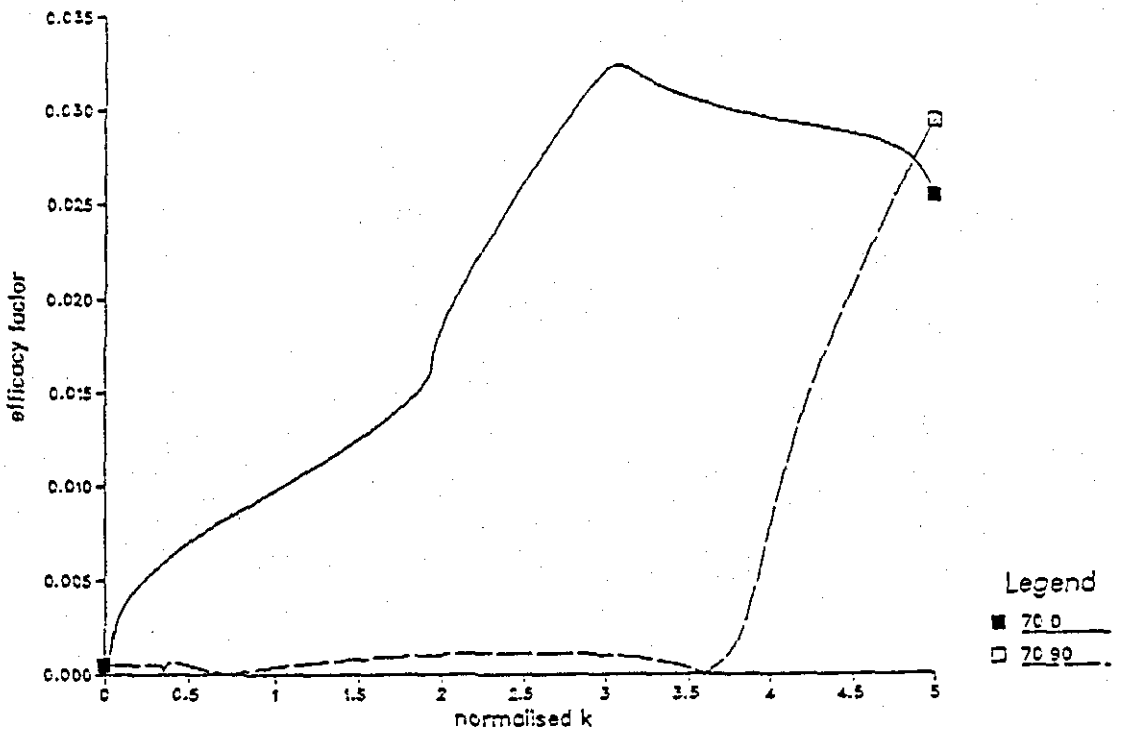
Overleaf Top b) At 1060 nm with $n_t = 0.1 + i4.0$

Overleaf Bottom c) At 249 nm with $n_t = 1.4 + i3.2$ and $\phi = 87^\circ$

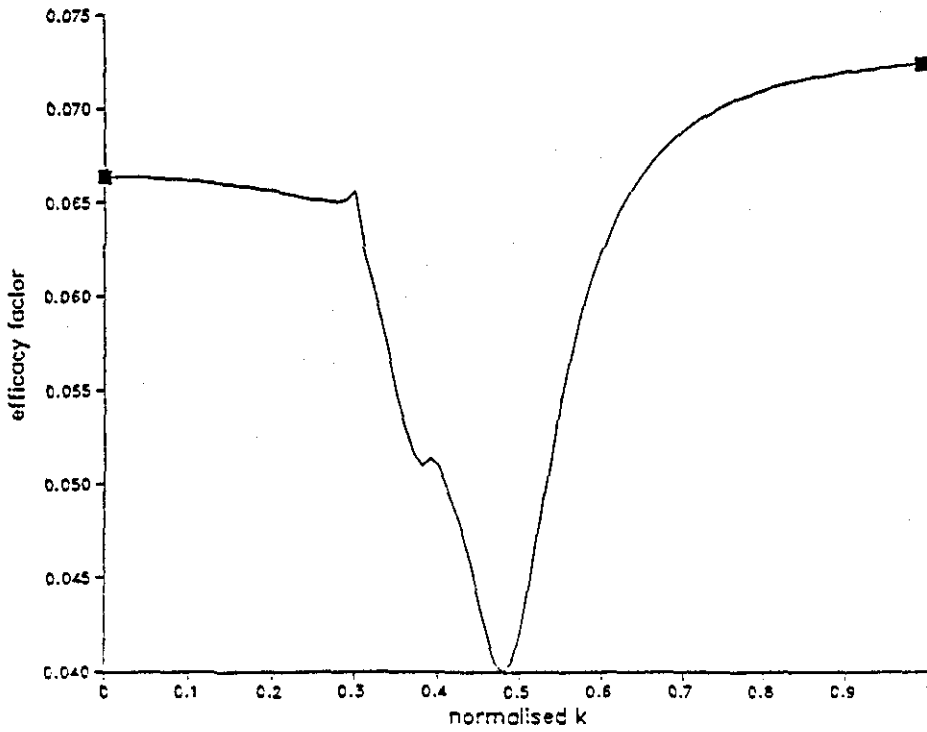
See fig 3.18 for explanation of the legend.

Note how the change to metallic behaviour (figs 3.19a-b) results in well defined peaks appearing in η . Also note the closely spaced double peak in fig 3.19c

efficacy factor...Ge...S pol
wavelength...1060.0....nm



efficacy factor...Ge...S pol
wavelength...249.0....nm



discrepancy and the absence on both Al and Ge of s^- fringes at certain angles is, once more, probably attributable to the role of feedback mechanisms in generating the ripples.

A careful examination of fig 3.14 reveals that the ripple spacing varies in the vertical direction being largest at the bottom of the picture which is nearer to the centre of the excimer beam than at the top where the spacing is somewhat less. This effect is attributed by the author to be due to the variation in the optical constants of Ge with temperature.

Since the morphology of the ripples (see 3.4.5.b) remains the same as the spacing varies, it is the author's contention that this means that all the ripples were formed by localised (inhomogeneous) melting (and subsequent resolidification) but with the surface E fields (selvedge and the normally refracted components) being initiated at different temperatures. Since Ge is a semiconductor it's optical constants will have a very strong temperature dependence, becoming more metallic as the temperature increases due to the increasing number of conduction band electrons. This variation in the optical constants manifests itself as a variation in the wavelength of the surface E fields which, when they interfere, results in a slight variation in the separation of the peaks and hence in ripple spacing. Calculations by the author of the efficacy factor of Ge as it's dielectric constant becomes more negative i.e. metallic, show that the predicted ripple spacing increases slightly in agreement with the above proposal.

The existence of two quite separate sets of ripples with different spacings and morphologies as evident from fig 3.11 is probably due to the formation of LIPSS on surfaces that have substantially different degrees of melting.

The narrower spaced ripples are formed away from the beam centre where the local fluence is lower and appear to have a narrow profile i.e. the surface appears to be sharply peaked in the vicinity of the ripples. On the other hand, the more widely spaced ripples are formed nearer the beam centre under higher fluences and have a very broad almost sinusoidal profile to them.

The spacing of the narrow ripples is consistent with that obtained using the solid state optical constants for Ge and their presence is thus attributed to LIPSS formation by localised (inhomogeneous) melting. The spacing of the larger ripples is consistent with that predicted by calculating the η factor for a liquid Ge surface (i.e. assuming homogeneous melting of the surface) using the optical constants determined from a free electron model (as discussed in 3.4.4b). A formation mechanism for ripples under the assumption of a liquid surface is discussed in the next section.

Whilst the use of a free electron model in the way done herein cannot be justified it has served the author's original intention, which was to show

that a large negative dielectric constant caused by heating the surface to or near melting can explain not only the existence of the two sets of ripple patterns but, accurately predict the spacings.

The origin of the well defined large spacing b type fringes that occur with S polarized light was initially unclear. Calculations of η for ϕ angles of $85-90^\circ$ (i.e. k almost perpendicular to kp) using different values of the shape (s), filling factors (f) and optical constants failed to reveal peaks at the extremely small k values that these widely spaced ripples correspond to. However, these calculations revealed that for $\phi \neq 90^\circ$ the efficacy factor has 2 closely spaced peaks.

Fig 3.19c shows η for S polarized light incident on Ge at 70° for $\phi = 87^\circ$, with peaks at 0.3 and $0.39k$. It is the author's suggestion that the "waves" corresponding to these peaks "interfere" (beat) with each other to give the observed ripples as now detailed. It is from the word beat that the author obtained the designation b type fringes.

Consider fig 3.19d which shows the positions on a surface at which individually the peaks in η would predict surface melting. The peaks at 0.3 and $0.39k$ correspond respectively to spacings of 0.83 and $0.638\mu\text{m}$ as shown in the diagram. If the "waves" start off at the same point, the location of the melting is usually different for each wave but, it is possible that after a sufficient number of periods (which will be different for each wave) the melt locations will once more overlap. In the case shown this occurs after 10 periods of the $0.3k$ wave i.e. $8.3\mu\text{m}$ which, corresponds exactly to 13 periods of the $0.39k$ wave. Hence every $8.3\mu\text{m}$ the amount of melting (to first order) will be doubled.

On a shot to shot basis it does not seem unreasonable to argue that due to this increased melting, more of the incident energy is coupled into these areas than into the areas where the melting is due to only one wave. Thus after multiple shot irradiation the most likely structures to have formed are those with a spacing corresponding to the distance between the points of double melting. Reference to fig 3.14 shows that the spacing of these b type fringes is $8\mu\text{m}$, in very good agreement with the above mechanism.

The generalised condition for the above to be a possible mechanism for generating LIPSS is simply that the values of k corresponding to the peaks in η can be written as the product of an integer and a factor common to both values. Under these conditions the predicted ripple spacing is given by the reciprocal of the common factor multiplied by the laser wavelength.

In the above case, the common factor is 0.03 with the integers being 10 and 13 and the predicted spacing being $1/0.03 * 0.249\mu\text{m}$ i.e. $8.3\mu\text{m}$.

Performing the same calculations for Al again revealed a double peak in η

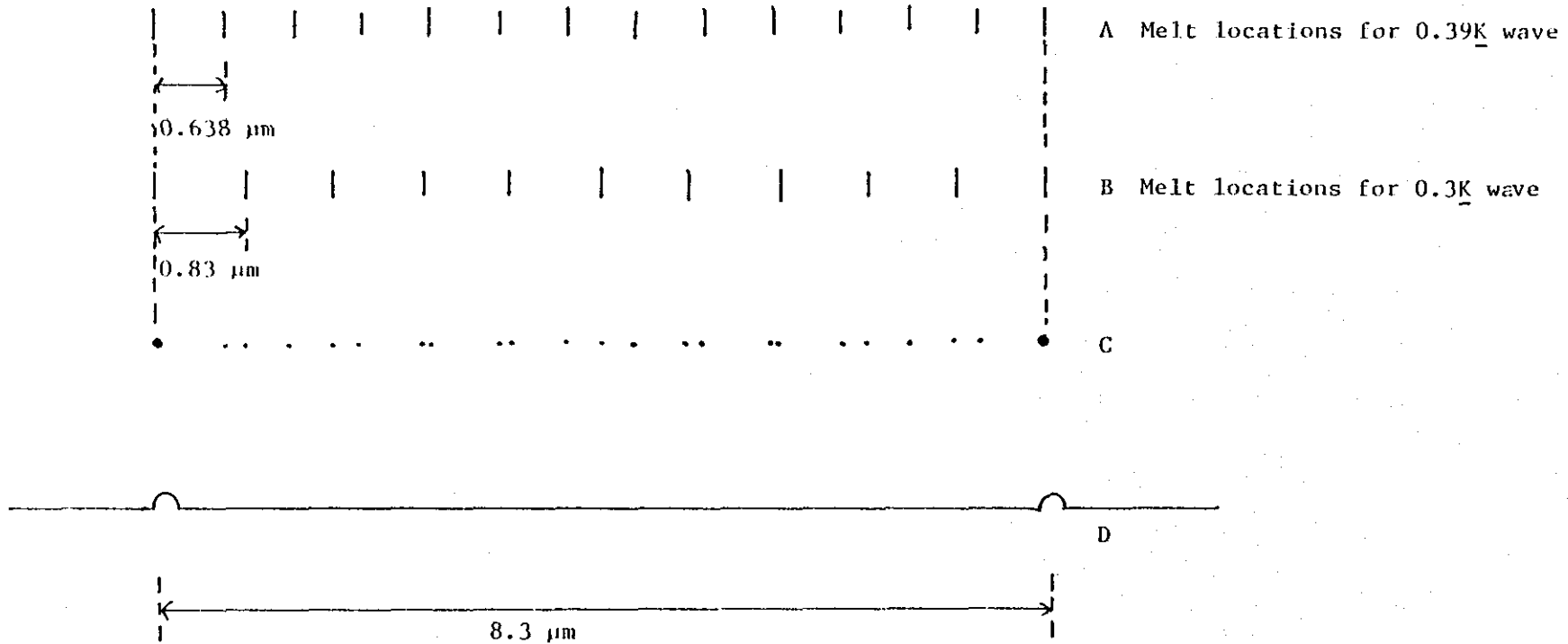


Fig 3.19d

C) Dots indicate amount of surface melting. Note double melting at the start and when the melt locations of the waves overlap.

D) After multiple shot irradiation, ripples form at points of double melting.

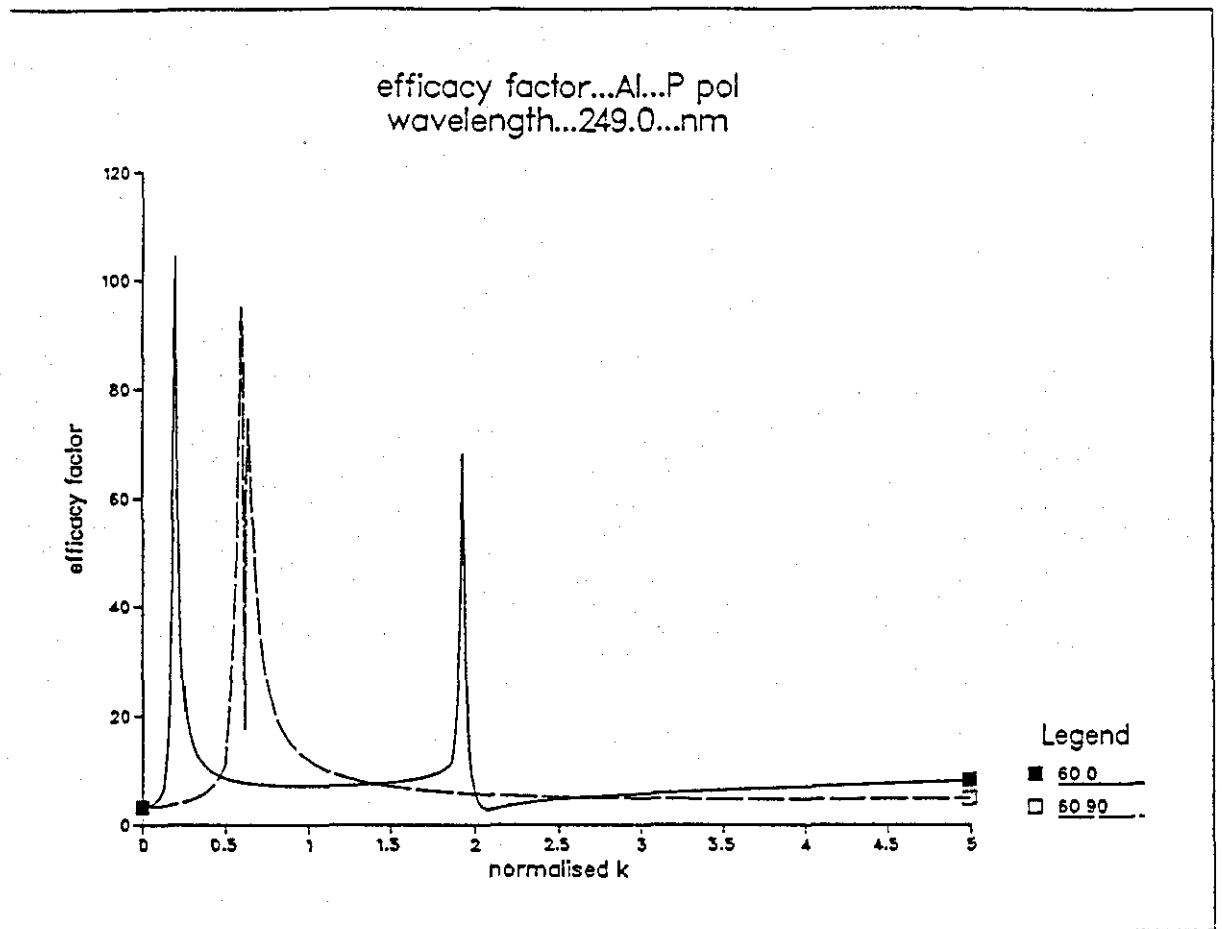


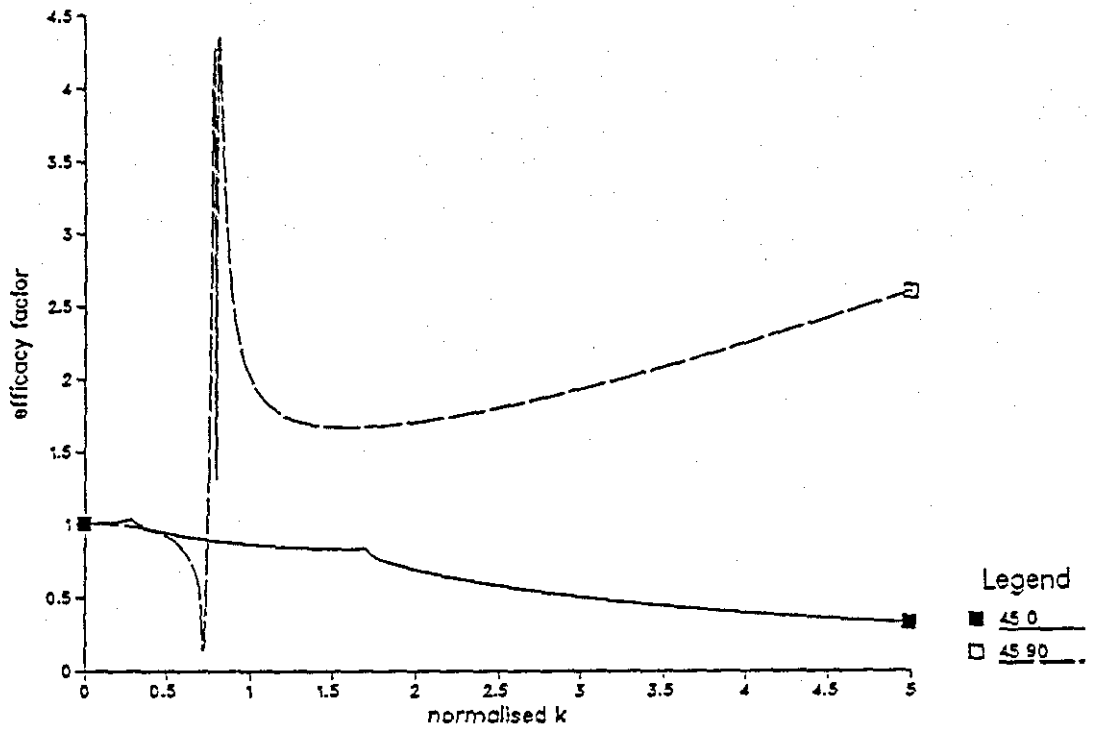
Fig 3.20

Efficacy factor as a function of normalised k and angle ϕ for Al under 249 nm irradiation with optical properties $n_t = 0.2 + i 2.9$ and surface properties $s = 0.4$ $f = 0.1$.

Above a) P polarised, 60° incidence
 Overleaf b) S polarised, 45° incidence

See fig 3.18 for explanation of the legend.

efficacy factor...Al...S pol
wavelength...249.0...nm



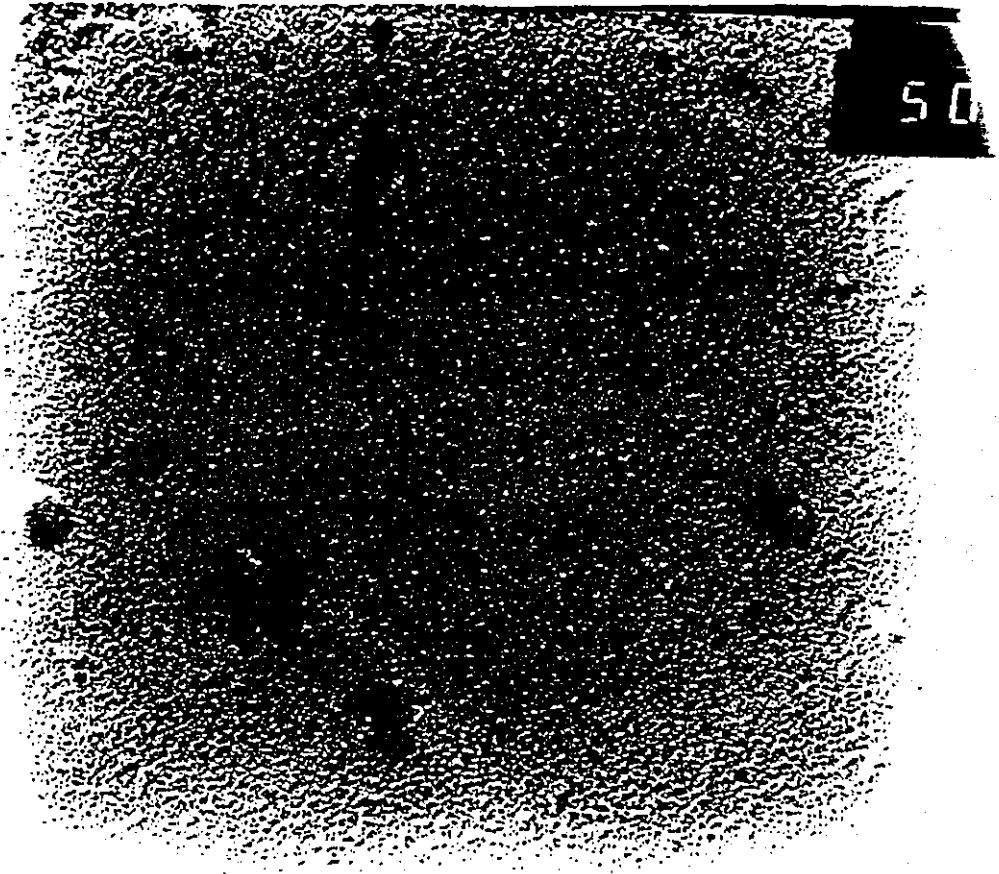
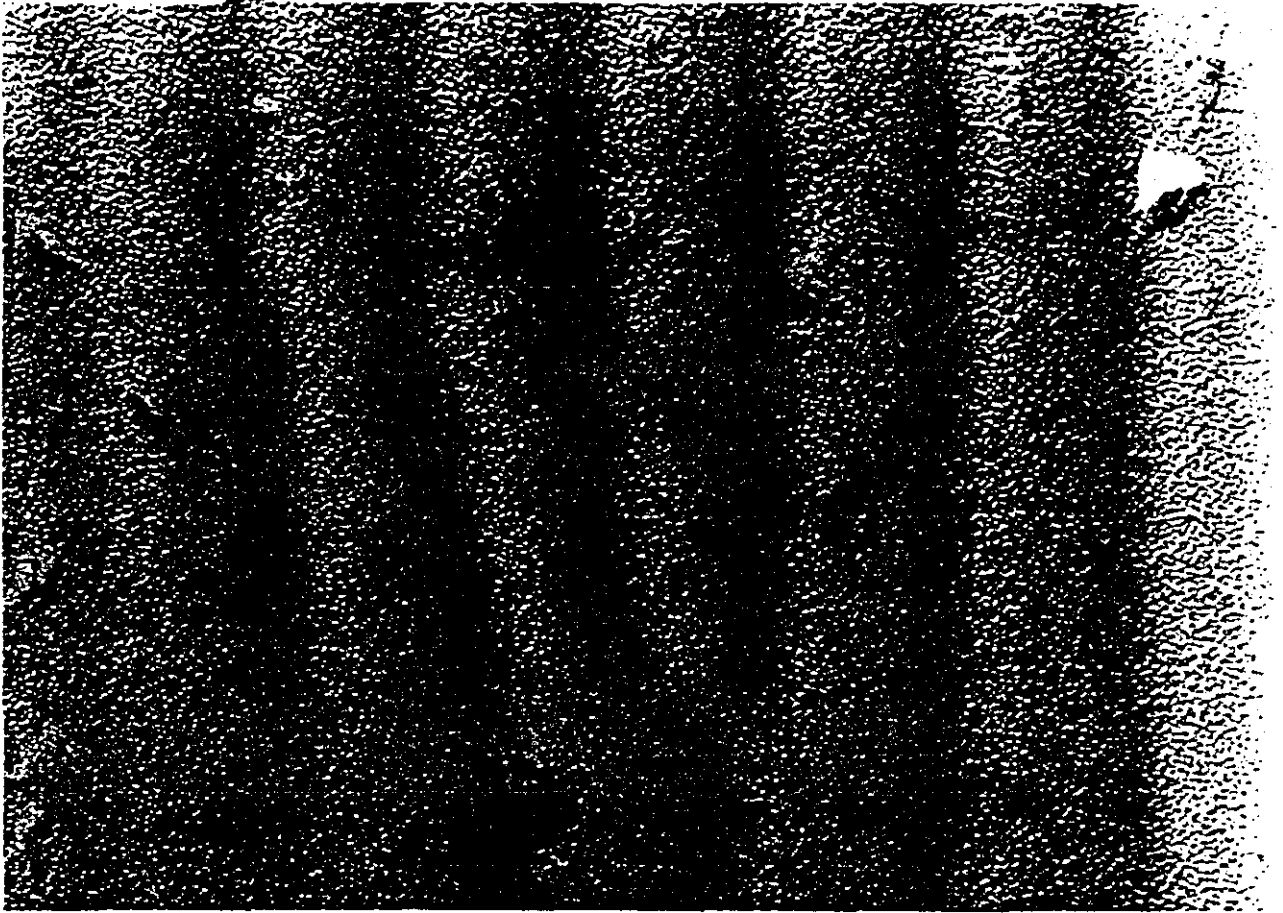


Fig 3.21

TEM pictures of sub wavelength ripples.

Above a) For Al S polarised light incident at 30° magnification X100,000
Overleaf b) For Ge S polarised light incident at 70° magnification X100,000

Note the fine grain structure which is due to the replication process.



which for $\phi = 86^\circ$ occurred at 0.75 and 0.87 μ m which, if slight allowances for the approximate nature of the calculation are made is almost exactly consistent with the peaks being related by a common factor of 0.125 i.e. a spacing of 2 μ m. Reference to fig 3.15b shows the spacing of the experimentally obtained ripples to be 2 μ m, giving further support to the proposed mechanism.

Overall from table 3.5 and the above it is evident that there is good agreement between theory and experiment which, if slight allowances are made for all the approximations used in the theory shows that the theory developed in I is essentially valid.

3.4.7b Ripple Morphology

Attempts were made to obtain pictures of the cross sectional profile of the ripples so that a comparison between ripples formed under different degrees of surface melting could be made. Essentially due to the polycrystalline nature of the samples and the limitations of the TEM these attempts failed and so the discussion in this section has to rely on inferences from the "face on" ripple pictures and the work by van Driel³¹ and Emmony et al³⁵.

Fig 3.11b shows quite clearly that the morphology of the large and small ripples is different as the former produces what appears to be a flat slowly varying ripple profile whereas the latter produces sharply peaked ripples. Figs 3.22a and b are the resolidification profiles calculated using the theory in 3.4.6c for Ge and Al. These quite clearly show that if an isolated area of the surface is melted it can resolidify to a different profile to that which it had originally.

By assuming melting only in the vicinity of the peaks in the surface E field i.e. each melt pool is totally separate from all others and that upon resolidification the surface takes up the profile in fig 3.22a, fig 3.23a shows the expected cross sectional profiles for ripples on Ge generated under conditions of localised melting and fig 3.23b shows that for Ge obtained experimentally by van Driel.

The agreement is excellent, in particular note the dip in the surface either side of the peak that is present experimentally and reproduced by the theory. It is the author's view that this agreement between theory and experiment (which is a more rigorous version of that performed in II, III and by van Driel) is sufficient to justify attributing the formation of these ripples to "localised melting".

The larger ripples (fig 3.11b) are attributed by the author to LIPSS formation on a molten surface as now detailed. Consider a Ge surface irradiated by a laser such that interference between the incident and

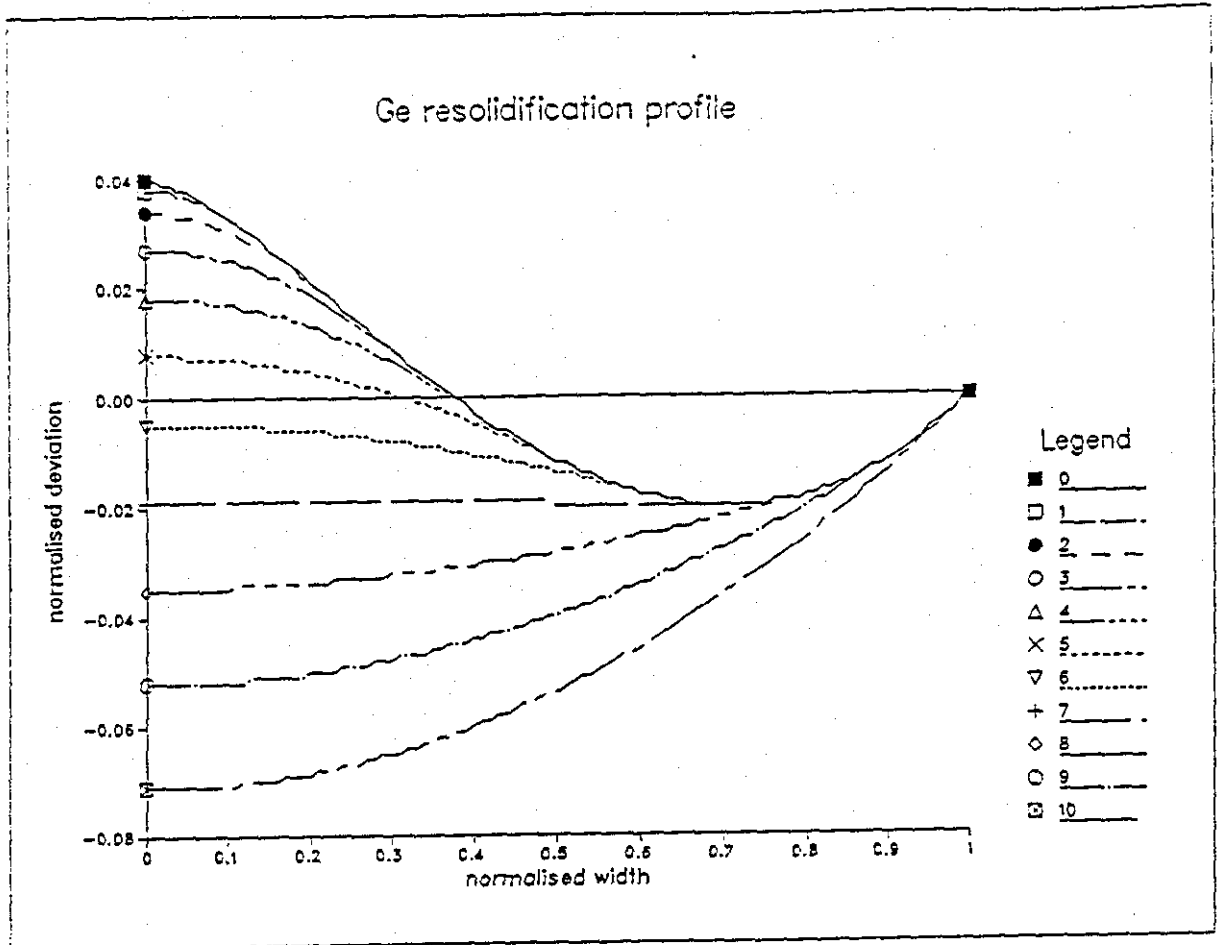


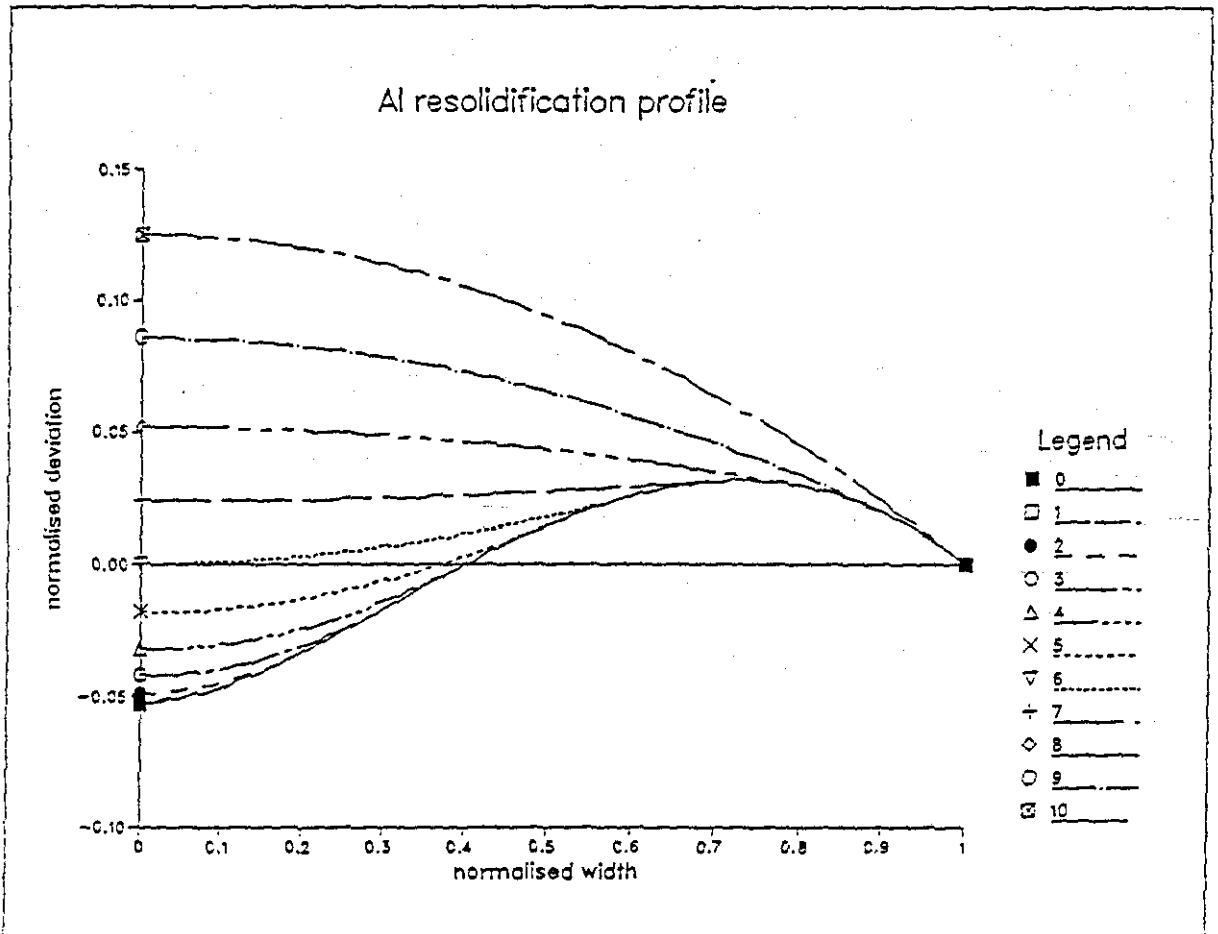
Fig 3.22

Computed deviation of the re-solidifying surface from the original flat profile as re-solidification proceeds. The vertical axis is a measure of the deviation from the original surface whilst the horizontal axis corresponds to the position across the solidifying pool.

The numbers in the legend correspond to the fraction, in tenths that is molten and hence 10 is completely molten and 0 is totally solid.

Above a) For Ge.
Overleaf b) For Al.

Al resolidification profile



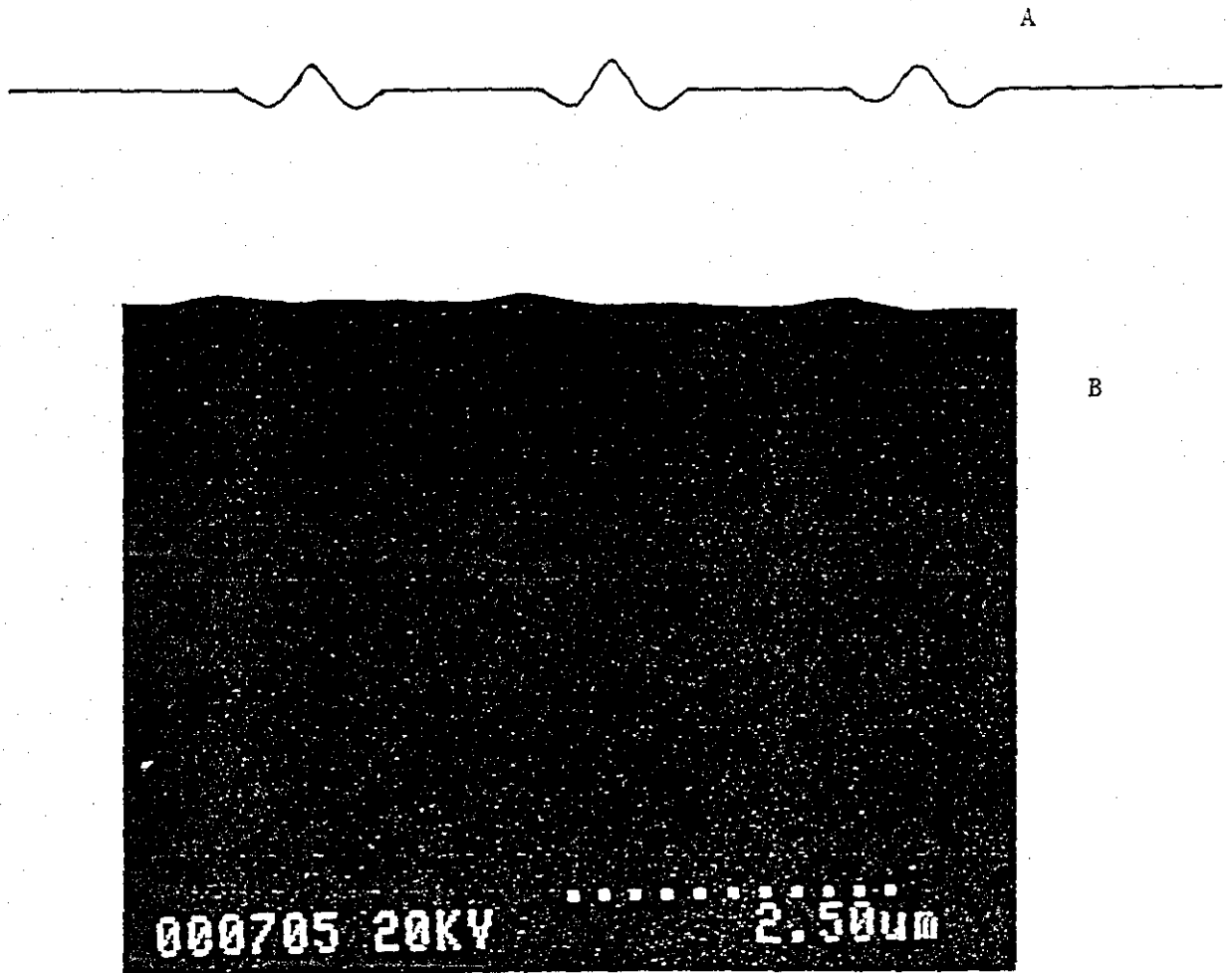


Fig 3.23

Resolidification profiles under localised melting conditions.

Above top a) Calculated profile using the shape from fig 3.22a.

Above b) Experimentally obtained profile by van Driel

induced E fields leads to a periodic temperature distribution on the surface. If the pulse fluence is high enough the surface will melt uniformly (i.e. to cause the localised melting in the vicinity of the E field maxima to spread out until the melt from adjacent maxima overlap at which point the entire surface will be melted) and take up an essentially flat profile. Due to the periodic temperature profile not only will the surface have "hot" and "cool" regions but, the melt depth will vary across the surface as shown in fig 3.24a.

Upon resolidification, the surface profile will remain flat until the point shown in 3,24b where the coolest points on the surface itself have frozen. After this point has been reached, the surface which is now effectively "pinned" by the solidified coolest parts must ripple upwards in order to conserve mass since, the density of Ge decreases upon solidification. The result of this is to leave the surface permanently rippled with a smooth almost sinusoidal profile as shown in 3.24c.

Without experiment cross sections of these high fluence ripples a proper test of the above hypothesis and attributing their formation to a "uniform melting" mechanism cannot be totally justified but, this mechanism at least is consistent with experimental observations to date.

3.4.7c Formation Fluence

The "normalised formation fluence" i.e. the fluence to form ripples at normal incidence for all s",p" and c type fringes is listed in table 3.4 where two clear trends can be seen.

Firstly there is a substantial difference between Al and Ge which ,at least in part will be due to their difference in normal incidence reflectivity and thermal properties.

Secondly and more difficult to explain is the fact that for each material there is agreement of the values within either S or P polarizations but not between them with, the formation fluence for S being lower than that for P particularly for Ge. It would not appear unreasonable to suggest that this difference is due in some way to a difference in the coupling of the light into the selvedge but, beyond this tentative statement the author has no explanation of this effect.

The good agreement within each polarization of the normalised formation fluence in the author's view confirms the basic idea that it is the absorbed fluence that is the important parameter in LIPSS formation and also gives support to the generally accepted idea that the ripples are formed on resolidification.

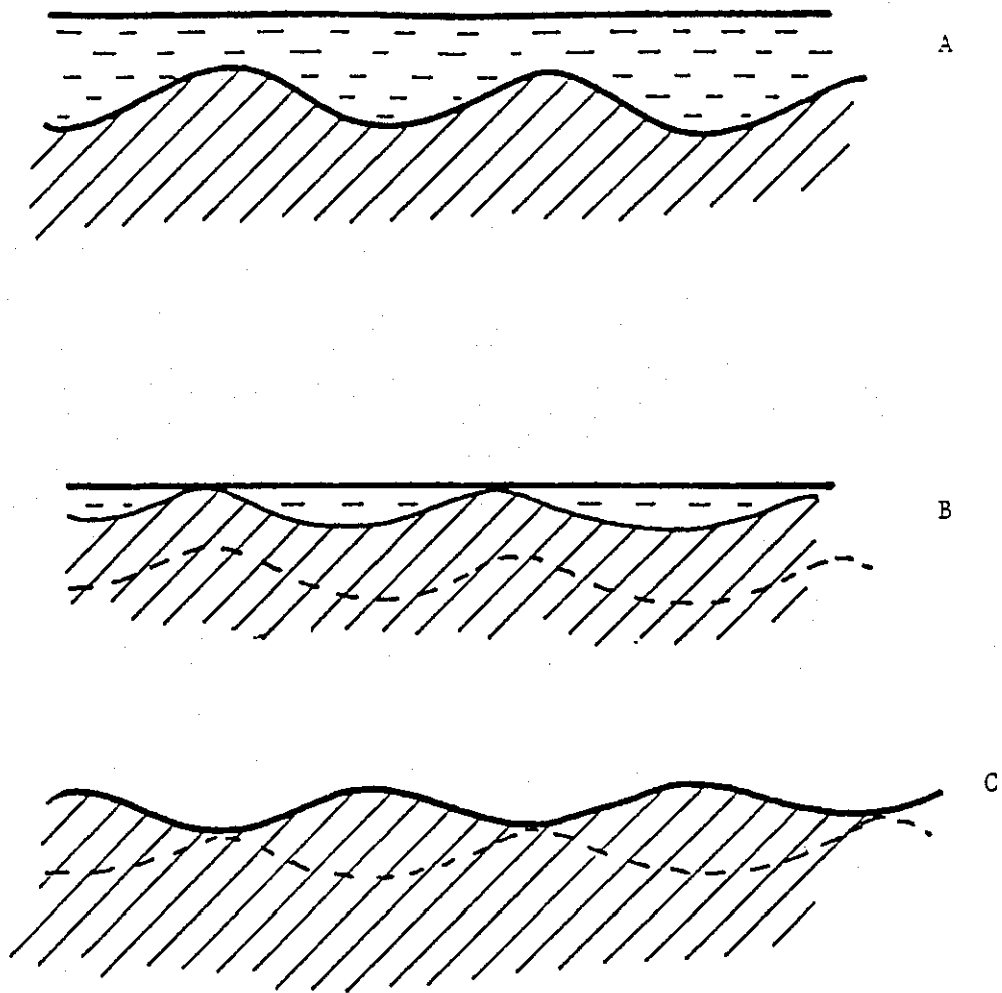


Fig 3.24

Proposed ripple formation mechanism from a liquid surface..

- a) Fully molten surface. Note how the depth of melting varies, following the periodic temperature profile on the surface.
- b) Coolest points just solidified and thereby "pin" the remaining molten material.
- c) Surface ripples to conserve mass which is upwards for Ge due to the decrease of density on solidification.

3.4.8 Conclusions

Good agreement between experiment and theory for the spacing, orientation and morphology of the ripples has been shown. Most of the discrepancies in the spacing and orientation of the ripples are probably linked to the role of feedback and the details of the damage mechanism for each material. The premiss that ripple formation depends upon reaching some degree of surface melting has been studied and it has been shown that the formation fluence is independent of angle of incidence but dependant on the polarization used. The formation of ripples on the Al samples which were of low optical finish shows that the degree of surface finish is at least to first order relatively unimportant.

3.5 ALIPSS

3.5.1 Introduction and Review

The work reported in the literature on LIPSS has been concerned essentially solely with those structures that had their spacings based on the normal (vacuum) wavelength on the incident laser beam. To the author's knowledge the only report of anomalous structures was that by Ursu et al³³ where ripples of large spacing with no obvious dependence on the wavelength or angle of incidence of the light used were reported. The fluences used to generate these ripples was such that they produced a large plasma above the sample's surface on every shot incident on the sample. It is the author's contention that at least in part these anomalous ripples can be explained by the presence of the plasma.

At these fluences, there will be considerable ablation of material and thus a considerable pressure wave on the surface due to the rapidly expanding volume of plasma both of which are possible "seeding" mechanisms for capillary waves. However, since the plasma will be generated very early in the pulse, the effect that it might have on the remaining light before it reaches the surface cannot be ignored.

It is the author's contention that under certain conditions, the plasma immediately above the surface can effectively alter the wavelength of the incident light as seen by the surface.

3.5.2 Propagation in a Plasma - Theory

The plasma above a surface will be formed by vaporization and subsequent ionization of material from the surface. Due to their large mass, the presence of (and therefore any effects from) the generated positive ions is ignored and only the generated electrons are considered. It is further

assumed that the fluences, plasma temperature etc are sufficiently low that the plasma can be treated as essentially a free electron gas. On this basis the dielectric constant and hence refractive index of the plasma reduces to that given by 30). It is apparent from 30) that for a given frequency, refractive indices n (where $n = \epsilon^{1/2}$) of less than 1 are possible.

Consider the situation where light of wavelength λ_a and thus frequency f_a (in air) is incident on a region of plasma immediately above a sample's surface. To satisfy the boundary conditions the frequency must be the same in both media and hence the wavelength in the plasma λ_p is given by..

$$\lambda_p = \lambda_a/n_p \quad (38)$$

where n_p is the refractive index of the plasma. Hence if $n_p < 1$ the wavelength in the plasma and thus that seen by the sample's surface is longer than that of the original light.

3.5.3 Experimental Results

In the course of attempting to generate LIPSS the author performed the same experiments as those in 3.4. but at fluences 2,3 or more times higher than the formation fluences in table 3.4.

The results under these conditions were badly damaged sites with a variety of complex structures of various orientations (fig 3.13) and spacings varying from less than 1 to more than 20 μm . In general these structures showed no simple dependence of their morphology, spacing and orientation on either the fluence used or the angle of incidence and polarization of the incident light. However, for certain fluences and numbers of shots structures whose orientation and dependence on angle of incidence was identical to that of p^- and c type fringes but, with a much larger spacing such as those in fig 3.25 were observed.

This picture (fig 3.25) of the structures generated on an Al sample by P polarized light incident at 60° of fluence 3.5 Jcm^{-2} shows two sets of ripples, one parallel to \underline{kp} labelled as p^{*-} type of spacing $10\mu\text{m}$ and the other perpendicular to \underline{kp} labelled as c^* type with a spacing of $\approx 3\mu\text{m}$ giving a ratio of 3.3 in the spacing of the two types of ripple. The morphology of the ripples is similar to that observed for the ripples formed on a liquid Ge surface (fig 3.11).

3.5.4 Analysis of Results

Calculation of the peaks in the efficacy factor for Al with P polarized light incident at 60° predict spacings of $1.25\mu\text{m}$ parallel to \underline{kp} and $0.39\mu\text{m}$

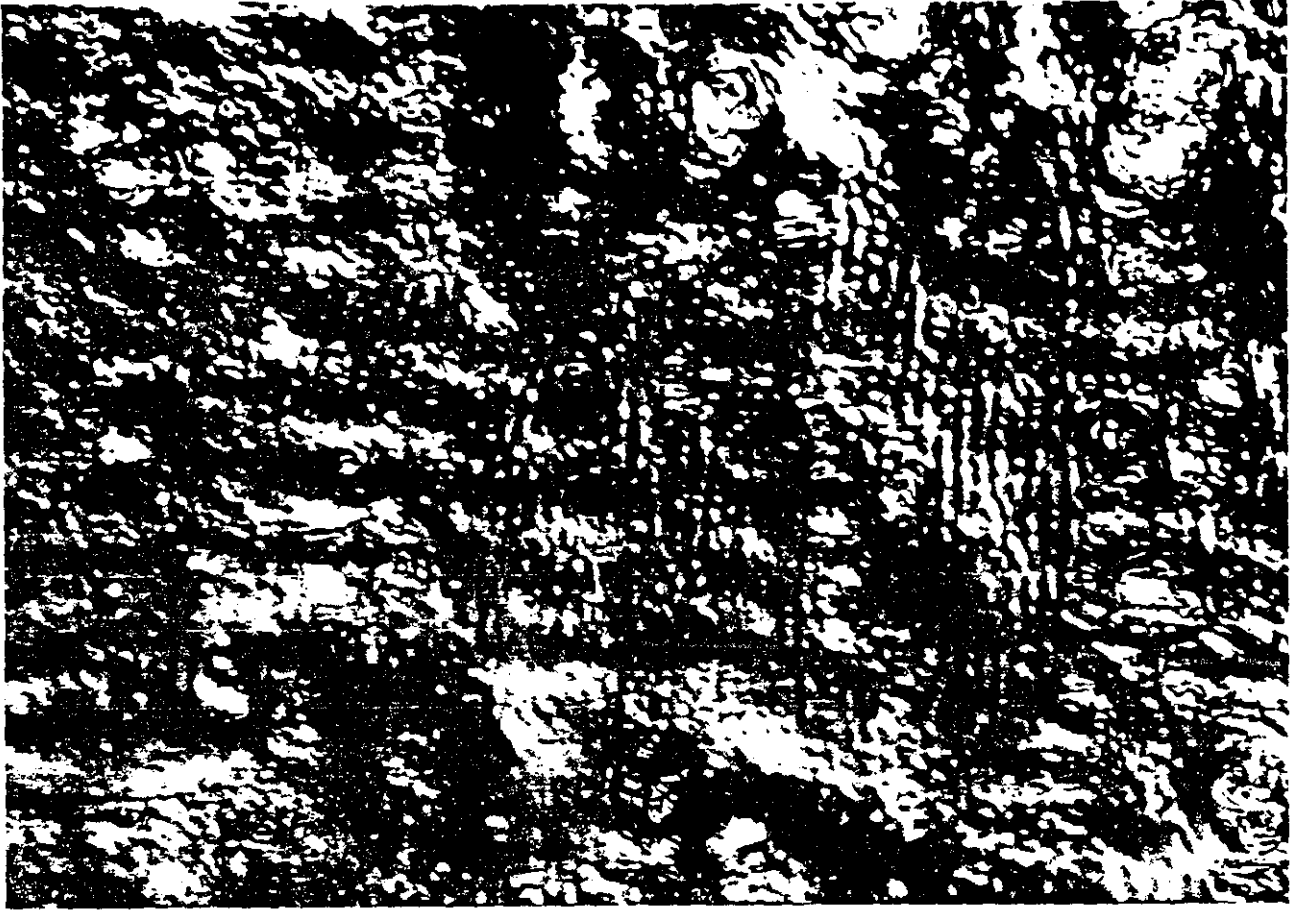


Fig 3.25

ALIPSS on Al P polarised light incident at 60° magnification X1250 Observe the p^* (horizontal with large spacing) and c^* (vertical and of smaller spacing) ripples. Ripple morphology similar to the normal LIPSS but the spacing is about 8 times too big.

perpendicular to k_p i.e. a ratio of 3.25 in the spacing of the two types.

Although the spacing of the experimental ripples is too large, it is too large by the essentially the same factor (times 8) for both sets of ripples with the ratio of their spacings being in good agreement with that predicted above.

In this case the sample had ripples over an area of order 0.4 by 0.1 mm. Assuming that the plasma that was produced by the laser pulse to be confined to an area of 0.4 by 0.1 mm and estimating its depth to be 10 μ m with the further assumption that 50% of the incident photons result in free electrons being generated then at the fluence used equation 30) yields a refractive index of 0.16 which, would result in the wavelength as seen by the surface being ≈ 6 times the vacuum value.

The purpose of this calculation was to show that at the fluences used and with "reasonable" estimates of the plasma volume and the fraction of photons that generate electrons, the plasma frequency ω_p and angular frequency ω of the incident light are nearly the same with the result that the refractive index of the plasma is much less than 1.

It is thus the author's contention that these observed ripples (p^* and c^* type) are LIPSS generated by a laser pulse whose wavelength as seen by the surface has been greatly increased due to passage through the plasma immediately above the surface and are thus designated as ALIPSS.

3.6 Comment

As all of the LIRS formed on non overcoated samples have a substantial depth to them it is the author's contention that the coating of Al on the non overcoated samples which is less than 100 nm thick is simply too thin to support any form of oscillation and hence shows no LICW/LIPSS.

3.7 Overall Conclusions

Three separate mechanisms mechanical buckling (LITS), capillary waves (LICW) and E field interference (LIPSS) for the generation of LIRS have been detailed.

Mechanical buckling was identified as the mechanism for ripple generation on overcoated (constrained) surfaces whilst on molten surfaces at least some of the ripples were shown to be attributable to induced capillary waves.

The work on LIPSS has shown that not only are such structures possible in the UV but that entirely new types of ripples can exist for S polarized light in the UV. A detailed comparison was made between theory (I) and

experiment which at least to first order validated the theory. Progress on understanding the morphology of the ripples was made, with two separate mechanisms being shown to generate ripples of differing morphologies but ones consistent with experimental observation.

The existence of ripples of anomalous spacing was shown together with a tentative explanation as to the origin of at least some of them.

There are inadequacies in both the theory used and experiments performed and whilst the former waits for an inspired theoretician, the author and colleagues here at Loughborough intend to redress the latter by studying the development of LIPSS both inter and intra shot (by use of a short pulse dye laser) and if equipment permits to perform time dependant diffraction studies.

References

- 1) S.S. Wiseall and D.C. Emmony, NBS Spec. Pub. 669 pp102 (1984)
- 2) "Laser Annealing of semiconductors" edited by J.M. Poate and J.W. Mayer, pub. by Academic Press 1982
- 3) J. Stephen et al in Laser and electron Beam Processing of Materials pp 639 pub. by Materials Research Society.
- 4) C. Hill and D.J. Godfrey, J. de Physique C4 suppl 5 41 79 (1980)
- 5) "Fluid mechanics" by L.D. Landau and E.m. Lifshitz, pub. by Pergamon Press 1986
- 6) "Vibrations and Waves in Physics" by I.G. Main, pub. by Cambridge universit press 1978
- 7) K. Rozniakowski, Mat. Res. Bul. 18 875 (1983)
- 8) F. Keilmann, Phys. Rev. Lett 51 2097 (1983)
- 9) J.F. Young, J.E. Sipe and H.M. van Driel, Phys. Rev B 30 2001 (1984)
- 10) "Effects of High Power Laser Radiation" by J.F. Ready, pub. by Academic Press 1971
- 11) "Heat Conduction in solids" 2nd ed by H.S. Carslaw and J.C. Jaegar, pub. by Clarendon Press 1986
- 12) J.A. Woodroffe et al, Appl. Phys. Lett 36 14 (1980)
- 13) Values taken from either "Tables of Physical and Chemical Constants" by G.W.C. Kaye and T.H. Laby, pub. by Longmans 1973 or from "The Handbook of Chemistry and Physics" edited by R.C. Weast, 62 ed pub. by CRC press 1982.
- 14) R. Bansal, J. Phys. C 6 3071 (1973)
- 15) D.C. Emmony et al, Appl. Phys. Lett 23 598 (1973)
- 16) Z. Guosheng et al, Phys. Rev. B 26 5366 (1982)
- 17) J.F. Young, J.S. Preston, H.M. van Driel and J.E. sipe, Phys. Rev. B 27 1155 (1983)
- 18) J.E. Sipe, J.F. Young, J.S. Preston and H.M. van Driel, Phys. Rev. B 27 1141 (1983)
- 19) M. Birbaum, J. Appl. Phys. 36 3688 (1965)
- 20) P.A. Temple and M.J. Soileau, NBS Spec. Pub. 462 p371 1976
- 21) " " " " , IEEE J. Quant. Elec 17 2067 (1981)
- 22) "Surface Polaritons" edited by V.M. Agranovich and D.L. Mills , pub. by North Holland 1982
- 23) M. Prokhorov et al , Sov. Tech. Phys. Lett 8 415 (1982)
- 24) V.I. Emel'yanov et al, Sov. J. Quant Electron 14 1515 (1984)
- 25) F. Keilmann and Y.H. Bai, Appl. Phys. A 29 9 (1982)
- 26) M.J. Soileau, IEEE J. quant. elec 20 464 (1984) and references therein
- 27) N. Mansour et al , NBS Spec. Pub. 727 (1986) p137 and references therein
- 28) D.J. Ehrlich and S.R.J. Brueck Appl. Phys. Lett 47 216 (1985)
- 29) P.M. Fauchet and A.E. Siegman, Appl. Phys. A 32 135 (1983)

- 30) D.J. Ehrlich et al, Appl. Phys. Lett 41 630 (1982)
- 31) H.M. van Driel, J.E. Sipe and J.F. Young, J. Lumin 30 446 (1985)
- 32) R.M. Osgood Jr. and D.J. Ehrlich, Opt. Lett. 7 385 (1982)
- 33) I. Ursu et al, J. Appl. Phys. 58 3909 (1985)
- 34) D.J. Ehrlich et al in MRS Symposia Proc. vol 13 p191 1982 and references therein
- 35) D.C. Emmony et al, J. Phys. D Appl. Phys. 8 1472 (1975)
- 36) L.J. Willis and D.C. Emmony, Optics and Laser Tec. October 1975 p222
- 37) Values for the optical constants taken from "Handbook of Optical Constants of Solids" edited by E.D. Palik, pub. by Academic Press 1985

Chapter 4

Early Detection of Laser Induced Damage

4.1) Introduction

By its very nature laser induced damage is a destructive phenomena and it is ironic that the practicalities of using high power lasers are such that a non destructive technique for the detection of the onset of LID before it becomes a "problem" would be of great use to many organisations. The early detection of laser induced damage (hereafter EDLID) is an idea whereby some form of "transient" non damaging effect associated with the interaction of the laser beam and the component in question is "monitored" so that predictions as to where and hopefully when the component will damage can be made.

In an ideal situation in which the surface of the sample under test was completely uniform i.e. no variation spatially in the optical, thermal and mechanical properties then the sample when it damaged would do so uniformly over the whole of the beam-target interaction area¹⁻³, however in reality where optical quality surfaces are used residual mechanical stress, strain and surface scratches together with contamination from polishing material and the general surroundings ensures that the samples do not have uniform surfaces with the consequence that damage tends to occur in isolated places corresponding to the location of these defects ("weak spots").

The presence of these defects means that there are areas of the surface that will damage more easily than others, a factor of 5-10 difference between "good" and "bad" areas not being uncommon and hence if the surface is probed by suitable technique it should be possible to find the spatial location of these weak spots.

4.2 Review

As discussed in chapter 1 in real situations LID for large beams i.e. larger than a few tens of square microns and long pulses i.e. longer than a nanosecond can be attributed to the presence of surface defects which absorb large amounts of the incident radiation which causes its temperature to rise excessively leading to damage via melting, vapourisation or mechanical failure of the surface.

Essentially then the failure mechanism is thermal in nature and the majority of previous workers have attempted to perform EDLID by means of thermal probes⁴⁻⁷ although other techniques such as predamage electron

emission phenomena⁸ have been studied. The thermal probes used rely on one of two effects, either the photoacoustic⁴⁻⁵ or photothermal⁶⁻⁷ effect (see chapter 1).

In the work by Jeen and Green⁴ the photoacoustic signal generated as a result of the laser sample interaction was studied as a function of incident laser fluence for NaCl and KBr crystals and was found to be linear at low fluences but at high enough fluence there was a step in the gradient of the signal versus fluence graph which the authors called a "super-linear rise". It was found that the fluence at which this step occurred could always be associated with the onset of LID although it was often very difficult to observe this damage. They then built a simple electronic system that would compare the photoacoustic voltage generated at any laser fluence with a preset voltage corresponding to the photoacoustic signal produced just below damage and, if it exceeded the preset level the system would activate a warning light. In this way they were able to setup a system that would warn the user about the onset of damage before it actually occurred. The major problem with this technique is that no information as to the spatial location of the defects/damage areas is obtainable however there is the advantage of continuously monitoring the sample for any signs of LID.

The work by Freese and Teegarden⁵ again used the photoacoustic signal but by using a suitably non damaging focused laser beam the samples, which were thin films of Ge deposited on Ge substrates, were scanned in raster fashion to yield the photoacoustic signal at any point. By doing this it is then possible to determine the spatial locations of defects before the damaging laser pulse is applied. Excellent correlation was seen ~~between~~ the defects picked out by the pre-damage photoacoustic scans and the location of the isolated low damage "weak spots" found after the sample was damaged by a 40 ns CO₂ laser pulse. This technique has the advantage of giving spatial information on the damage but as it is a scanning technique is very time consuming requiring many laser pulses which themselves may change the surface and, it cannot be used to monitor the surface in real time.

The work performed by Abate and Roides⁶ uses a computer controlled raster scanning technique, but probes by means of the photothermal effect. The beam from a cw Argon ion laser that is chopped at several kilohertz is used as the non damaging photothermal excitation source and is raster scanned across the surface causing localised heating. A cw HeNe laser beam is then skimmed across the surface so as to interact with this localised heating causing the HeNe beam to be deflected from its original path. A position sensor connected to a set of lock-in amplifiers arranged in quadrature was used to detect both the amplitude and phase of the deflected signal. A computer was then

used to convert the signals into an absorption map of the surface. Once the surface had been mapped out it was exposed to a damaging pulse from a frequency triple Nd:Yag laser at approximately 20% above its single shot threshold. Moderate correlation between the sites measured as having high absorption and the actual damage sites was obtained, significant numbers of areas not showing up as having high absorption were also found to have damaged. Based upon the work to be discussed later in the chapter it is the author's contention that the "moderate" performance of this technique is due to either using different pulse lengths in the probing and damaging parts of the experiment or to using a laser fluence that is too much in excess of the damage fluence such that once damage has been initiated, intrapulse feedback processes result in the damage "spreading out" across the surface away from the original defect position giving seemingly poor correlation between the predicted and actual damage sites, whereas if the fluence had been lower i.e. at the damage level there would not have been enough energy for the damage to "spread" out thereby preserving good correlation. This technique gives information as to the spatial location of the defects but due to its scanning basis is time consuming (requiring about 8 hours to scan an area 0.1 by 0.1 mm) and requires many heating pulses on the surface which may themselves change the surface.

The work performed by Mars and Porteus⁷ was a first attempt to devise a technique that would allow a single non damaging laser pulse to yield the spatial location of "all" the defects. The technique was based upon the direct imaging of the defects in the surface by imaging the scattered (non specularly reflected) light off of the surface. The light source for both the imaging and damaging was a tunable dye laser operating between 460 and 530 nm, and in both cases the light was suitably focussed and arranged to be near normally incident on the sample. The surface was then viewed at an angle of 15 deg. to the normal i.e. such that only light scattered from the defects could reach the camera. A non damaging laser pulse was then applied to the surface which enabled the spatial location of the defects to be recorded via the video system. A pulse approximately 200% greater than the non damaging pulse was then used to actually damage the sample and provide a picture of the surface scattering after damage. In general correlation between the areas that damaged and the illuminated defects was good, but sometimes it took several non damaging shots of increasing energy to make the defects visible and by no means all the defects damaged and no indication of which one(s) would damage first was possible although, there was some suggestion of a correlation with defect size. It is by no means clear that

1) the use of several so called non damaging pulses is not in itself damaging the sample and

2) the direct use of scattered light will yield information about areas that are optically perfect but contain thermal defects since it will be shown in the chapter that areas that are optically good can contain thermal flaws which result in them damaging early.

In addition the lack of information concerning which area(s) will damage first is another significant drawback.

4.3) System Requirements

Ideally the technique for EDLID should be able to

i) Obtain spatial information about the surface weak spots both optical and or thermal in nature, preferably in the form of a direct picture of the surface.

ii) Require only a single non damaging laser pulse to obtain all the above information

iii) Use the same laser for both non damaging and damaging pulses to avoid any possible problems relating to different pulse lengths or beam profiles

iv) Use a recording system such as a video framestore that can be interfaced to a computer so that data processing e.g. the comparison of two images can easily, accurately and quickly be performed

4.4 The use of Interferometry

Since the heating of the air surrounding a sample that has absorbed significant amounts of laser radiation causes the refractive index in the locality and hence the optical path to change the first approach adopted was to look at the possibility of using interferometry to perform EDLID.

Following close discussions and collaboration with Dr. P.A. Temple⁹ a piezoelectrically controllable Fabry Perot Interferometer (FPI) was setup. The basis of this was that by using high reflectivity mirrors a high finesse system and therefore one very sensitive to small phase changes and hence changes in optical path could be built. Two configurations of the system depending upon whether the sample was a reflective or transmissive were envisaged as shown in fig 4.1. The sample surface illuminated by a 50mw HeNe laser that was fitted with an etalon to ensure single longitudinal mode operation was imaged using the video imaging system described in chapter 2.

Although the mirrors used were 95% reflecting and flat to $\lambda/100$ which according to theory¹⁰ would give a finesse of around 60 and in reality gave a figure around 30, the overall finesse of the system would in practice be

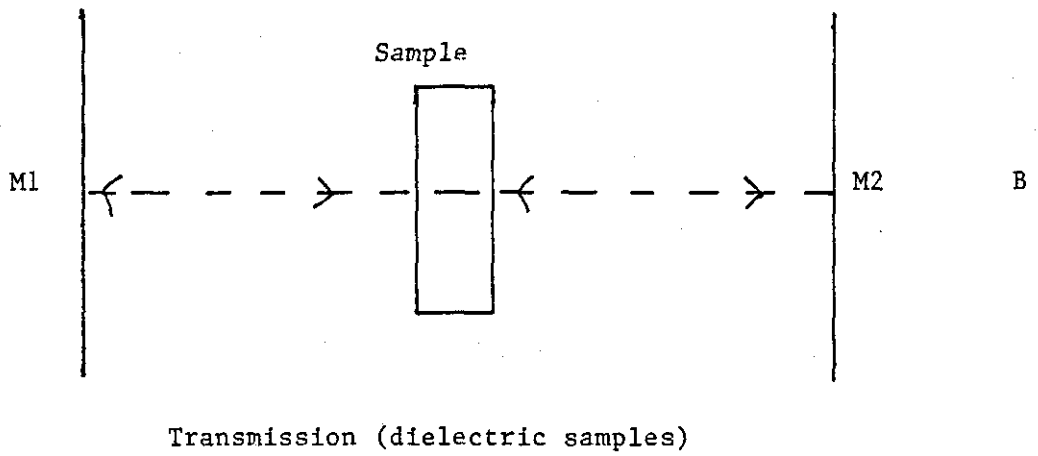
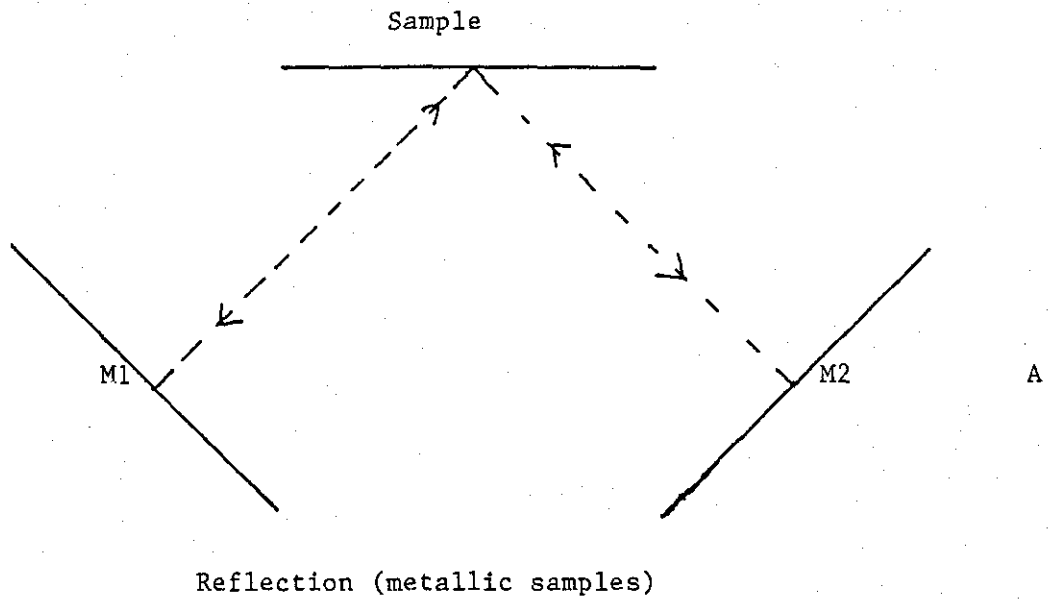


Fig 4.1 FPI Setup

limited by the flatness (typically $\lambda/10$ or worse) of the sample used, particularly if it were transmissive when, there is the possibility of a secondary etalon being formed by the sample itself.

Initial experiments and calculations¹⁰ indicated that with the experimental setup available the modal stability and divergence of the HeNe beam would not cause a problem, however the problem of multiple images in the FPI is much more serious.

It can be shown¹¹ that to first order the finesse in a FPI corresponds to the number of round cavity passes that a beam of light makes before it has totally emerged from the FPI and consequently the imaging of a surface within the FPI can be thought of as the summation of multiple images of the surface each separated in space by a distance $2l$ where l is the cavity length. Thus although the camera may be focussed onto the surface i.e. the first image, subsequent images will be out of focus, leading to a final image that is a smeared out version of the original object. Experiments were carried out using a thin piece of wire as test object to determine the severity of this problem.

Fig 4.2a shows the image that is produced from a single thin piece of wire, note the evenly spaced vertical "lines" either side of the wire which are characteristic of a badly out of focus object. In fig 4.2b the intensity profile across the wire is shown on the image and it can be seen that the profile in the vicinity of the wire is similar to the Fresnel diffraction pattern from a single opaque strip as in fig 4.2c. This similarity can at least qualitatively be explained by simple diffraction ideas as follows. The light from the first image will be perfectly focussed on the plane that contains the camera chip, subsequent images will be focussed back from this plane towards the focus of the imaging lens as effectively they come from objects successively further behind the lens. Light then has to propagate from each of these focus planes to the plane that contains the camera chip and as it does so its propagation can be described by the simplified Kirchoff diffraction integral¹² which given the experimental dimensions involved is well approximated by the form appropriate to Fresnel diffraction. Thus the image seen is the sum of a replica of the object and a number (i.e. the finesse - 1) of images that are similar to the Fresnel diffraction pattern of the object giving a final image that will have similarities to the Fresnel diffraction pattern of the original object.

The variation of intensity with path length in the FPI follows the well known Airy pattern¹⁰ and fig 4.3 shows the actual pattern obtained for the FPI used. In order to detect both increases and decreases of optical path one of the mirrors of the FPI was adjusted piezoelectrically so that the output

Fig 4.2a

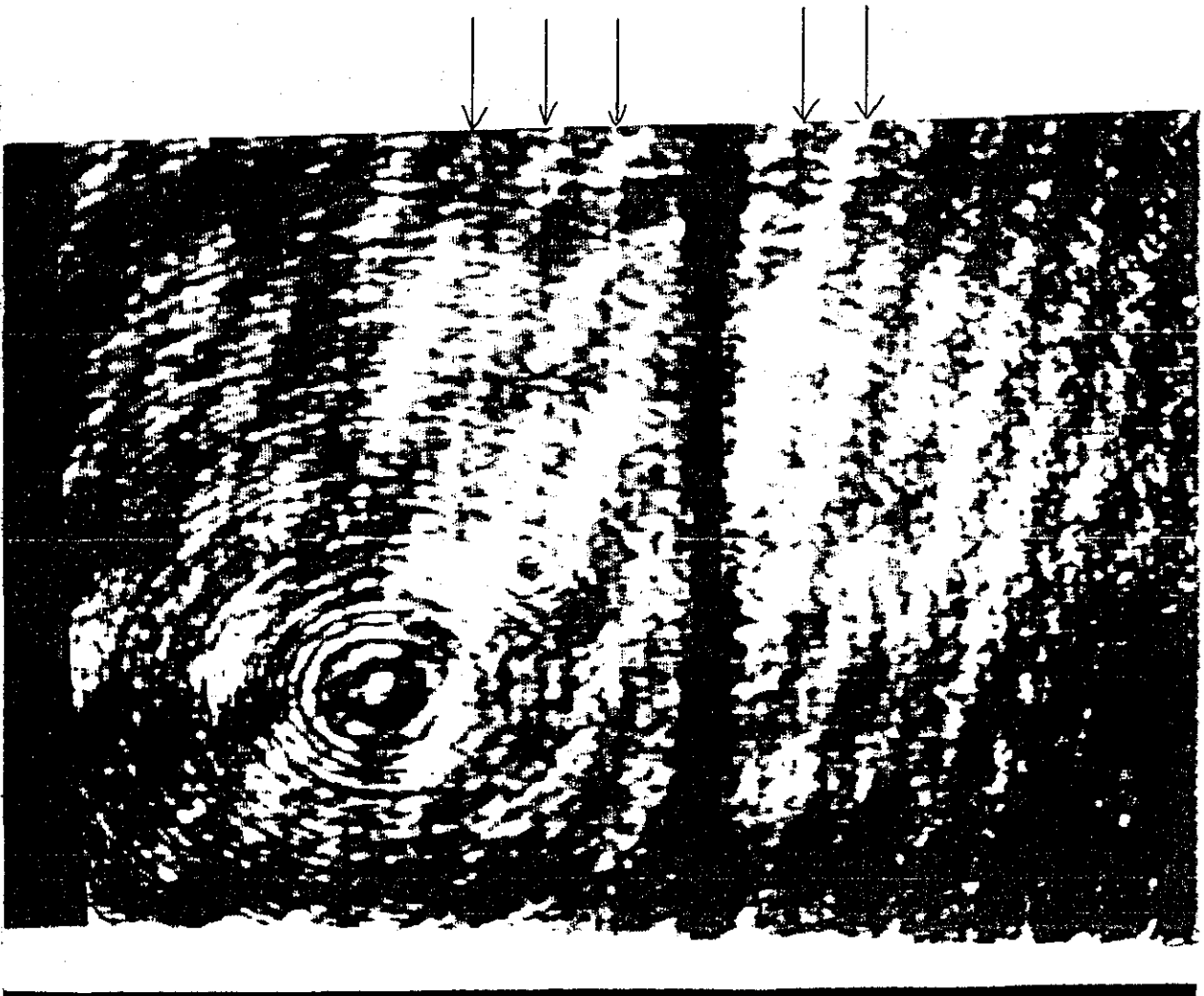
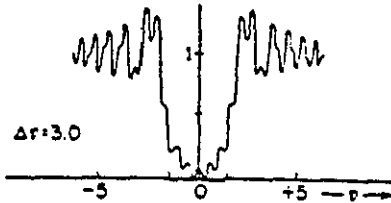
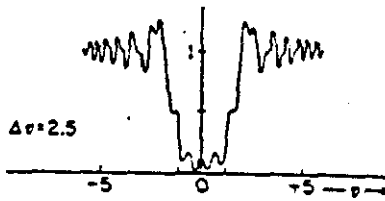
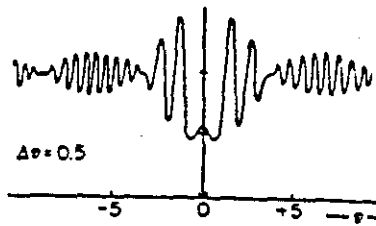
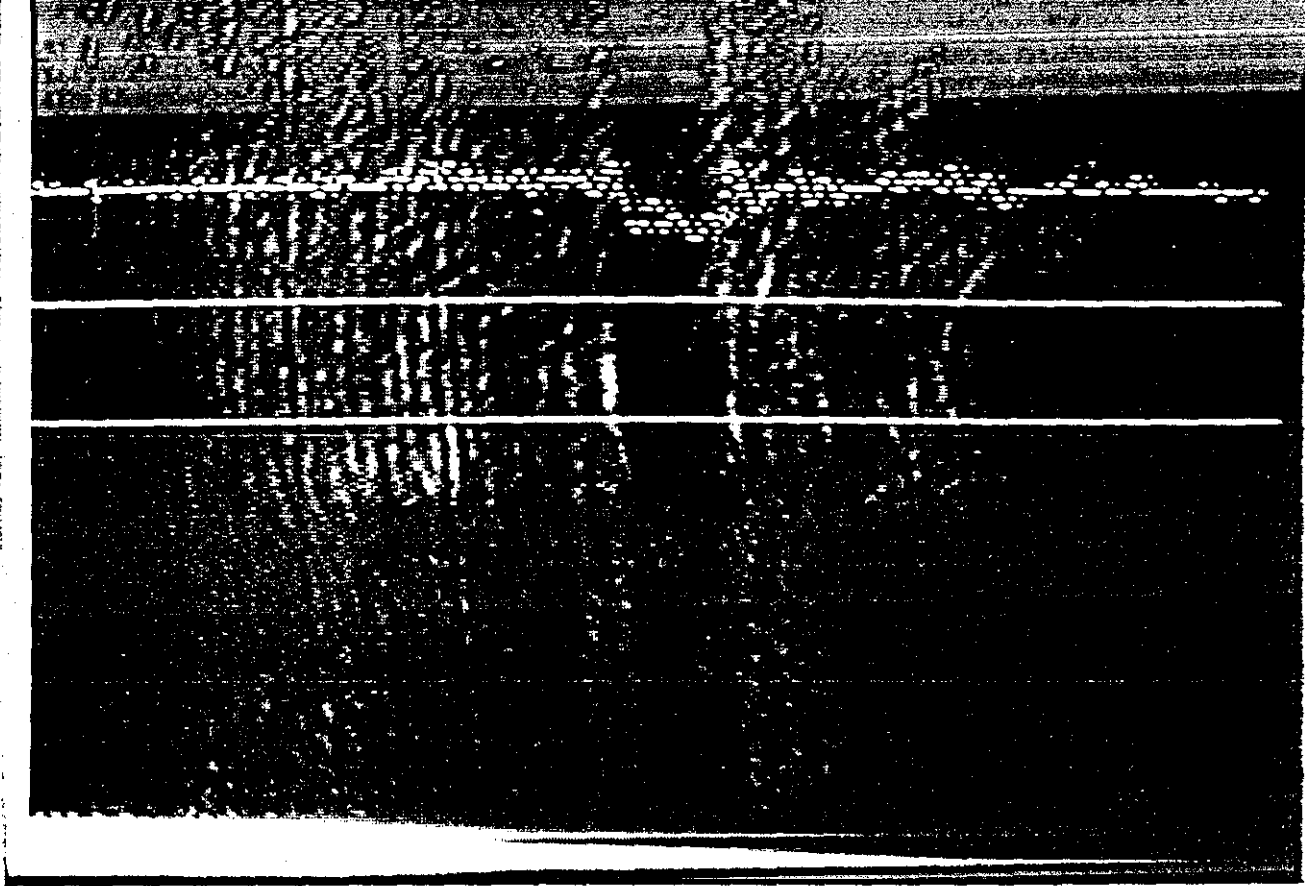


Image of Single Wire in FPI

Note multiple lines (images) as arrowed



Above Fig 4.2c Fresnel diffraction by single opaque strips
from ref 19

Above top Fig 4.2b Profile of the image of a single wire in
the FPI obtained by use of the video processing
system

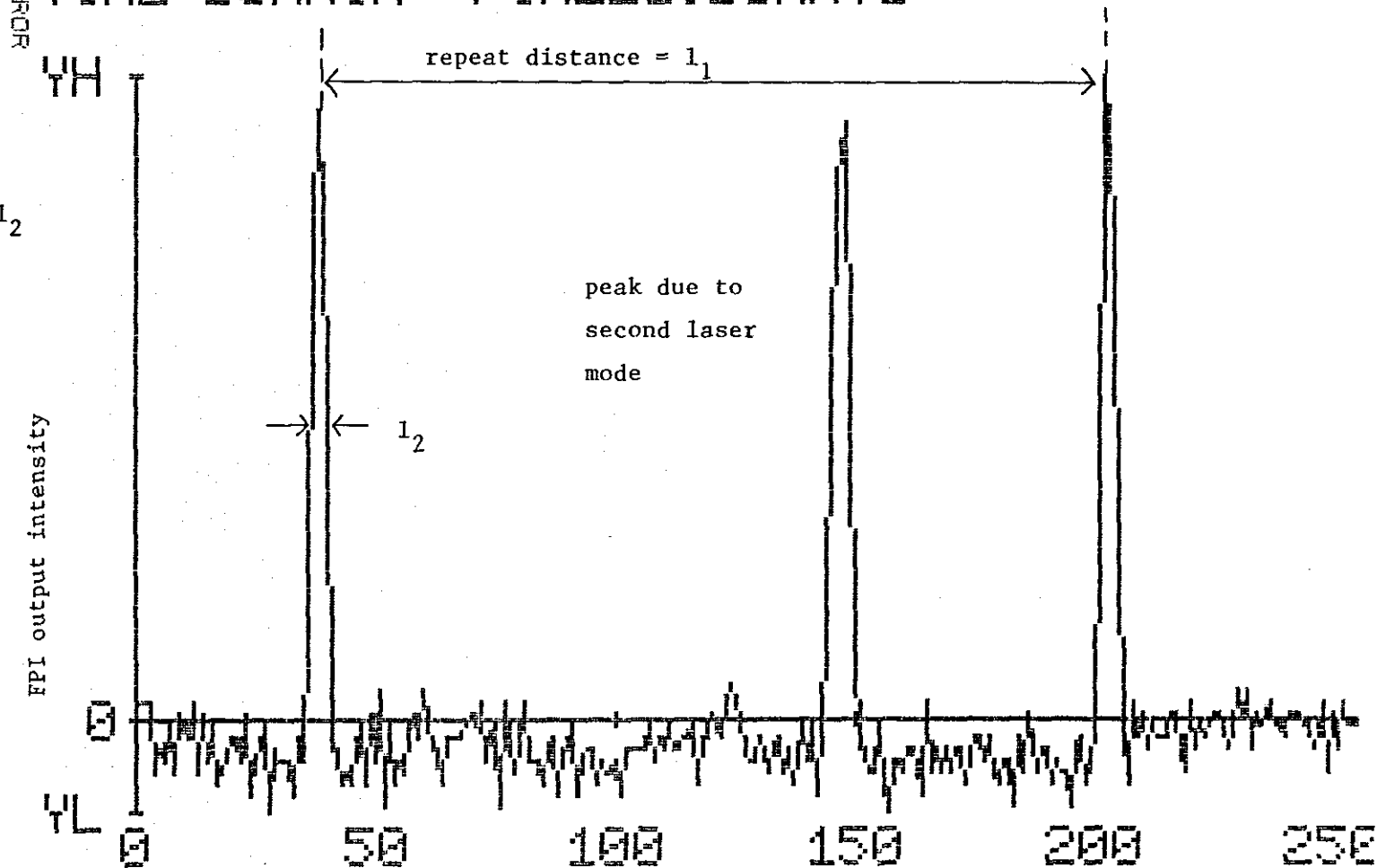
Fig 4.3

X AXIS SCALE IS 4.000 E1 US / DIV
 YL = -2.500 E-2 YH = 1.84375 E-1
 TIME DOMAIN PIXEL1:DUMMY1

$l_2 = \text{FWHM}$

$F (\text{finesse}) = l_1 / l_2$

$= 40$



Type any key to continue

Horizontal Axis = Path Length in FPI

corresponded to being half way up any given peak. This of course meant that a significant amount of light was always on the camera which, in practice meant that the difference between areas that showed the transient, often intense but short lived effects (in comparison to the 20 ms imaging time per field of the video camera) and the unaffected areas was considerably reduced by the continuous background light, since effectively the transient changes were a small a.c. effect on a large d.c. background. Ideally the image would have been a uniformly illuminated field but in practice the image always had rings (or parts of rings) of light and dark.

Despite these problems it was possible to detect some transient effects in highly absorbing samples e.g. metals as shown in fig 4.4 although unfortunately to detect these changes the laser energy that was used had to be very close ($\approx 95\%$) of the damage threshold with the result that the sample would often damage in the attempt to detect transient effects.

The requirement to work so close to the damage threshold clearly made this system an impractical technique and emphasis was placed on finding an alternative method for EDLID.

4.5 A Pulsed Laser Schlieren Imaging System

The work on the FPI had shown that cw illumination with long imaging times was not a suitable technique leaving two choices either cw illumination with a gated camera or, pulsed illumination with ordinary cameras, of which the latter was chosen due to the availability of equipment.

The basic experimental setup was to use the excimer laser for both the damaging pulse(s) and non damaging pulse where it would transiently heat both the surface and, by re-radiation from the surface the surrounding air whilst, a separate pulsed dye laser would be used to image the surface. The imaging light would be focussed by a first lens which had a knife edge suitably placed in its focal plane so as to remove the undeflected beam to form an intermediate image (i.e. Schlieren imaging). This intermediate image would then via suitable microscope objectives be imaged onto the camera chip and recorded by the video processing system.

The physical idea that the experiment was based on is that the "weak spots" on the surface would absorb anomalous amounts of the incident excimer radiation and thereby become hotter than the surrounding areas which would cause more deflection of the pulsed dye laser in these areas due to the temperature dependence of the refractive index of air and thus show up as anomalously bright or dark areas in the Schlieren image.

Since the samples to be used were real optical surfaces they inevitably contained scratches, small defects etc. which would mean that the initial

initial image



image showing transient change

Schlieren image was not totally uniform across the field of view, allowance for the light in the image before the transient heating would have to be made necessitating 3 images of the surface being taken. The images would contain the following

- i) The initial surface
- ii) The surface showing transient effects
- iii) The surface after the transient effects have decayed to zero.

with each image being stored in the video processing system described in chapter 2. Processing (see later in this chapter) of images ii and i would yield only transient effects whilst comparison of i and iii would reveal if the pulse had caused any permanent damage.

Due to the way in which the video framestore works and the requirement to avoid the half pixel shift that exists between the even and odd fields in CCD cameras the dye (imaging) laser would be fired 3 times with a gap of 80 ms between pulses thus enabling 3 separate images (but all on either the odd or even video field) to be captured, digitised and stored in the framestore. Initial experiments and previous experience showed that background vibrations, air currents etc would not cause any movement of the image between dye pulses on this time scale nor indeed on the scale of the 1 to 5 minutes it would take to expose the sample to a suitable number of excimer pulses for damage to occur.

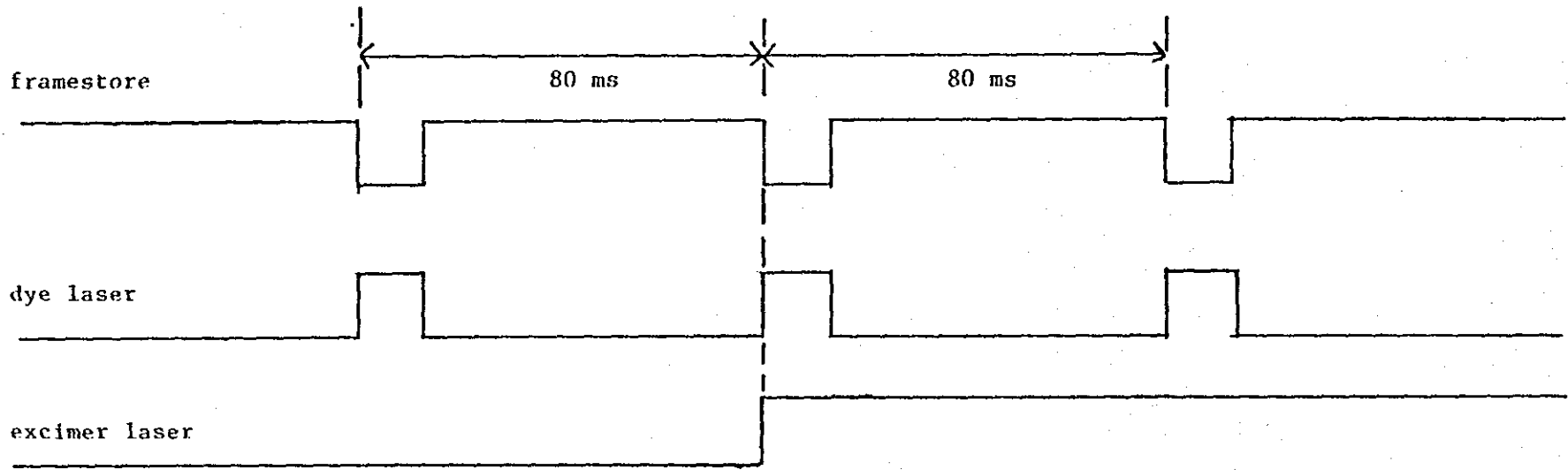
The excimer laser would be fired only during the the second imaging period with the dye laser being fired a known "short" time after the excimer and so capture the transient heating of the surface. Fig 4.5 shows the timing sequence of the framestore, dye laser and excimer laser. Based upon the general "belief" amongst workers in the field that the heating of sample and surrounding air lasts of the order of a microsecond and that the peak of the heating is around the end of the laser pulse, modifications to the video/dye laser triggering box were made so that initially the dye laser could be triggered up to 400ns after the excimer laser although later, after measurements of the time duration of the transient heating effects for the excimer fluences used had been made this was extended to 10 μ s.

4.5.1 System Configuration

4.5.1a Hardware

The available pulsed laser was a model LN103 nitrogen pumped dye laser manufactured by PRA of Canada. The system consisted of a 80 μ J N₂ laser

A



B

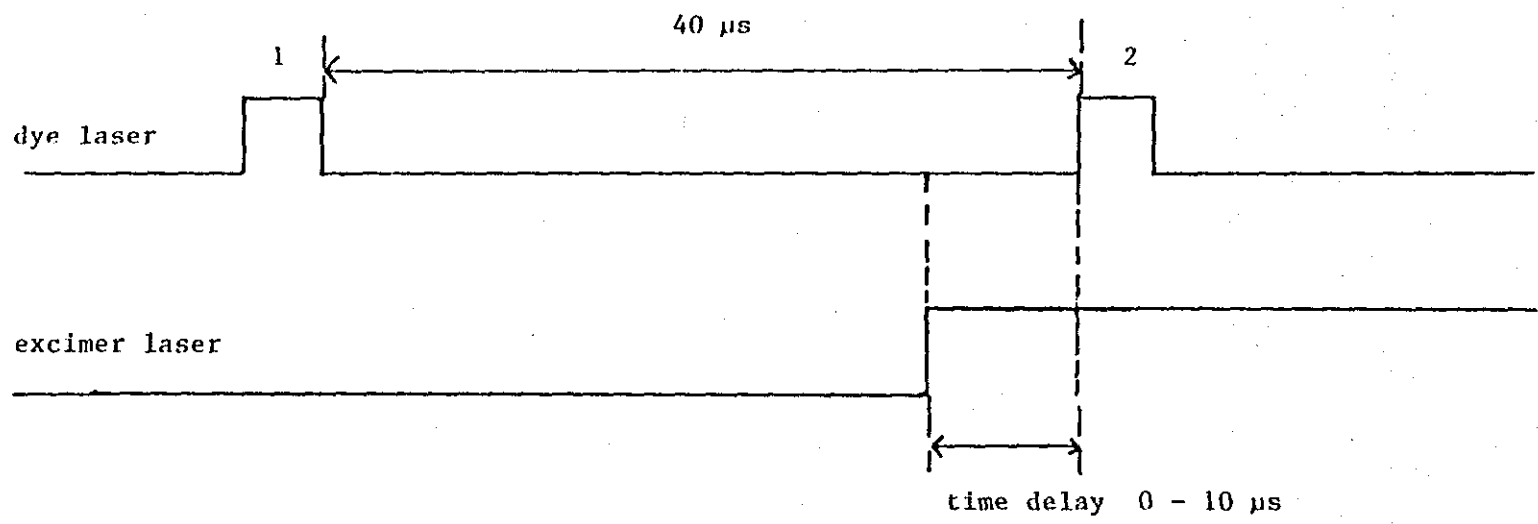


Fig 4.5 Timing Sequences

operating at 337 nm that would pump a small dye laser attachment which could give up to 12 μJ of energy depending upon the dye and operating wavelength used, with a pulse length of around 0.5 ns i.e. much shorter than the excimer pulse (25ns). The spatial profile of the beam from the dye laser was, after being spatially filtered, to a good approximation Gaussian. As with the general damage work, the triggering of this laser, through a purpose built electronic delay box (based on TTL chips 74121) was linked to the field sync pulses coming from the video camera.

Whenever 2 separate lasers are required to produce pulses that are temporally linked to each other there will be some temporal jitter i.e. variation in the time between the two pulses on a shot to shot basis caused by small variations in a number of parameters such as the time the discharge breakdown to initiate laser action, noise on the electronics of each device etc and in particular noise on the electronics of the device that is controlling both lasers. Experiments using a fast uv photodiode and suitable storage oscilloscope which measured the time between the application of a suitable triggering pulse to the excimer and the arrival of the excimer radiation at a point in space very close to where one of the monitoring photodiodes that this experiment requires would be situated, showed that the jitter in the excimer laser (relative to the triggering pulse) was less than 6ns.

The dye laser has two modes of operation one with a trigger jitter of ≈ 60 ns (regular mode) and the other, the "low jitter" mode of ≈ 1 ns (manufacturers figures). Operation in the low jitter mode requires two TTL trigger pulses to separate parts of the trigger circuit with the second between 30 and 50 μs after the first. Further modifications were made to the video/dye triggering box so that two pulses 40 μs apart could be applied to the dye laser. After extensive optimisation of the settings for discharge voltage, gas pressure, lasing volume and the spark gap of the N_2 pumping laser, the trigger jitter of the dye laser on a shot to shot basis was measured in the same way as that of the excimer and found to vary randomly between 10 and 30 ns.

The controlling electronics was found to have approximately 10 ns of jitter shot to shot between its output of the pulses to either laser on any given single shot.

Overall therefore the jitter shot to shot between the two lasers varied from about 10 ns to 40 ns. A fast photodiode pd1 in fig 4.6a linked to a fast digital storage scope was used to monitor the delay on each pulse between the excimer and dye lasers. For the purposes of these experiments the time delay between the excimer and dye pulses is defined as the time between the

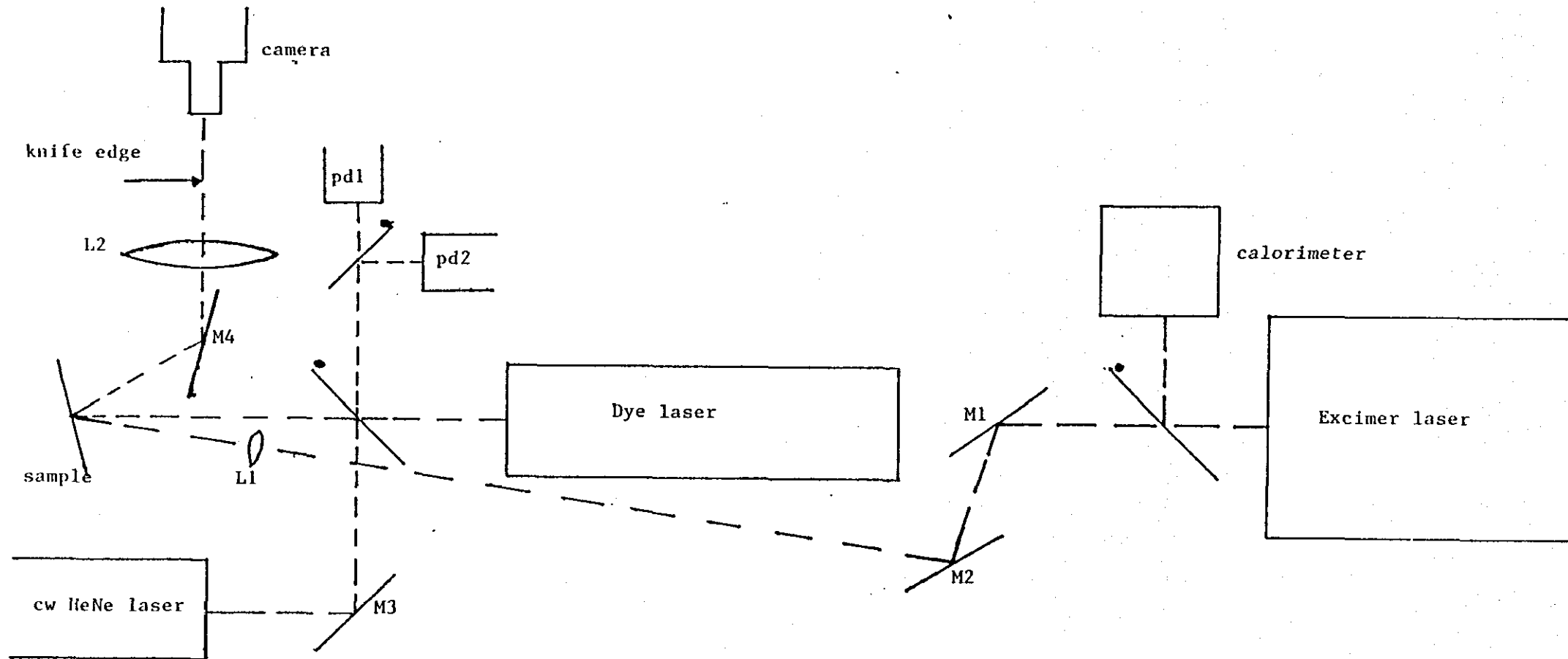



Fig 4.6a

Schematic of experimental layout

 = beamsplitter

M? = mirror

pd1 = photodiode to monitor dye laser energy

pd2 = photodiode to monitor delay between excimer and dye

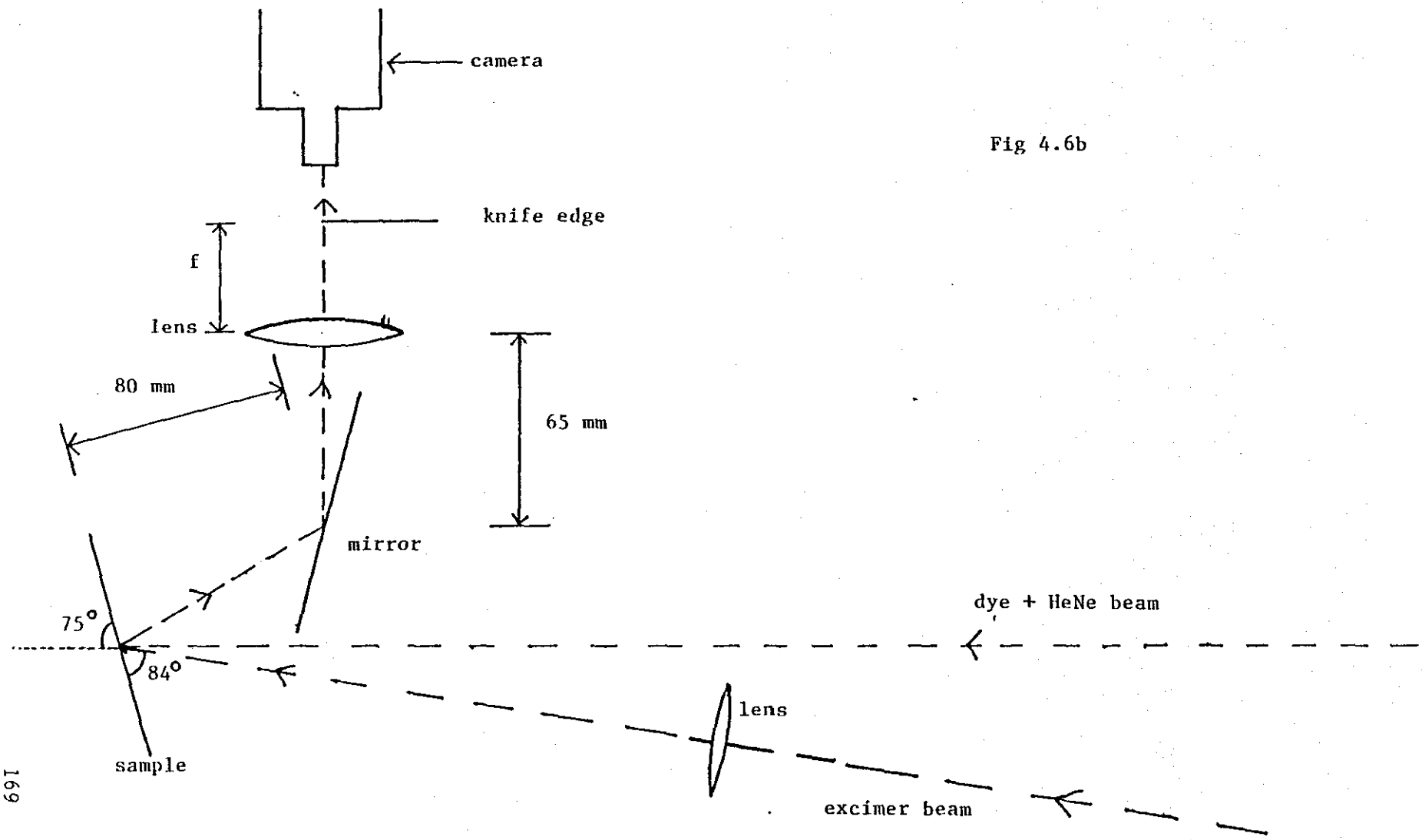


Fig 4.6b

end of the excimer pulse and the start of the dye pulse. A small fraction (8%) of the dye beam was split off and directed onto pd1 whilst there was enough scattered excimer radiation for the diode to respond to without specifically having to split a fraction off of the beam.

Initial experiments showed that during the first 100 ns or so the magnitude of the transient effect decayed very rapidly and so early experiments were conducted with the delay set at the maximum then available (375 ns). Later experiments and in particular (see 4.5.9) the one to investigate how the transient effect changes with increasing number of shots as the multiple shot (arbitrarily chosen as either 10 or 50 shots) damage threshold is approached were conducted with a longer delay ($\approx 1 \mu\text{s}$) to reduce the problem of variation in the transient effect caused by jitter in the system.

A number of images of the same area on a sample were taken and analysed on the video system revealing a large shot to shot variation in the output energy of the dye laser. In order to usefully compare the required images either the energy in each would have to be the same or some measure of the difference between the images would be required, and since experiment clearly showed the former was impossible, a system was devised to indicate at least the relative energy of each dye shot.

Ideally some form of calorimeter or pyroelectric device would be used to measure absolutely the beam energy but no device capable of measuring the few μJ 's output of the laser was available therefore only the change in energy between pulses could be found, this was done by using a simple fast photodiode, pd2 in fig 4.6a. The scope used to record the output of pd2 was a Philips PM 3311 60 MHz digital storage scope as this model has built in the facility to capture and store in separate memories 4 successive incoming signals without any manual intervention, however there must be at least 50 ns between each of the signals which is part of the reason that the time between each of the three dye laser pulses was set at 80 ns. Unfortunately, this scope would not work in single shot mode faster than 200 ns per division which corresponds to a sampling time of 8 ns per point. The time duration of the dye pulses when measured with the fastest available scope (a Tektronix 466) and using pd2 suitably terminated (when according to the manufacturers figures the rise time should be 0.5 ns) was found to be $\approx 5\text{ns}$, a time corresponding to the limitations of the scope electronics rather than to the actual pulse length which, from manufacturers data obtained by a Streak camera is of the order of 0.5 ns. Thus given that the scope to be used for measuring the dye pulse energies could only sample to 8 ns i.e. provide only 1-2 sample points per dye pulse it was decided to use pd2 without a 50 Ω terminator which gave a rise time of around 1 μs when connected directly

into the Philips scope. This meant that pd2 would effectively act like a capacitor discharging through the high impedance of the scope where, the total initial charge induced into pd2 by the incident laser pulse during its 0.5 ns duration equalled the number of created electrons which are directly dependant on the number of incident photons and hence to the incident energy. Since the charge takes of the order of 1 μ s to start to leak out of pd2 i.e. much longer than the time to create it, the voltage output of pd2 at any time is a measure of the incident energy, thus to obtain a measure of the energy in each pulse the output of pd2 was recorded and the voltage produced at a convenient time usually 2-3 μ s after the dye pulse was used as a measure of the energy of each pulse.

A small 10 mW cw HeNe laser was arranged via suitable mirrors and beamsplitters (fig 4.6a) to be co-axial with the dye beam so that it could be used for both aligning the system i.e. camera(s), sample and excimer laser which, is substantially easier and more accurately performed with cw illumination than with pulsed light and to allow real time direct monitoring of the surface during the multiple shot experiments.

Initially when only one camera was available a piece of card suitably mounted was used to block either the cw or pulsed beam depending upon which was required. At this stage to avoid having to refocus the camera when changing light sources the dye laser was set to work at the HeNe wavelength 632.8 nm. The major disadvantage of this single camera setup is that it does not allow the simultaneous real time cw monitoring of the surface (for any obvious signs of damage) and the capture of transient information using the dye laser. Towards the end of the experiment a second camera became available which allowed the setup to be modified (fig 4.7) so that one camera imaged the surface in real time using the cw HeNe and the other was linked to the video system and recorded the dye pulses. The dye laser was, in this setup set to operate at 514 nm. The separation of the co-axial beams was achieved by means of narrow band pass filters, one set at 632.8 nm and one at 514 nm and experiments readily showed that no detectable amount of the 514 light was passed by the 633 filter and vice versa. To allow for the differences in intensity of each laser and the responsivity of the cameras at the two wavelengths, each camera was provided with its own knife edge so that the cut off point could be set independently on each camera.

The cameras used were all EEV CCD frame transfer devices which would image an object for 18 ms out of the 20 ms field time and take the other 2 ms to transfer the image to the store area on the chip from where it would be read out during the next field.

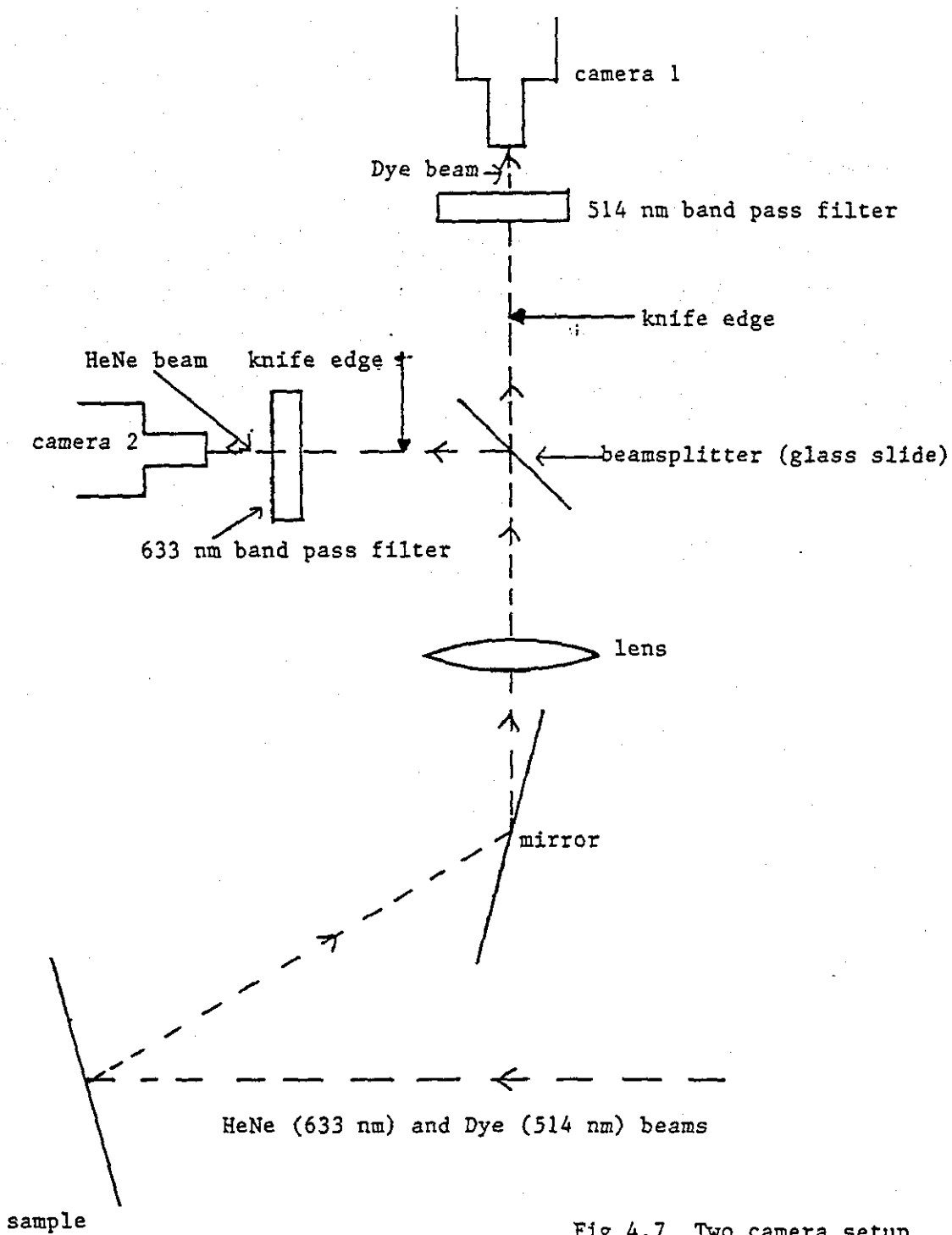


Fig 4.7 Two camera setup

4.5.1b Optics

Due to experiments being performed by colleagues the space available to setup the experiment was limited particularly once the dye laser had been inserted on the optical table. Initially metallic samples were to be used and the system was setup to work with the light that was reflected from the sample although at the end transparent dielectric samples were used which required the system to be rearranged to view the transmitted light. With space limited the optics at the target area were setup as shown in detail in fig 4.6b.

In order to use Schlieren imaging, access to the focal plane of the imaging lens is required and it was deemed that this would be far easier if an intermediate lens 12 was used rather than trying to perform Schlieren in the focal plane of the microscope objectives that were to be attached to the camera to provide the required magnification. The knife edge was accordingly setup in the focal plane of this intermediate lens. The use of an intermediate lens would also have the advantage that some initial magnification could be accomplished which would mean objectives of less power could be used with a reduction in the problems of depth of focus and aberrations caused by having "the tube length" of the system not exactly correct.

The diameter of lens 12 and that of mirror m4 were chosen so that with the space available they would be consistent with resolving objects of the order of 3 μm in diameter, this was done by reference to the Abbe theory of resolution¹².

The Abbe theory requires that in order to resolve an object of diameter d using light of wavelength λ the detecting system must have an acceptance angle for the light coming from the object at least equal and preferably greater than θ , where $\sin\theta = \lambda/d$ i.e. at least the first order diffracted ray from the object must be detected in order to resolve the object. With $d=3 \mu\text{m}$ and $\lambda=633\text{nm}$ requires $\theta \geq 12.2$ deg.

The distance p between the sample and the reflecting mirror m4 was limited to 80 mm thus with $\theta= 12.2$ deg requires the radius of the mirror given by $D = p \cdot \tan\theta$ to be greater than 18 mm. A 60 mm diameter mirror was used.

The distance to lens 12 was limited to a minimum of 145 mm thus requiring a diameter 63 mm or greater. A 63 mm diameter lens with a focal length of 100 mm was used.

By means of a suitable adapter tube a microscope nose turret with 2 objectives on it was attached to the camera(s). The adapter tube was made such that the distance from the shoulder of the objectives, which were all parfocal types, to the plane containing the CCD chip was approximately 160

mm i.e. such that the image formed by the objectives was focussed onto the CCD chip at a working distance close to that for which the objectives were designed so that 1) aberrations are kept to a minimum, 2) advantage of the parfocal nature of the objectives could be taken to avoid gross re-focussing on changing objective and 3) the magnification produced by the objectives was close to the value that they would have when used in a normal microscope. The objectives chosen were a X4 with a numerical aperture (n.a.) of 0.12 and a X10 with n.a. of 0.17. The X10 is consistent with a resolution of around 4 μm and was the only readily available version of this objective. The spacing of the pixels on the video camera was 22 μm which when the image has been through a X10 objective corresponds to 2.2 μm well inside the 4 μm limitation of the objectives. Thus overall the resolution of the system was limited by the microscope objective to a value around 4 μm .

4.5.1c Computing

On the basis that on each shot 3 images (i.e. 192 kbytes of information) of the surface of a sample would have to be stored and processed the Sperry PC/IT with its fast 80286 processor, 45 Mbyte hard disk and 1.2 MByte floppy disks was chosen over the slower Amstrad PC 1512 to be the system computer.

The programming was divided into two areas..

1) General data handling such as image capture, storage to disc and simple processing such as image subtraction and enhancement were written in assembly level language for maximum speed of operation.

2) Advanced image processing e.g. subtraction where allowance is made for the variation in the energy of the laser pulse that was used to record each image, because this involves multiplication and division was written in compiled BASIC.

Details of area 1 are given in chapter 2 and a discussion of the concepts and mathematical expressions (but no actual programming) used in area 2 is given below.

A program was written to subtract and then automatically enhance the 3 images on the assumption that as the pulse energy increased/decreased the recorded greylevel increased/decreased linearly. With the framestore in the 256 by 256 pixel mode giving 4 stores a,b,c,d the images were always arranged to be (when processing transient data)

Store a....image of initial surface
 Store b....image of the transient effect
 Store c....image of the surface after the transient has decayed
 Store d....spare, used as a temporary store during processing

or if the data collection was complete i.e. the transient data collected, sample damaged and an image of the damaged surface collected the images by suitable manipulation of the stored data would be arranged to be

Store a...image of initial surface
 Store b...image of transient effect
 Store c...image of damaged surface
 Store d...spare

Denoting the energy of each image by E_a, E_b, E_c , the greylevel at any point in the image by a in store a, b in store b etc.., taking the energy of the first image to be the "correct energy" and noting that as before in order to represent both increased and decreased light levels that the zero point is set at level 31 the program would process the images on a point by point basis so that the greylevel at a point appropriately rounded is given as below...

Store a...unchanged...image of initial surface
 Store b... $31+f*(a-(E_a/E_b*b))$ i.e. $a-b$...net transient effect
 Store c... $31+f*(a-(E_a/E_c*c))$ i.e. $a-c$...net permanent change
 Store d... $31+f*E_a*(b/E_b-c/E_c)$ i.e. $b-c$...transient-permanent

where f is a factor determined by the program that enhances the deviation of the processed images from level 31 such that the maximum deviation from 31 be it positive or negative is set to 31 levels making the overall signal either 0 if the maximum deviation was <31 or 63 if it was >31 . The use of this factor f automatically enhances the images to their maximum contrast level.

Processing in this way requires only the fractional change in energy shot to shot rather than the absolute energy of each shot and therefore the use of $\mu p 2$ is adequate, and enables the net transient and permanent effects to be separated from the dust, scratches etc that are present on the initial image.

Experience of recording the dye laser images showed that if the energy varied by more than about 40% shot to shot the spatial location of the centre of the beam moved e.g. if the centre i.e. most intense part of the

dye beam was arranged to be in the centre of the image then if the pulse energy varied too much the centre part could move up to a few mm as viewed on the monitor. Thus in practice if the shot to shot energy variation exceeded 20% the images which were recorded were ignored and new ones with less than 20% variation obtained.

Whilst this simple approach of merely subtracting the images worked very well in correcting for the shot to shot energy variations it made no allowance for the spatial variation of either the dye laser beam or the excimer beam.

In order to calibrate the magnification of the system and therefore the size of the imaged area as viewed on the television monitor connected to the framestore, a ruled graticule was placed in the same position as the surface of a test sample would have been and then imaged through the video system yielding the following values

Table 4.1

Objective	Magnification	Comments
X4	260	Low mag setup
X10	680	High mag setup

These figure meant that since the viewing area on the TV monitor was approximately (horizontal figure first then vertical) 280mm by 210 mm the area actually imaged was 1.08 mm by 0.81 mm in the low mag setup and 0.41mm by 0.31 mm in the high mag setup. At the target plane used the excimer beam was approximately Gaussian with a $1/e^2$ diameters of 2mm horizontally and 0.5 mm vertically. The dye laser was spatially filtered such that it had a $1/e^2$ diameter of approximately 2mm. Using the formula

$$I = I_0 \exp(-2*(r/r_0)^2) \quad 1)$$

where r_0 is the $1/e^2$ radius and noting that the images from the dye laser were always arranged to be centred on the screen as was the excimer beam, tables 4.2a and 4.2b show the variation in intensity in each beam across the viewed area.

Table 4.2a

Variation of dye laser intensity with distance and magnification

Magnification	Axis	Max r mm	Edge ratio*
low	Horizontal	0.54	0.56
"	Vertical	0.41	0.71
High	Horizontal	0.21	0.92
"	Vertical	0.16	0.95

* The edge ratio is the ratio of the intensity at a distance corresponding to the edge of the observed area in each direction (i.e. the values in the max r column above) and that at the centre ($r = 0$) of the pattern using 1) with $r_0 = 1\text{mm}$.

Table 4.2b

Variation of excimer intensity with magnification and distance

Magnification	Axis	Max r mm	Edge ratio*
Low	Horizontal	0.54	0.56
"	Vertical	0.41	0.004
High	Horizontal	0.21	0.92
"	Vertical	0.16	0.44

*edge ratio calculated as in table 4.2a but using $r_0 = 1\text{ mm}$ for the horizontal axis and 0.25 mm vertically.

It is immediately obvious that in the high mag setup which was the one that almost all the data was recorded in, the variation in the dye beam can be ignored whilst that of the excimer can only be ignored along the horizontal axis. To correct for the variation in the excimer a factor of the order of $\exp(-2*(r_0/r)^2)$ would have to be applied to the net transient signal, but this requires knowing the spatial centre of the excimer beam relative to the centre of the recorded image which is in practice very hard to obtain. Thus in the first instance it was decided not to use any correction factors and process the images by simple subtraction.

Hence in the recorded images the heating and damaging intensities

(fluences) are essentially uniform in the horizontal direction but vary vertically across the image such that if the surface was completely uniform the damage would occur first along a horizontal axis in the centre of the image due to the higher intensity there, but if damage is found to first occur away from the centre vertically then it must be due to the presence of "weak spots" in the surface and as such is a good indication of the presence of these defects.

A second high level program was written to allow a form of semi-automatic monitoring of the transient signal at user selected points as a function of the number of excimer laser pulses that had been incident on the target (see 4.5.9).

4.5.2 Samples

4.5.2a Samples Used

Three sets of samples were used...

- 1) Commercially pure Al (99% Al min.) polished commercially as described in chapter 3
- 2) Laser quality (for CO₂ use) polished to $\lambda/10$ polycrystalline Ge
- 3) Multi layer (quarter wave) dielectric stacks on fused silica substrates setup to be high reflectors at a wavelength of 249 nm

4.5.2b Pre-Imaging Cleaning of the Samples

The Al samples were found to have surfaces that were contaminated by dirt, dust and polishing residue that would cause large scale damage at very low fluences, leading to intolerably large variations of the damage threshold area to area unless, such contamination was removed from the surface by exposure to very low level "cleaning" pulses from the excimer laser. Typically the surface would be pre cleaned before imaging by 5 pulses at a fluence around 0.1 J cm^{-2} . These very low fluence pulses would essentially cause the dirt etc to be "blown off" the surface leaving a much cleaner and more uniform surface.

The Ge and dielectric samples were also pre cleaned at appropriate levels to ensure that dust etc picked up on the surface would not cause poor reproducibility of the results.

4.5.3 Experiments Performed

Three types of experiments were performed on the samples although not all types were performed on all samples...

- 1) To investigate the correlation between areas of anomalous absorption as revealed by the transient signal and the areas that first damage on a single shot damage basis
- 2) As in 1 but on a multiple shot basis
- 3) To investigate the magnitude of the transient effect as the number of incident shots approaches the number required to cause damage at the fluence used to determine if it is possible to predict the onset of damage i.e. how many more shots are required to cause damage by monitoring the transient signal.

4.5.4 Experimental Operation of the System

Once the samples had been inserted into the correct position the dye, HeNe and excimer laser beams had to be aligned. This was performed by using a very high excimer fluence which would damage the sample over a large area i.e. almost the complete field of view when set to high magnification. The dye and HeNe lasers were then set so that the centre of each passed through the centre of this large damage area and the cameras adjusted and focussed so that the image of this area was centred on the tv screen. The excimer laser fluence was then reduced to appropriate levels (see 4.5.5).

The knife edge(s) were set so as to cut off half of the beam (i.e. half the zeroth order and all the spatial frequencies on one side) for both dye and HeNe lasers. The setting for the dye laser was particularly critical since the transient effects could show up as either increases or decreases in the amount of light passing the knife edge and if this was set too far either way i.e. too much or too little was cut off the system would lose its sensitivity. Experience proved to be the best way to set the knife edge so that for both the amount of light reaching the surface and the nature of the surface, the system was at its most sensitive.

Each time a new area was chosen it was pre-cleaned as described before it was imaged. The video system was then placed in its triple frame grabbing mode and the transient signal recorded. If it appeared after processing that the sample had permanently damaged the process was repeated on a new area. If there was more than 20% variation in the dye pulse energy between the shots another transient signal was recorded on the same site, should three attempts on any one site to record the transient signal fail, a new site was chosen and the process started again.

Once the transient signal was recorded the sample was damaged either by a single shot or by multiple shots and an image of the damaged sample using the dye laser was obtained so that comparisons between the initial, transient and damaged states could be made. For the experiments where the magnitude of

transient as a function of number of excimer pulses were being performed a transient signal (using the appropriate excimer fluence) was recorded every 5 or 10 excimer pulses.

The excimer fluence of the transient and damaging pulses, shot to shot dye pulse energy and the time between excimer and dye pulses were also recorded each time. Use of the computer processing routines would enable the net transient and damage signals to be determined and compared.

4.5.5 Selection of Excimer Fluence

The fluence used to record the transient image was that found by experiment to be the lowest fluence i.e. furthest below the single shot damage threshold that would give a readily detectable signal and for the metallic samples was about 1/3 to 1/4 of the single shot threshold level whilst being slightly higher at about 1/2 the single shot level for the dielectric samples. The limit as to how far below threshold the system would work at is controlled by 2 parameters, 1) the sensitivity of the camera and 2) the output energy of the dye laser.

The output energy of the dye laser was such that in operation, the camera was typically operating in the lower half of its dynamic range with the result that the system was "short" of light. A more powerful dye laser coupled with a more sensitive camera would not only increase the amount of light in the system but, also allow smaller changes to be detected and therefore the use of reduced excimer fluences.

4.5.6 Noise Considerations

There are 2 sources of noise in the system, the variation in the output of the dye laser and the noise associated with the electronics i.e. the camera and framestore. The noise in the camera has been discussed in chapter 2 where it is estimated as 1 to 2 greylevels. As already mentioned with the variation in output energy of the dye laser came a slight variation in the spatial location of the beam which could mean that 2 images of an identical surface would not yield a uniformly blank screen if subtracted but one that had a random distribution of speckles on it. Somewhat of a more serious problem is that as the output energy varies, light from certain areas would be detected when the output was high but a small reduction e.g. 10% in output would mean that the light from these areas would no longer be detected and the image would show a dark spot. This would result in the processed image showing significant change associated with these areas indicative of damage but actually due to the simple failure to detect the light due to reduction of the dye pulse energy.

4.5.7 Results 1

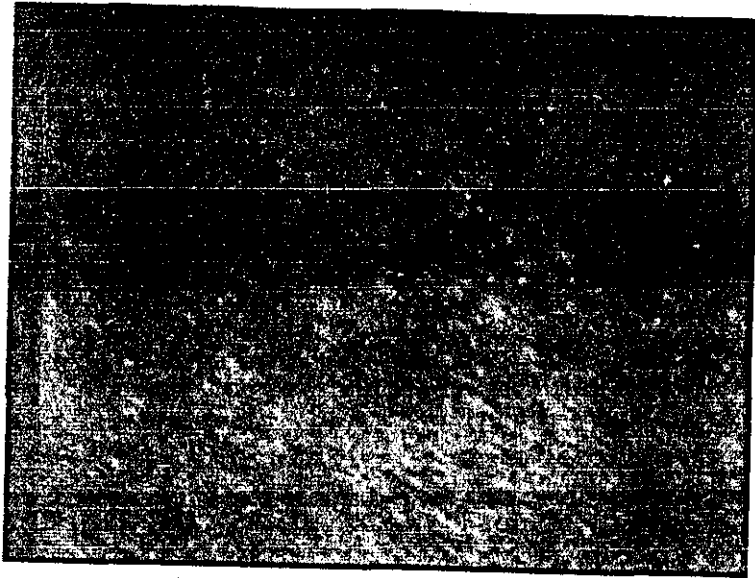
4.5.7a Al single shot

The first set of experiments were performed on one of the Al samples with a damaging fluence of approximately 0.7 J cm^{-2} , a transient recording fluence of $\approx 0.22 \text{ J cm}^{-2}$ and a delay of 375 ns between the excimer and dye pulses, the results of which are shown in figs 4.8a-e. Comparison of figs 4.8a and c shows that at the fluence used, the surface damaged over the width of the field of view and vertically between the points marked by the arrows. Closer inspection of fig 4.8c reveals a darker area within the 2 bright lines which at first sight possibly seems not to have damaged, this appearance is due to using Schlieren based imaging. A Schlieren system is only sensitive to discontinuities in the optical properties of a surface and thus only areas near to such discontinuities will show up as bright patches in the image whilst areas of uniformity will show up as dark patches. Hence the changeover from non-damaged to damaged areas shows up brightly but areas that are either uniformly damaged or undamaged will be imaged as dark patches. Processing of figs 4.8a and b yields the net transient signal (fig 4.8d) and that of figs 4.8a and c yields fig 4.8e the net damage. It is obvious from figs 4.8d and e that the correlation between the transient and damaged images is quite poor. This poor correlation is attributed to excessively damaging the surface with the excimer.

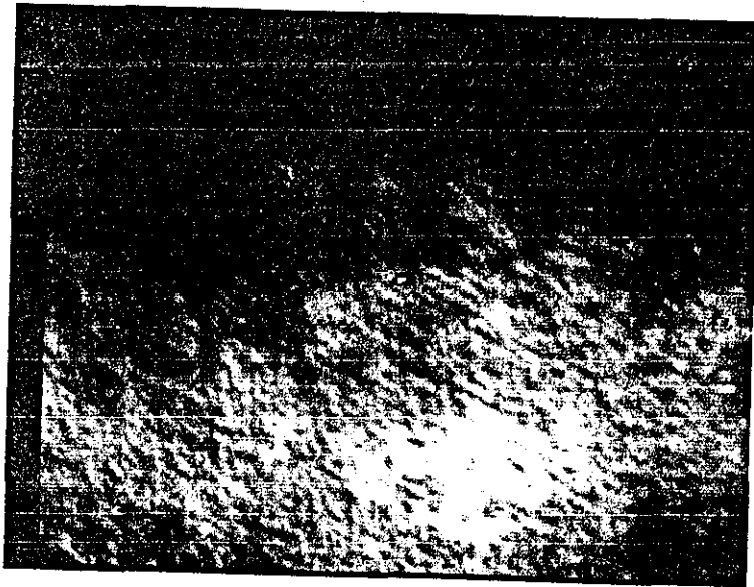
The damaging fluence was gradually reduced to a value of 0.56 J cm^{-2} at which level, when the surface damaged it did so at isolated points which then enabled comparisons between the transient and damaged images to be made. A fluence of 0.18 J cm^{-2} was used to record the transient effects. The time delay between excimer and dye pulses remained constant at 375 ns.

Typical of the results obtained when the excimer fluence was sufficiently low to cause damage only at isolated points is shown in fig 4.8 f-1 obtained in the high mag (X680) setup. Fig's 4.8f-i show respectively the image of the surface after the cleaning pulses, the image of the transient heating, the surface after the transient pulse and the damaged surface after one damaging excimer pulse. Processing of the difference between figs 4.8f and h to give fig 4.8j reveals a maximum difference between the images of only 1 level which is within the noise of the recorded image. Processing of figs 4.8 f and g to give fig 4.8k shows the net transient signal, where as indicated on the diagram there are several areas of anomalous absorption which show up as a decrease in the processed signal i.e. the area becomes blacker. Processing of figs 4.8f and i to give 4.8l shows the areas that have actually damaged. Comparison of figs 4.8k and l shows that there is excellent

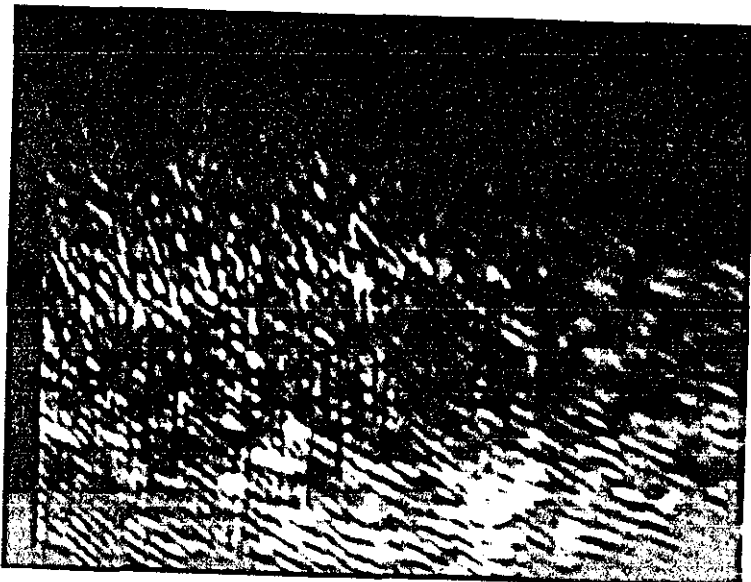
Fig 4.8 (a-e)
Al single shot
high fluence



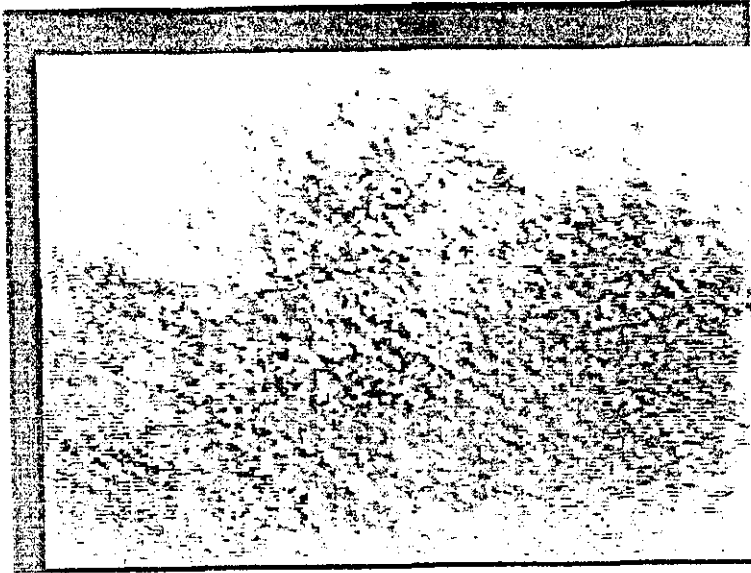
A



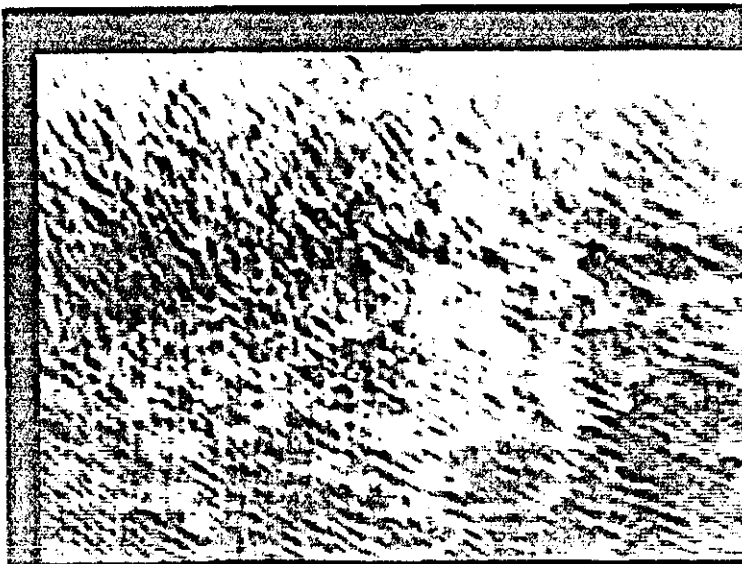
B



C

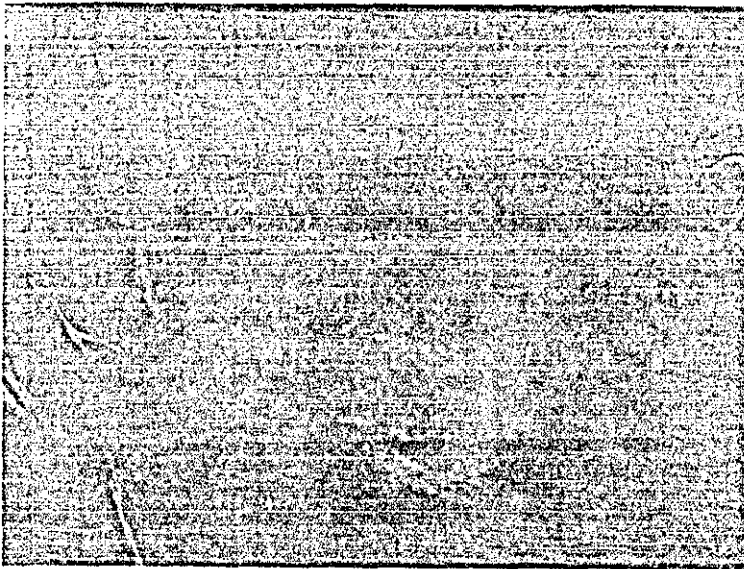


D

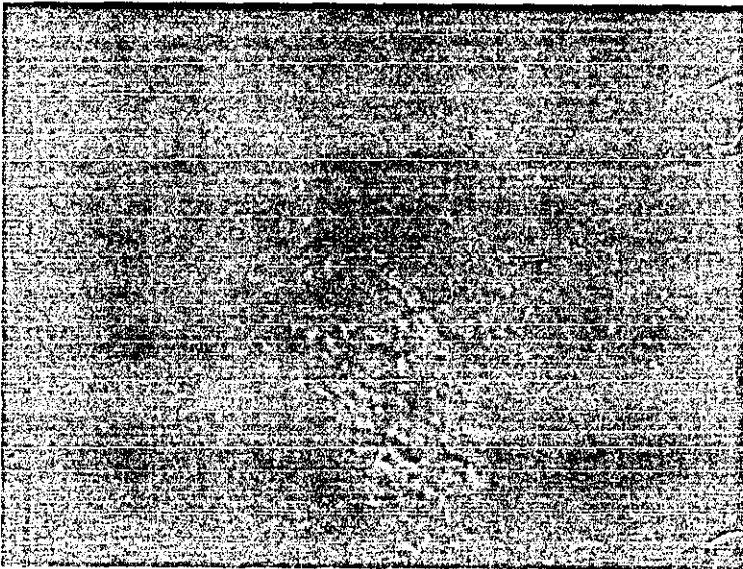


E

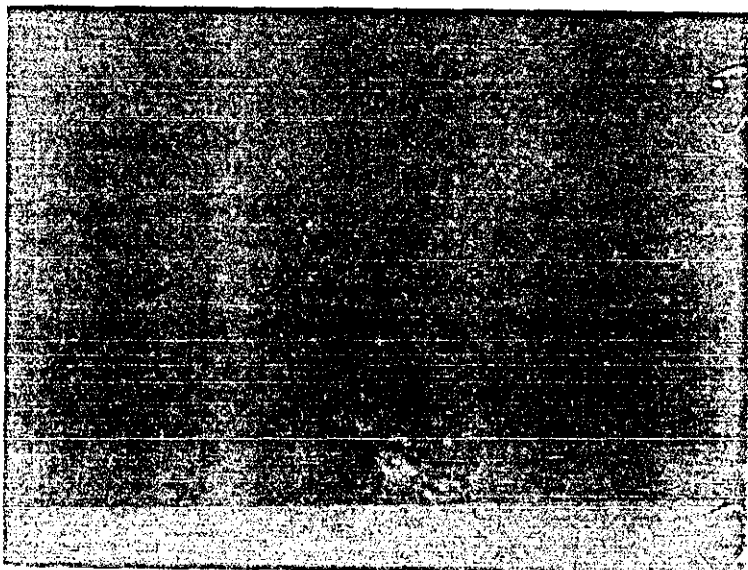
Fig 4.8 (f-1)
Al single shot
low fluence



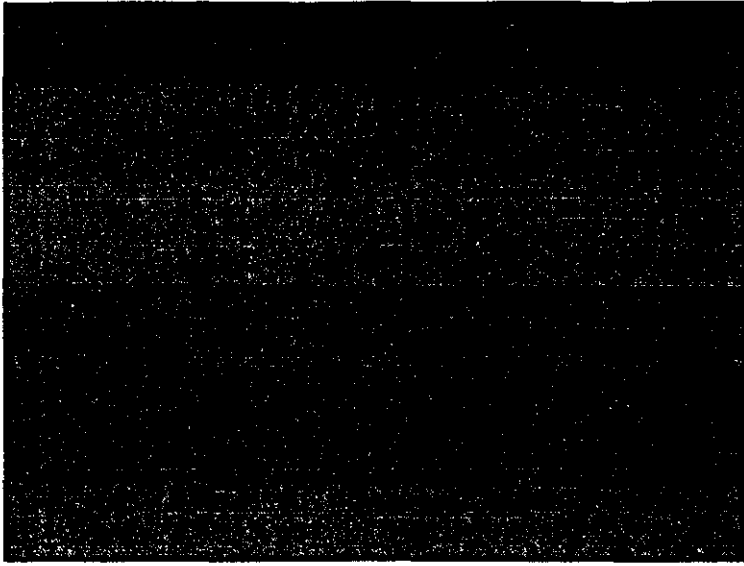
F



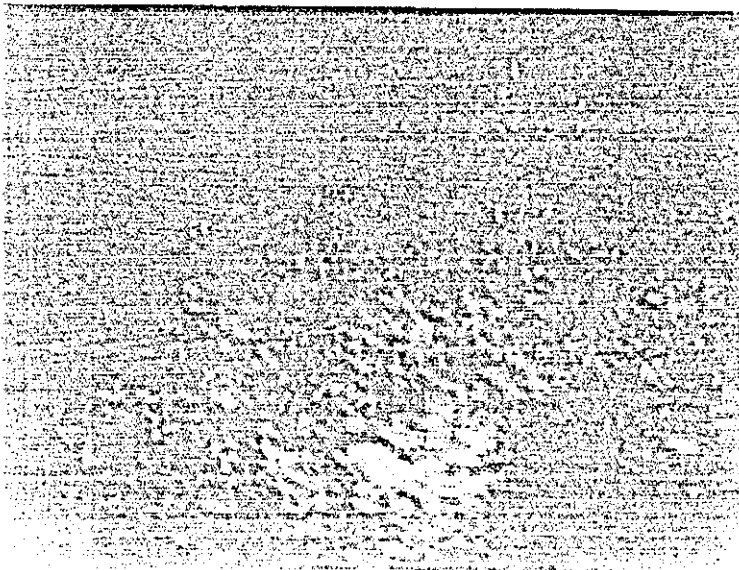
G



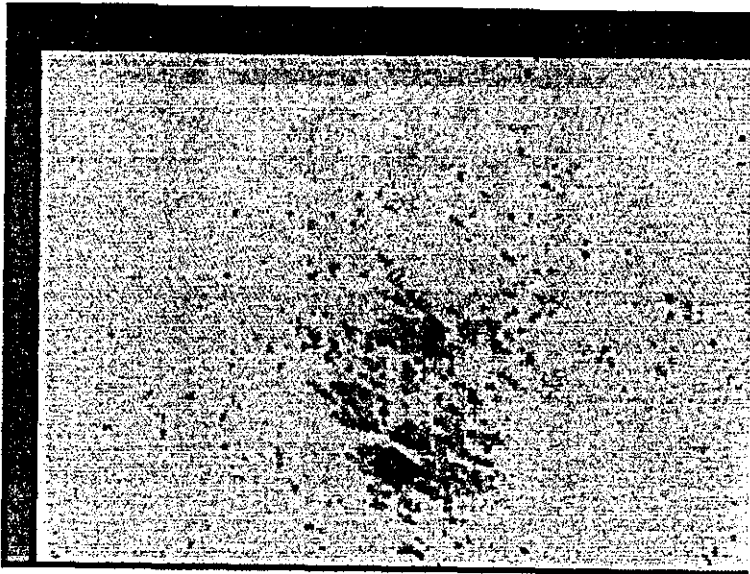
H



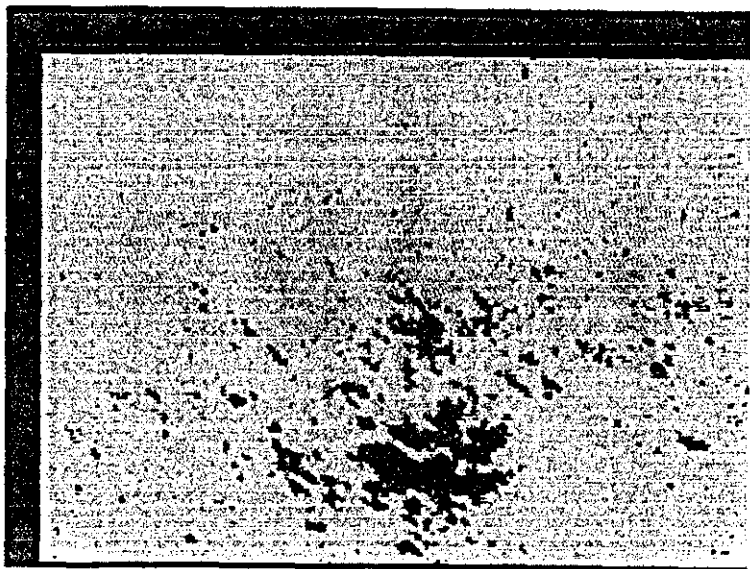
J



I



K



L

agreement between not only the areas of large anomalous absorption and the more heavily damaged areas but between the general location of the damage sites and the somewhat less black areas of the transient image and it is encouraging to see that in these secondary areas, the "amount" of damage is less than that predicted and seen in the blackest areas of the images. The fact that damage occurs over areas that are well separated from each other particularly in the vertical direction over which the excimer intensity can vary significantly indicates that the damage has been caused by the presence of localised defects on the surface rather than by thermal induced changes on a uniform surface.

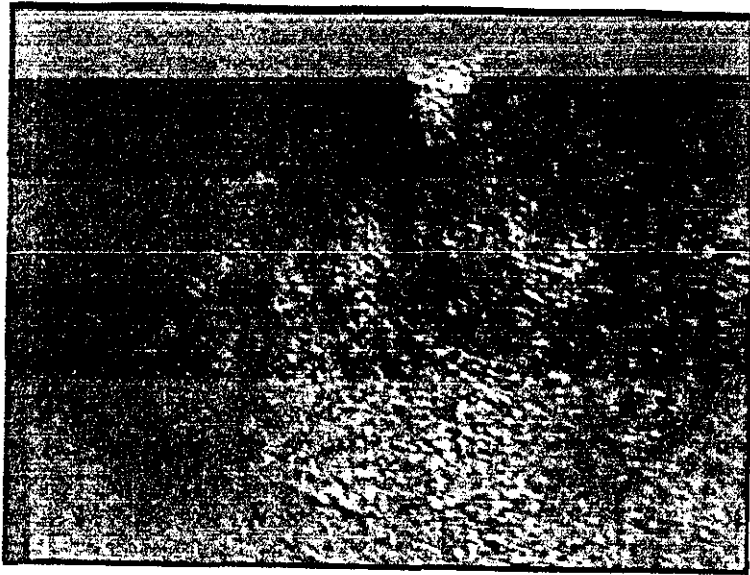
4.5.7b A1 multiple shot

The excimer fluence was gradually reduced to a value of 0.27 Jcm^{-2} at which point it would require approximately 10 shots to damage the sample, the transient effect was still recorded at 0.18 Jcm^{-2} with a delay of 375 ns between excimer and dye pulses. Figs 4.9a-e show a sequence recorded at such excimer fluences in the low mag setup with fig 4.9a being the initial surface, 4.9b the transient effect and 4.9c the surface after 10 excimer pulses. Processing of fig 4.9a and b to give the net transient effect (4.9d) and 4.9a and c to give fig 4.9e the net damage, shows the excellent correlation between the areas of greatest absorption in the transient signal and the first areas to damage.

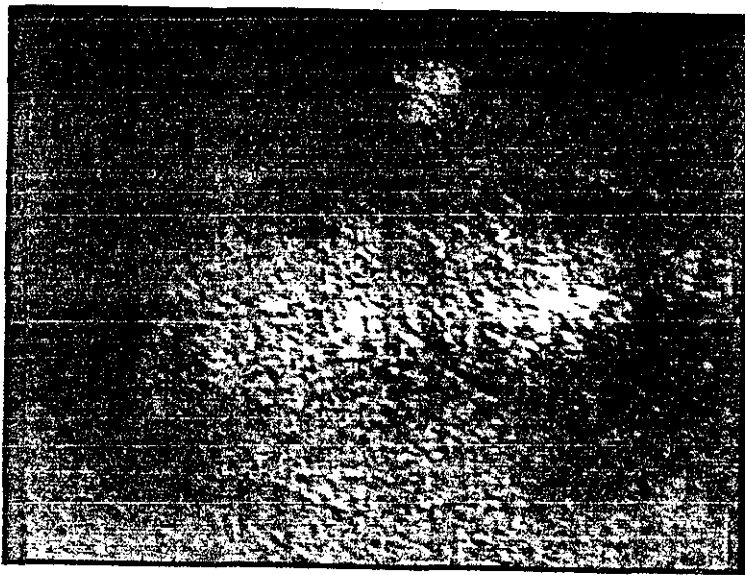
The damage in this case initiates along a line in the centre of the field of view corresponding to the most intense part of the excimer beam but it is important to note that the damage occurs in 2 separate areas with no damage in between. The lack of damage vertically (in which direction the excimer intensity decays rapidly in the low mag setup) away from the centre implies that the damage is being caused by thermal effects but gives no information about damage at defects. More significantly, the fact that in the horizontal direction the damage occurs in two well separated areas with nothing between when the spatial profile of the excimer is known to have no localised hot spots in it is indicative of damage due to the presence of defects ("weak spots") on the surface.

Further reductions in the damaging fluence were made until it was equal to the fluence used to record the transient image (0.18 Jcm^{-2}) at which point it required approximately 50 excimer pulses to damage the surface of the sample. Figs 4.10a-j obtained in the low mag (X260) setup show a typical multi shot sequence at this fluence, with fig 4.10a being the initial surface, 4.10b the transient effect and 4.10c the surface immediately after recording the transient pulse. Comparison of fig 4.10c which was recorded at

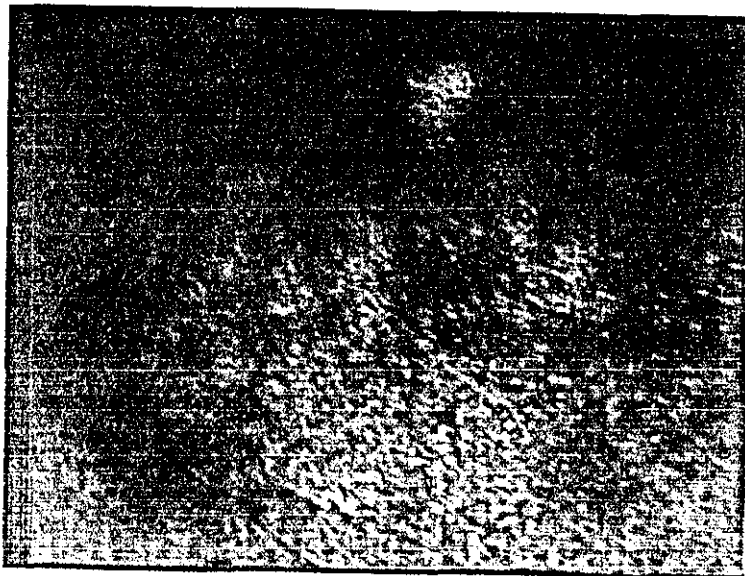
Fig 4.9
Al multiple shot



A

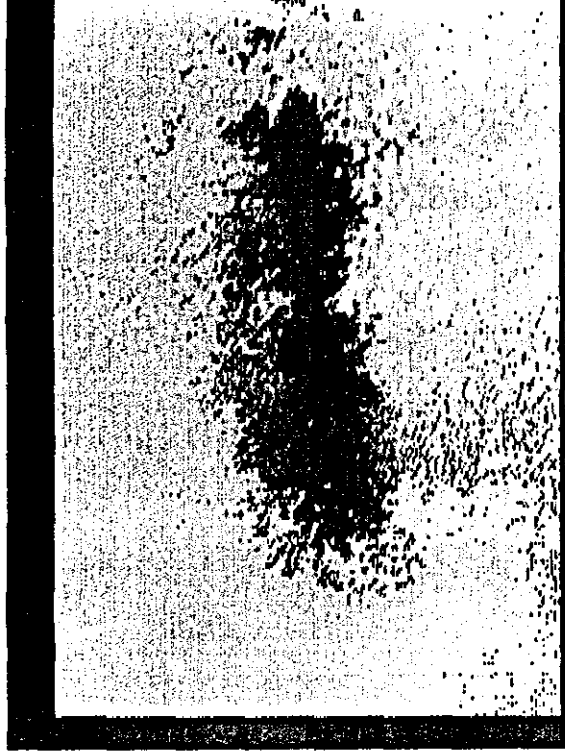


B

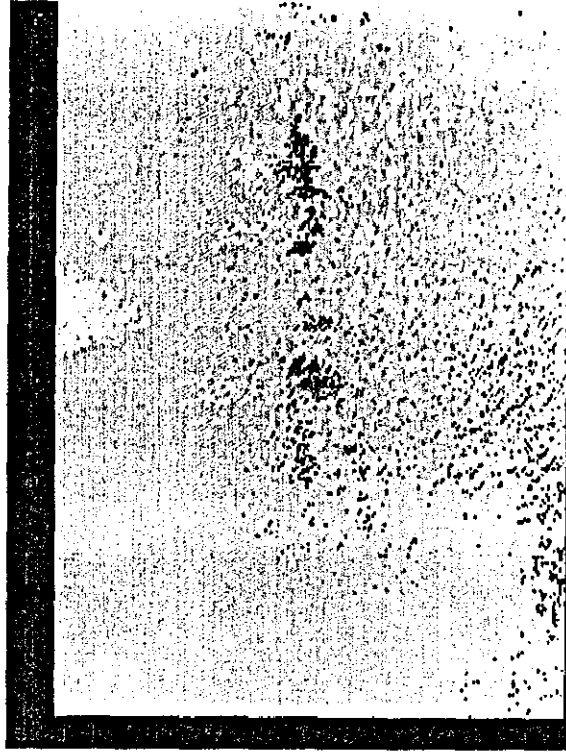


C

Fig 4.9

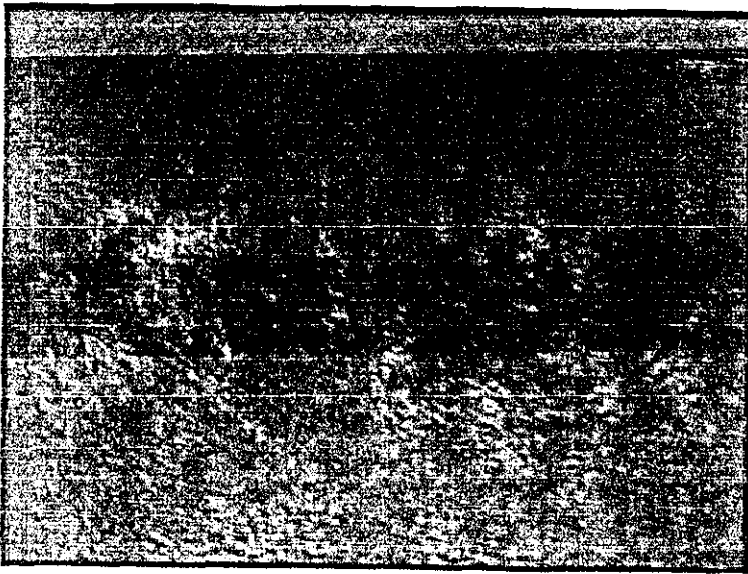


D

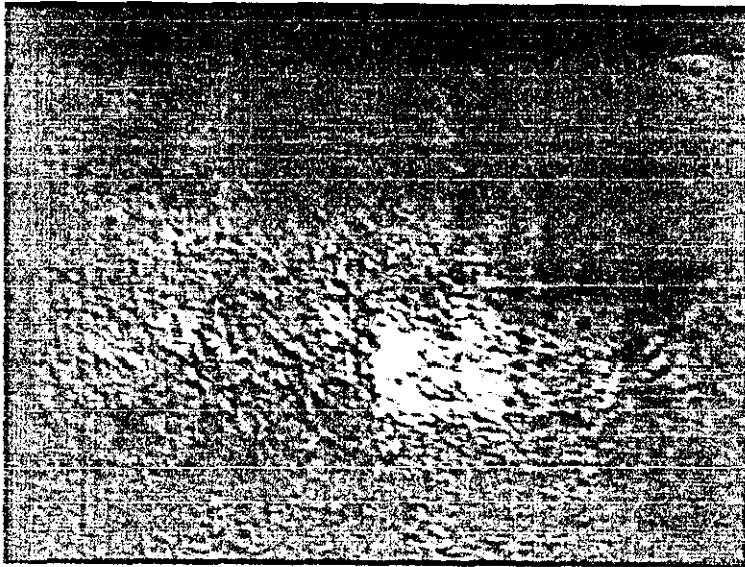


E

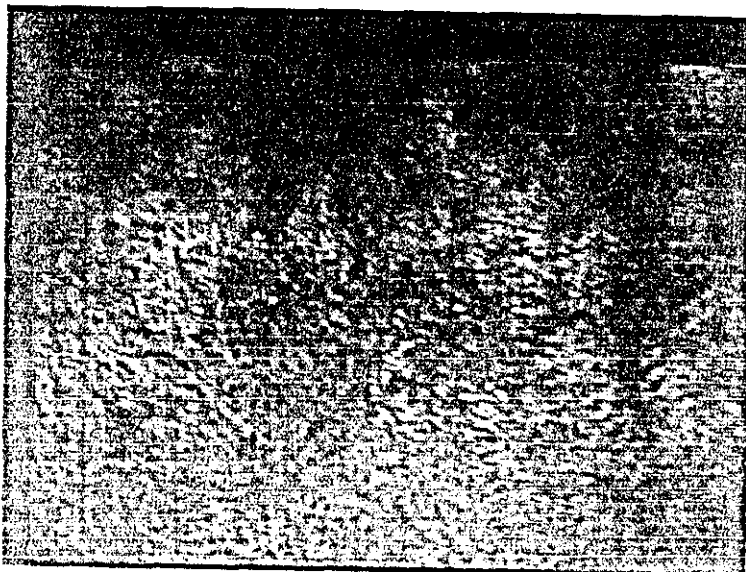
Fig 4.10
Al multiple shot
low fluence



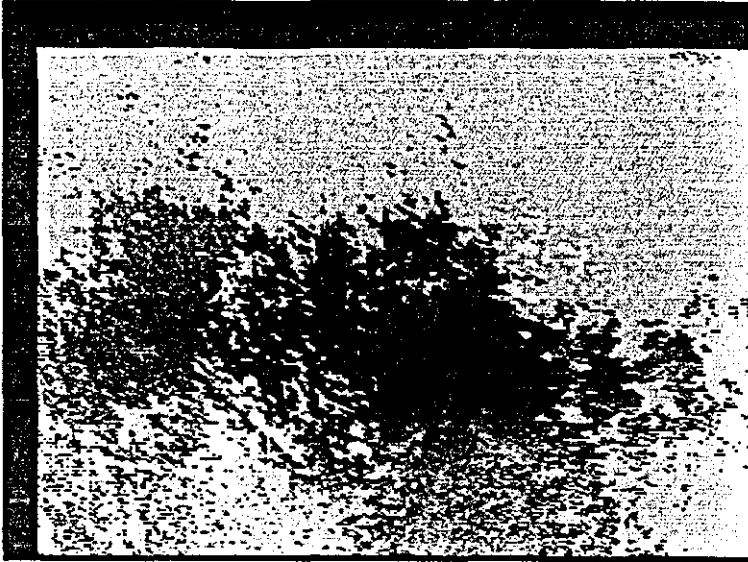
A



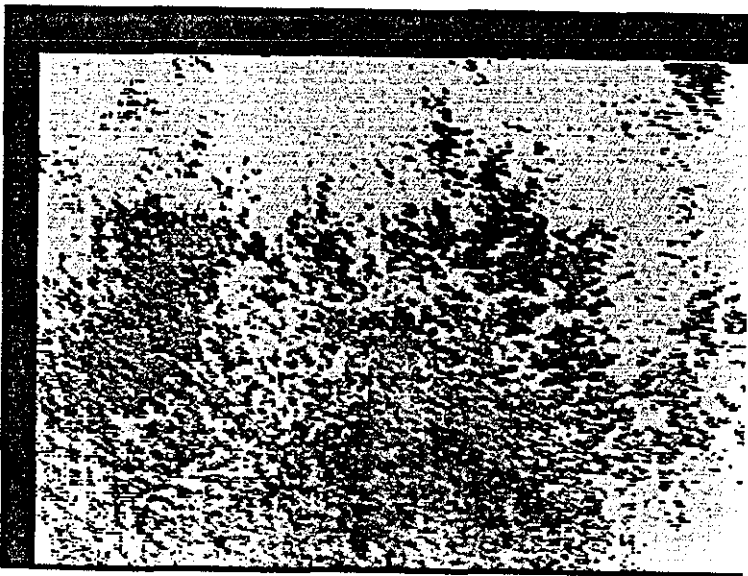
B



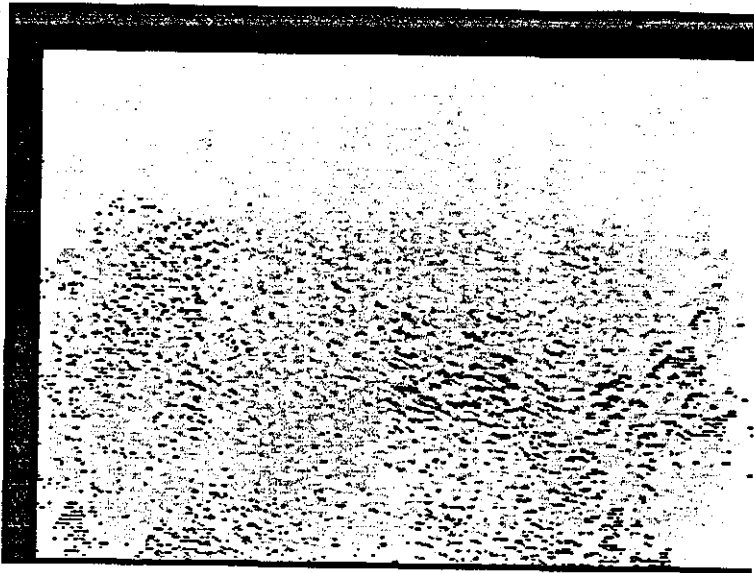
C



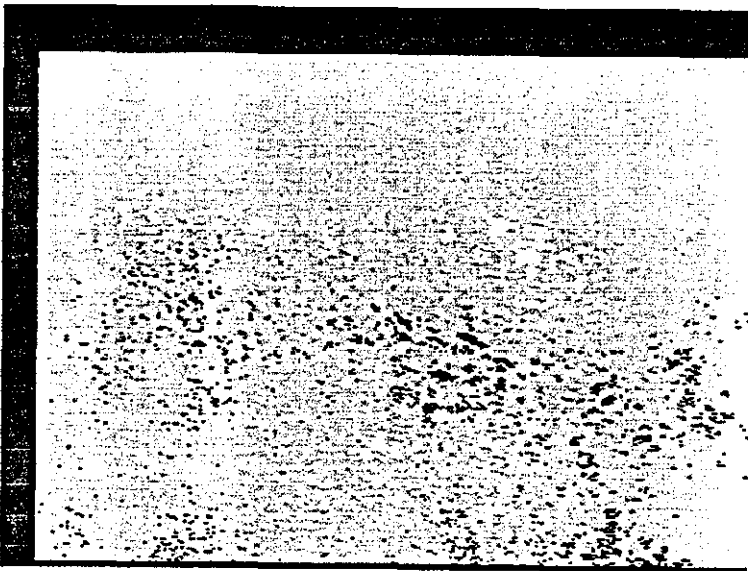
D



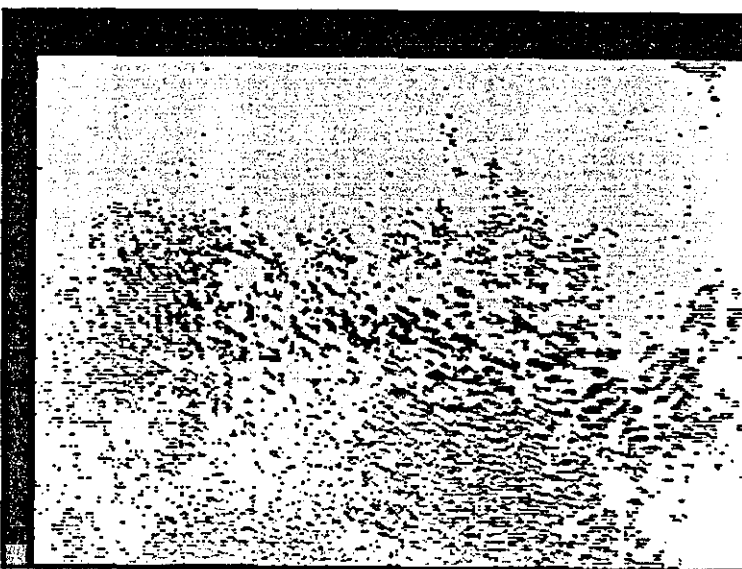
E



F



G



H



I



J

≈15% more energy than 4.10a shows immediately the problem of new areas of light (e.g. those circled in fig 4.10c) being detected simply due to a change in the dye pulse energy between shots. Processing of figs 4.10a and c giving fig 4.10d shows that in the centre of the field where both the laser fluence and detected transient effect are greatest no damage has occurred and that the areas of change between 4.10a and c are simply due to a change of dye pulse energy. This conclusion is reinforced by the extra light that appears in the top right hand corner of 4.10c which in real space corresponds to an area where the excimer beam is $<1/100$ of that at the centre where no damage was found.

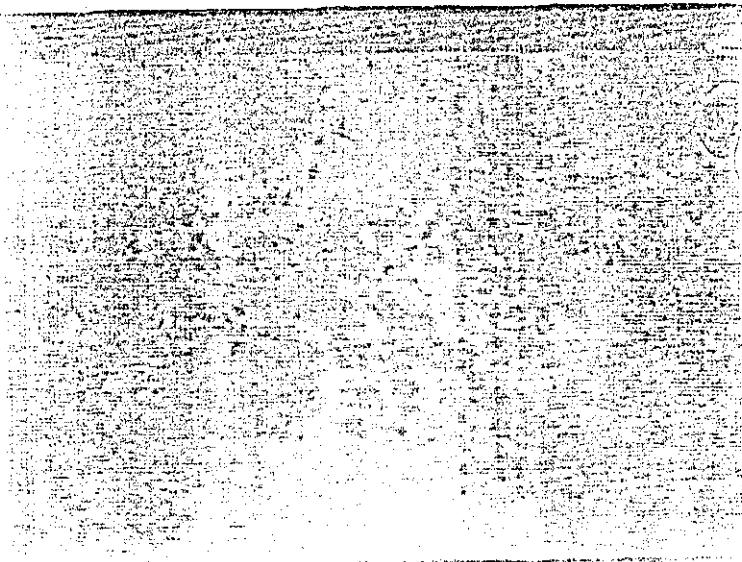
Figs 4.10 f-j are respectively the processed difference between the initial surface and 50,100,150,200 and 400 excimer pulses incident on the sample surface. The sequence shows the gradual development of the damage site from a few isolated areas in the centre to a heavily damaged area with dimensions corresponding to that of the transient effect. It can be seen (fig 4.10f) that the first stages of damage occur in the centre areas that according to the transient effect (fig 4.10e) shows the greatest heating effect and that the damage as it propagates across the field of view generally first occurs in those places that on the transient image are darkest.

The relatively poor correlation between the areas of damage on the left hand side of fig 4.10j and the transient (4.10e) is probably due to a form of positive feedback whereby once initiated at a particular site damage can more easily increase around this area and propagate outwards than it could on an as yet undamaged site.

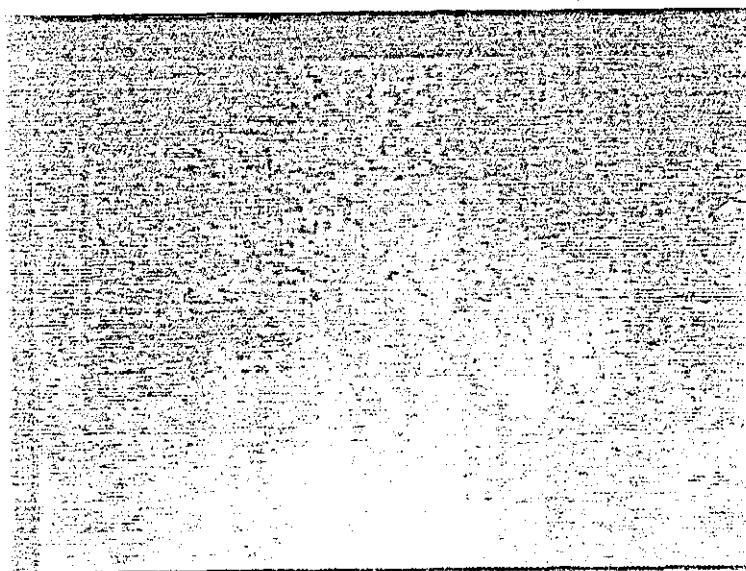
4.5.7c Ge single shot

The Ge sample was first pre-cleaned by 5 excimer pulses at 0.03 Jcm^{-2} and then the transient effects were imaged using an excimer fluence of 0.06 Jcm^{-2} and a time delay of 375 ns between excimer and dye pulses. The sample was then damaged at a fluence of 0.18 Jcm^{-2} and a further image recorded. Fig 4.11 a-e recorded in the high mag setup show a sequence at this fluence.

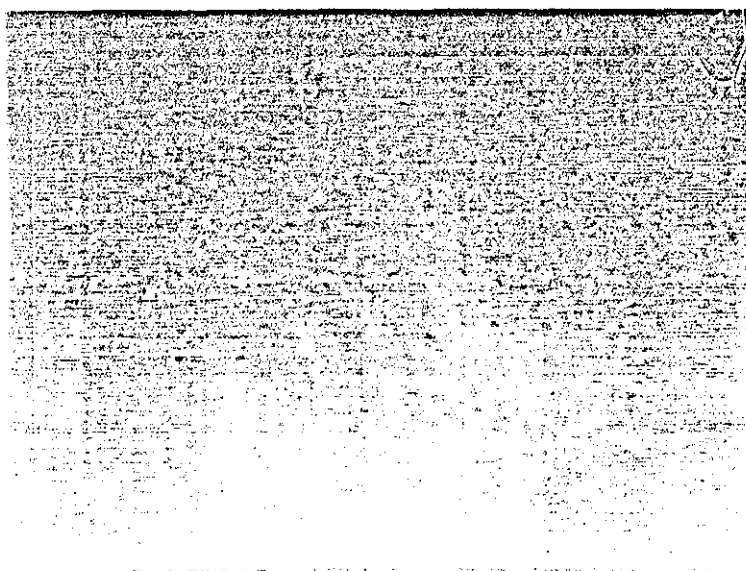
The greylevels found in the images were all quite low in the range 0 -16 and so do not show up well on the video copier prints as this is only a 16 level device which essentially means it takes 4 greylevels to correspond to one level on the printout and thus the actual surface images 4.11a-c are not very clear on the printouts. After processing to give the net transient signal (4.11d) and the actual damage (4.11e) excellent agreement is seen between the areas of anomalous absorption in the transient and those areas that actually damaged. The fact that the damage is essentially in isolated areas particularly well spread out in the vertical direction in which the



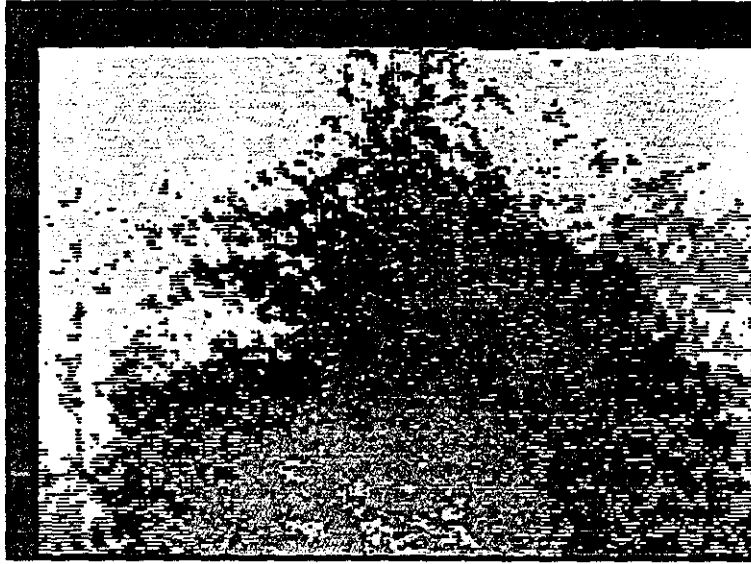
A



B



C



D



E

excimer fluence varies by about a factor of 2.5 over the field of view shows that defects are again the prime cause of damage.

The energy required to damage the Ge sample was approximately 1/3 that for the Al which is consistent with the assumption that damage is thermal in nature since, at 249 nm, Ge absorbs around 3 times as much light as Al, and thus (with the further assumption of similar overall thermal properties) should damage at about 1/3 the fluence of Al.

4.5.7d Dielectric single shot

Although the dielectric sample used was an HR at 249nm it was transparent at visible wavelengths thus requiring the system to be rearranged to work with the dye laser light transmitted through the sample as opposed to the $\approx 4\%$ reflected of the front surface which initial experiments showed was undetectable.

Earlier work had shown that these samples had a single shot threshold of around 3.5 Jcm^{-2} so initially the fluence for transient recording was set to $\approx 1 \text{ Jcm}^{-2}$ but this proved insufficient to produce a detectable transient effect. The fluence was gradually increased to a value of 1.7 Jcm^{-2} at which point there was a detectable transient signal.

The transient effect on the sample was then recorded in the high mag setup at a fluence of 1.7 Jcm^{-2} and a delay of 375 ns between the excimer and dye pulses before the sample was damaged at 3.5 Jcm^{-2} and re-imaged. Figs 4.12a-g show such a sequence with 4.12a being the initial surface b the transient effect c the surface after the transient pulse and d after one damaging excimer pulse.

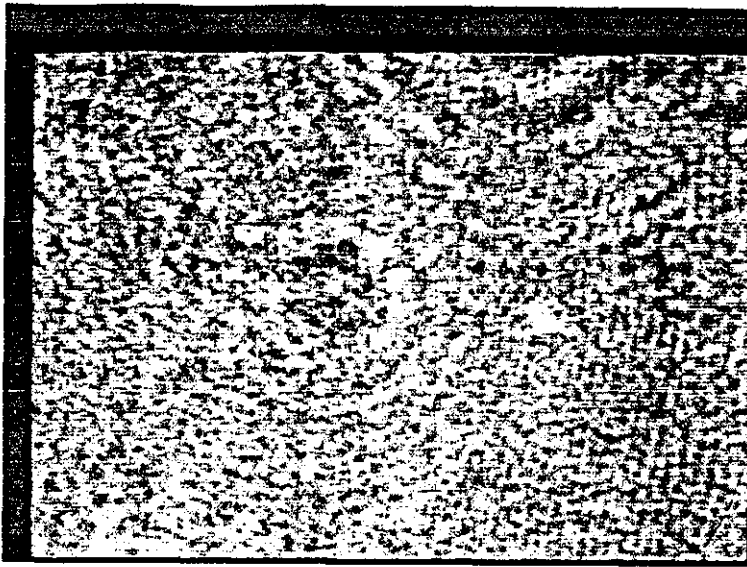
Processing 4.12a and c to give 4.12e shows an essentially uniform image with only the slightest indication in the area marked as 1 that a small change occurred although no change was detected on the surface as a result of the excimer pulse when viewed in real time on the second camera. There is absolutely no indication of a change in area 2 which is where most of the damage occurs.

Figs 4.12a and b are processed to give 4.12f the net transient effect and 4.12a and d to give 4.12g the net damage. Initial comparison of 4.12f and g reveals excellent agreement between the area (marked as 1 in fig 4.12f) of anomalous absorption in the transient and the area of main damage (marked as 1 in fig 4.12g). A more detailed study of 4.12f and g reveals the transient signal shows definite effects at the sites of the 3 secondary areas of damage (labelled 2-4) although they are not as well defined as that at area 1.

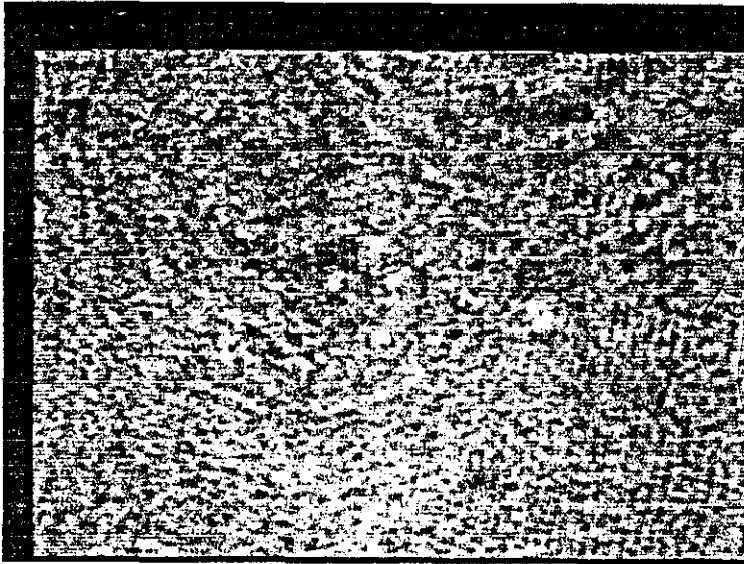
The nature of the damage on this sample i.e. very isolated areas of point

Fig 4.12

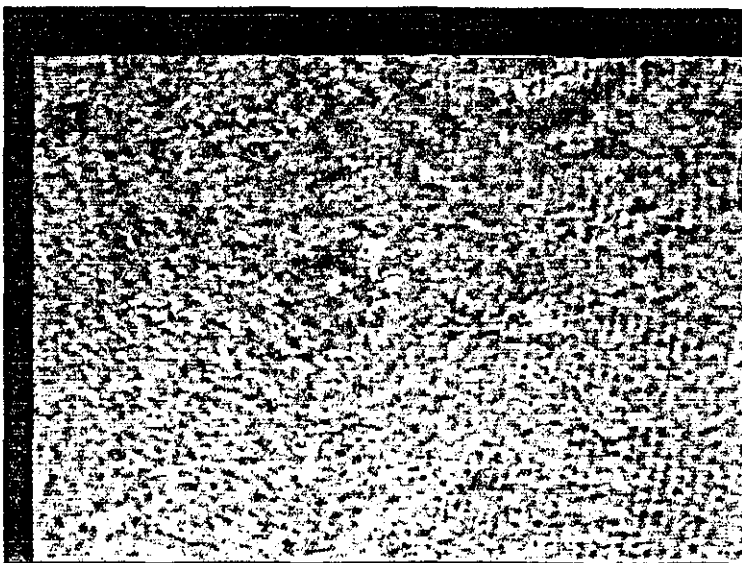
Dielectric
single shot



A

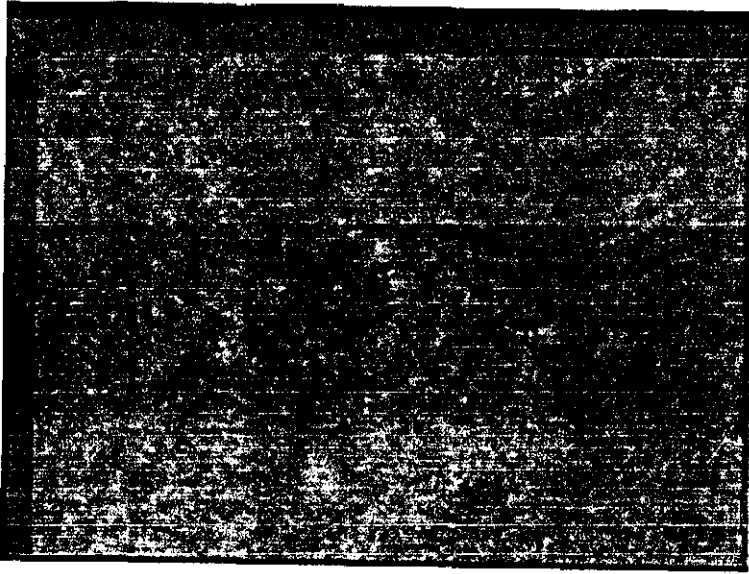


B

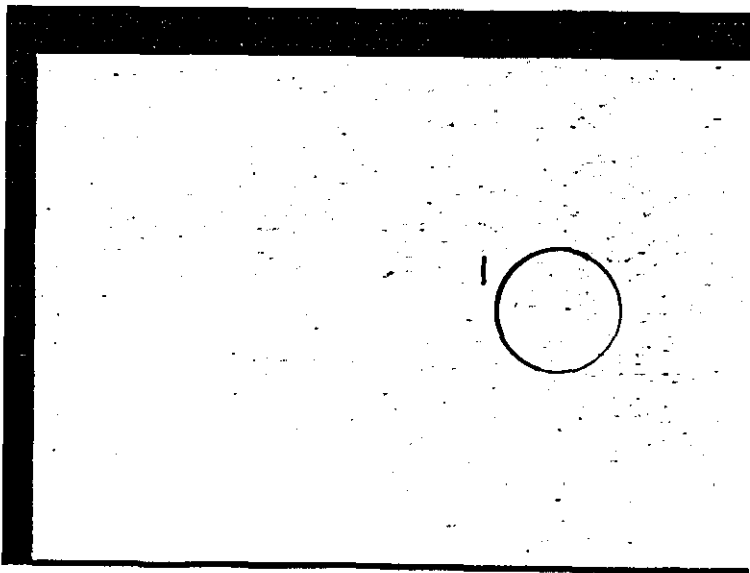


C

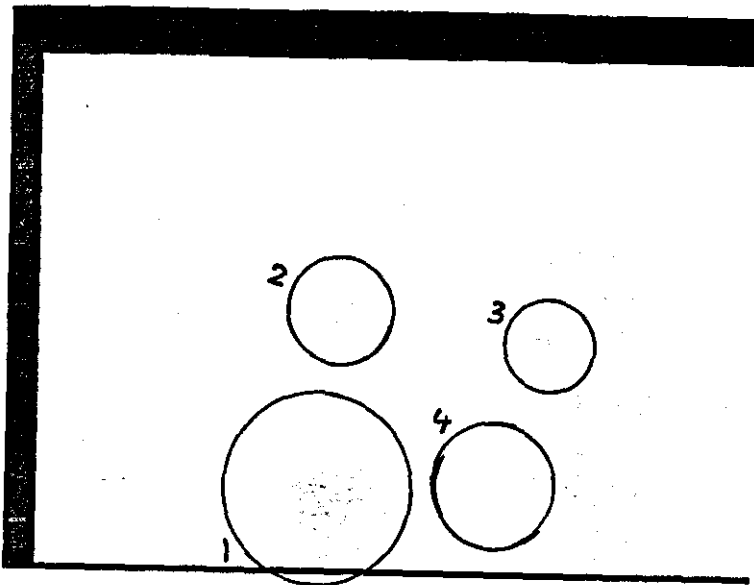
Fig 4.12



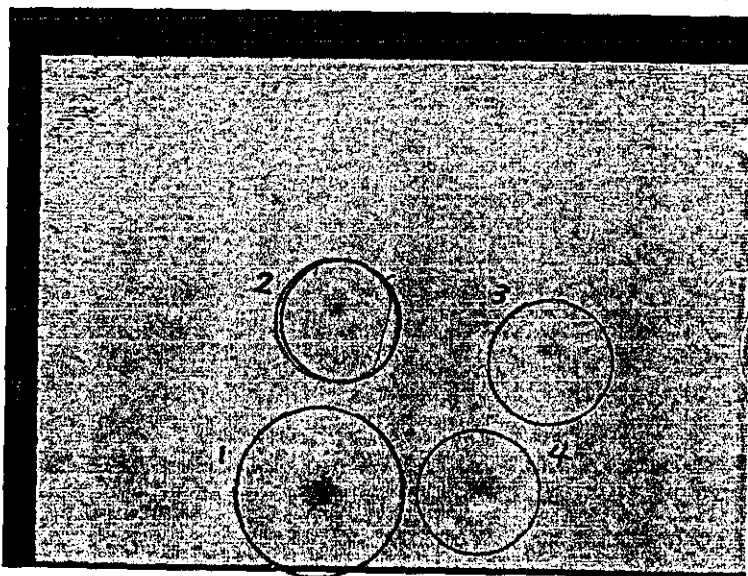
D



E



F



G

damage very strongly indicates that thermal failure through the presence of surface defects such as scratches, voids, inclusions etc is the main cause of damage on these samples.

4.5.8 Time Dependence of Transient Effects

The results obtained in 4.5.7 were all obtained at a delay of around 375 ns between excimer and dye laser pulses but this was subject to variation of up to 40 ns which would make comparison of the transient shots on the same site impossible as the magnitude of the transient has significant variation over the time delay range 350 to 410 ns. To overcome the effect of the jitter the transients should be recorded at a long enough time delay such that 40 ns of jitter has negligible effect on the magnitude of the signal. To determine if this was possible an investigation into the time duration of the transient effect was made.

Two types of experiments were performed to determine the time dependence of the transient heating on Al samples....

1) Use of a photomultiplier set to record the transient deflection (heating effect) of a cw HeNe laser beam that had been arranged to illuminate the same area on the sample as the dye laser if it were used.

2) Use of the video system and dye laser to record images of the transient effect as a function of the delay between excimer and dye pulses.

4.5.8a Time Dependence - Photomultiplier monitoring

A second 10 mW cw HeNe laser was used to provide a beam that illuminated the same area on the sample as the dye laser did. After reflection from the surface the beam was passed through a suitable lens with a knife edge in its focal plane. A suitable fast photomultiplier was situated immediately behind the knife edge so that only the deflected light would be incident on the front surface of the photomultiplier. An Al sample was used as the test surface and was pre cleaned in the standard way with the time duration signals being recorded at the excimer fluence used to image the transient effects (0.18 Jcm^{-2}).

Figs 4.13 a and b show typical traces of transient signal (vertical axis) versus time (horizontal axis). It should be noted that the photomultiplier produces a signal that becomes more negative as the light level increases. From 4.13a it can be seen that detectable signal persists for $\approx 40 \mu\text{s}$ although a more realistic measure of the signal is the time to the half peak point which is of the order of $10 \mu\text{s}$. Reference to fig 4.13b shows that

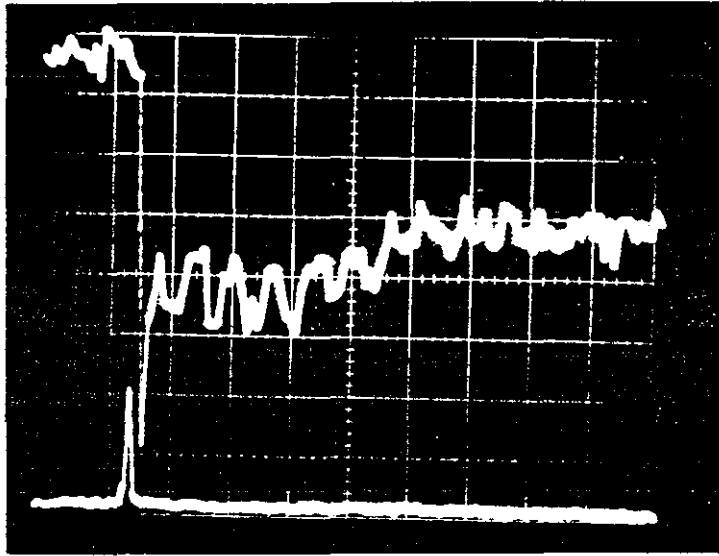


Fig 4.13b
time base 200 ns/div

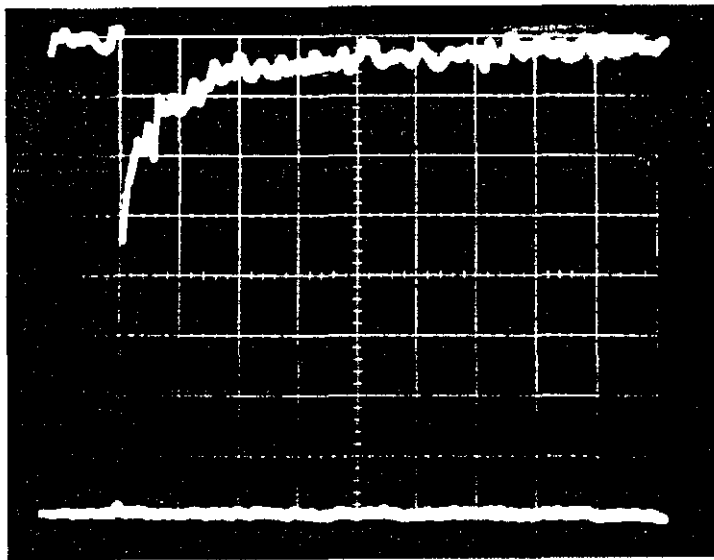


Fig 4.13a
time base 10 μ s/div

the peak of the signal occurs almost immediately after the end of the excimer pulse (lower trace) and after a rapid initial decay during the first 200 ns the decay rate slows drastically and the signal is almost constant from ≈ 800 ns to $2\mu\text{s}$. Thus for a delay of around $1\mu\text{s}$ between excimer and dye pulses the transient signal is essentially constant and therefore insensitive to the 40 ns of jitter present in the system, thereby making it possible to accurately compare transient images at this delay.

4.5.8b Time Dependence of the Transient Image

Sequences of transient images as a function of the delay between excimer and dye pulses on a given site were recorded. Figs 4.14a-f show such a sequence recorded at delays of 30 ns, 80 ns, 320 ns, 950 ns, $3\mu\text{s}$ and $9\mu\text{s}$ respectively. They quite clearly show that the transient is very large at small delays but from ≈ 320 ns to $1\mu\text{s}$ the magnitude of the signal does not change appreciably but that after $\approx 10\mu\text{s}$ the signal has all but decayed.

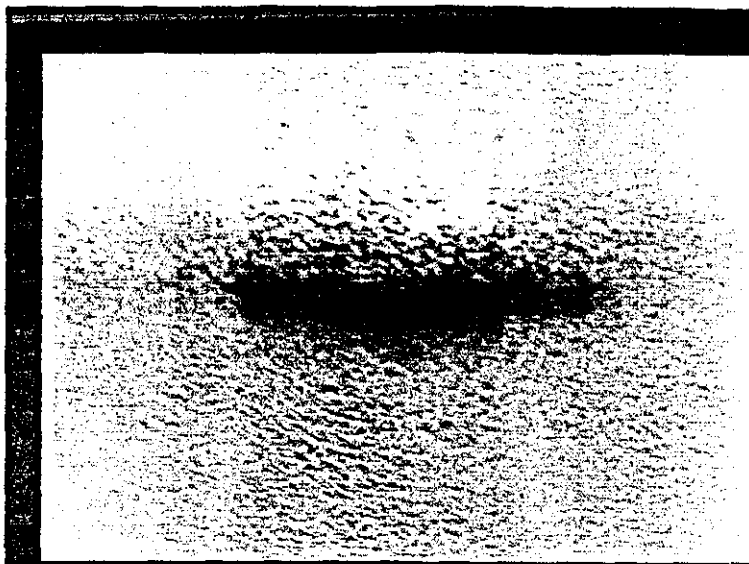
4.5.9 Results 2 - Transient Variations

The variation of the transient signal as a function of the number of incident excimer pulses was investigated using a sample of Ge with a delay of $1\mu\text{s}$ between excimer and dye pulses. The excimer fluence was set at 0.18 Jcm^{-2} i.e. the same fluence was used to record both the transient image and, after a sufficient number of pulses to damage the sample. The initial transient image was processed to yield the areas of high absorption. By means of a computer program a the co-ordinates of a number of these points e.g. 10 were stored in memory. Transient images were then recorded every 5 or 10 excimer pulses and the magnitude of the signal at the stored co ordinates was compared to the initial transient signal.

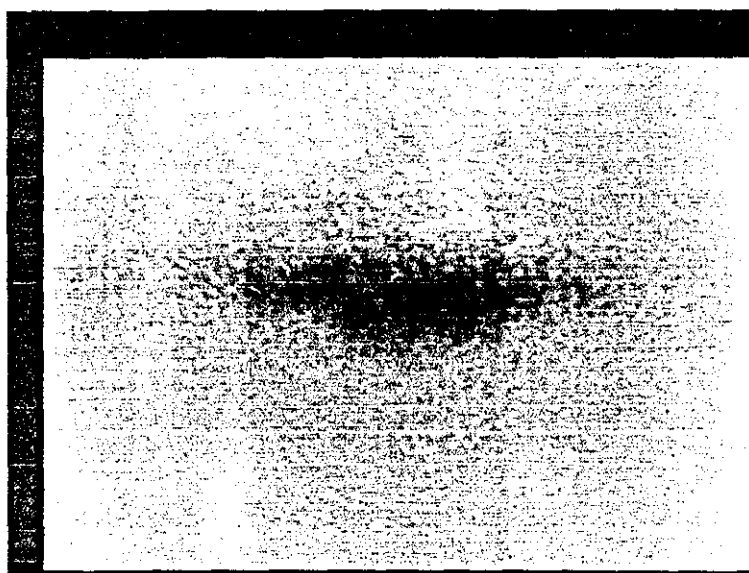
In all cases the same trend in the transient signal as a function of the number of excimer pulses was found i.e. the transient signal would increase with increasing excimer pulses up to the first point that damage occurred where it would decrease substantially. Further excimer pulses would cause the transient to either dramatically decrease or increase.

Fig 4.15 shows the data for several points from a typical sequence. The first row concerns the initial image of the surface and can be ignored for this discussion, the second row corresponds to the initial transient image of the surface. The column headed deviation shows the net change (after allowing for the energy variation shot to shot of the dye laser) in greylevels of the transients from the initial steady state image of the surface. The last column headed ratio is the ratio of the transient effect after n shots and that at the start after 1 shot and represents the

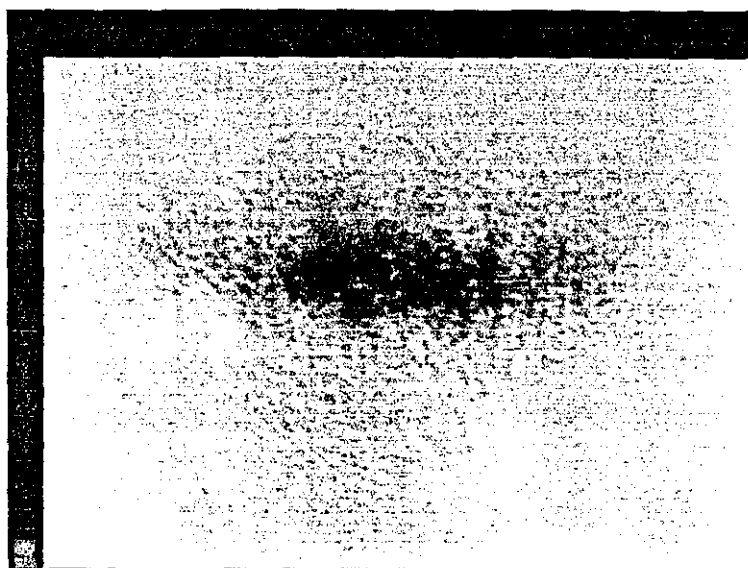
Fig 4.14
Transient as a
function of
delay time



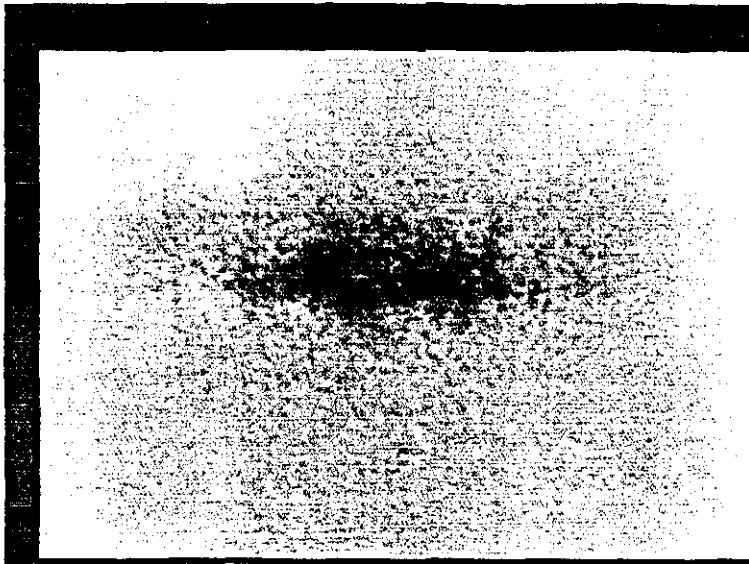
A



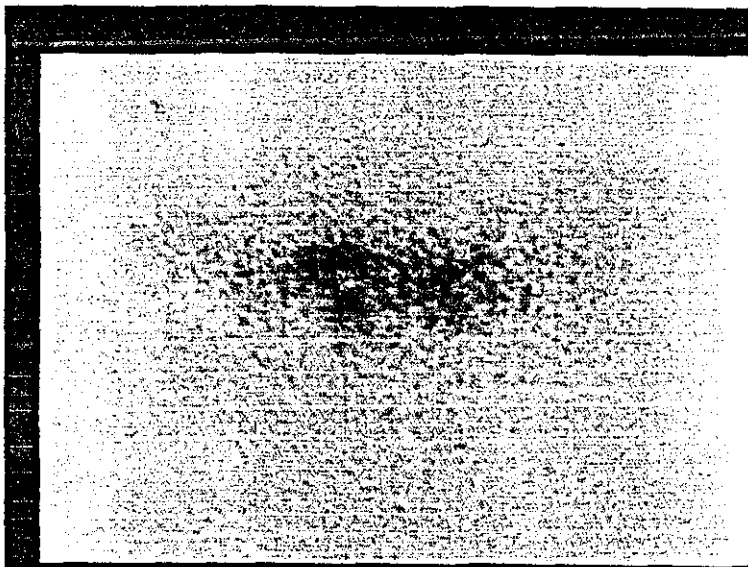
B



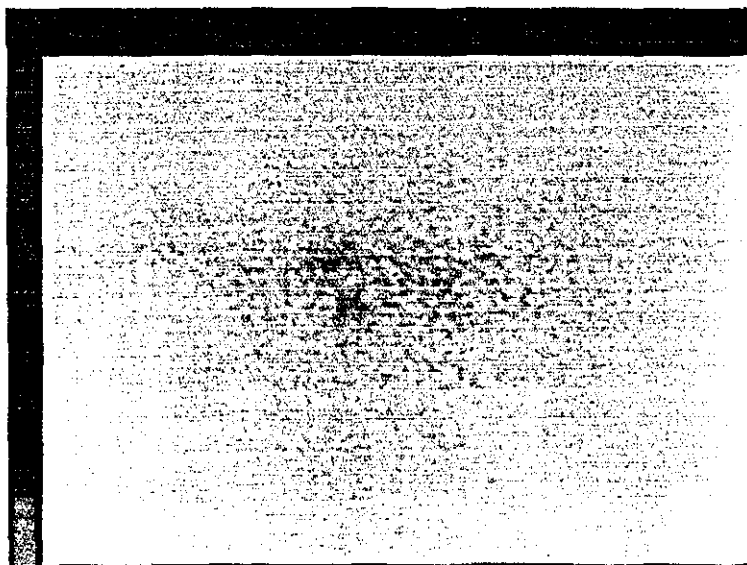
C



D



E



F

Fig 4.15

point 1

Dicrimination level set to ... %

No. shots.....	Deviation.....	ratio..
0	6.222222	1
1	1.333333	1
10	1.733334	1.300001
20	3.333333	2.500001
30	1.692307	1.269231
45	1.333333	1
55	1.111111	.8333336
65	-.6666665	-1.5

type any key to continue..

point 2

Dicrimination level set to ... %

No. shots.....	Deviation.....	ratio..
0	7.777778	1
1	1.777778	1
10	2.933332	1.649999
20	4.4	2.474999
30	1.435898	.8076924
45	4	2.25
55	3.666667	2.0625
65	2.444445	1.375

type any key to continue..

point 3

Dicrimination level set to ... %

No. shots.....	Deviation.....	ratio..
0	6.333334	1
1	2.111111	1
10	3.066667	1.452632
20	3.733333	1.768421
30	2.102564	.9959514
45	3.333333	1.578947
55	4.222222	2
65	3.333333	1.578947

type any key to continue..

fractional change in the transient effect with increasing numbers of excimer pulses.

For each of the points the transient signal as recorded after 10 and 20 excimer pulses has increased but after 30 pulses has decreased. Detailed analysis of steady state images obtained following each transient image show that there was no detectable damage on the images after 10 or 20 excimer pulses but that there was damage present after 30 pulses.

At the time of writing it had not been possible to determine if the transient signal simply increased until damage occurred or if it started to decrease before a permanent change occurred on the surface, in which case monitoring of the transient signal would provide a way to determine how many more pulses a sample could survive before damaging and further work into this needs to be done.

4.6 The Physical Basis of the Deflection of the Probe Beam

Subsequent to the investigation of the time dependence of the transient heating an investigation was made to attempt to determine the contribution to the deflection of a probe beam that each of the following made....

- 1) The heated surface itself
- 2) The re-radiation of heat from the surface into the surrounding air
- 3) The small volume of plasma that is always present when damage occurs

In order to distinguish between the effects, some of the experiments were performed in a purpose built vacuum chamber capable of reaching a pressure of 10^{-5} torr. The chamber had suitable windows so that the excimer radiation could be focussed onto the sample whilst a cw HeNe laser was used to probe the surface. The experiments of section 4.5.8a were then repeated at suitable damaging and non damaging fluences.

4.6.1 Damaging Fluences at Atmospheric Pressure

The signal recorded by the photomultiplier as a function of the excimer energy was now investigated. Since it was not the alignment laser that was providing the light source for the photomultiplier, it was possible to use the alignment HeNe to monitor the surface in both real time and via the framestore for any signs of damage whilst the variation in photomultiplier signal with excimer energy was being recorded. The signal so obtained increased with increasing laser energy but remained identical in pulse shape

until the damage threshold was reached.

Signals recorded at or very slightly above the damage threshold had a very significantly different pulse shape to those recorded either below the damage level or at the damage level but after enough excimer pulses had been incident on the surface to sufficiently "surface harden" it, so that at the fluence used no more damage occurred. Typical of these signals is fig 4.16 where the upper trace shows the signal recorded during a pulse in which damage was detected whilst the lower trace is recorded on the same site and fluence as the upper one but after 15 shots by which point further excimer pulses produced no further damage. The upper trace has 2 clearly separate peaks marked 1 and 2. For reasons to be justified in the next section the peak in the lower trace is attributed to the heating of the air by re-radiation of heat from the Al surface across the whole of the excimer target interaction area. This gives a moderate sized effect but one that persists for a "long" time.

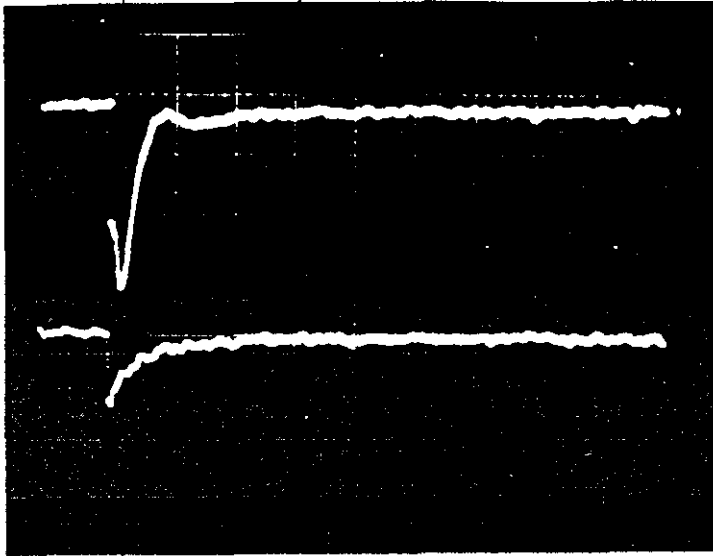
The shape of the upper trace is attributed to the combination of 2 effects, the overall heating from the surface which gives peak 2 and a short lived intense effect which gives peak 1. This short lived effect is attributed to the small volume of plasma that is produced whenever the surface damages. This plasma which is produced by the vaporization of a small fraction of the excimer target interaction area contributes to the observed signal in 2 ways. Firstly the plasma emits its own intense broad band light, of which, only some of the component at 632.8 nm will be detected by the photomultiplier as the light is emitted in all directions. Secondly, the plasma will cause a very intense but due to the small volume of plasma relatively short lived heating effect in the air surrounding the plasma volume, which will give a large deflected signal. The first peak in 4.16 is attributed to the plasma but after $\approx 10 \mu\text{s}$ this effect has decayed sufficiently for the less intense but longer lived heating effect from the whole surface area to become detectable.

4.6.2 Beam Deflection Dependence on Pressure

In order to determine the contribution of the heating of the air to the deflection of the probe beam as discussed in the previous section an Al sample was placed in the vacuum chamber and subjected to damaging and non damaging excimer fluences.

Typical of non damaging excimer fluences is fig 4.17a. Trace a was obtained at atmospheric pressure, the system was then evacuated and trace b obtained and finally after returning the system back to atmospheric pressure trace c was obtained. Since no damage occurred there is only 1 peak in trace a and c

Peak 1 Peak 2

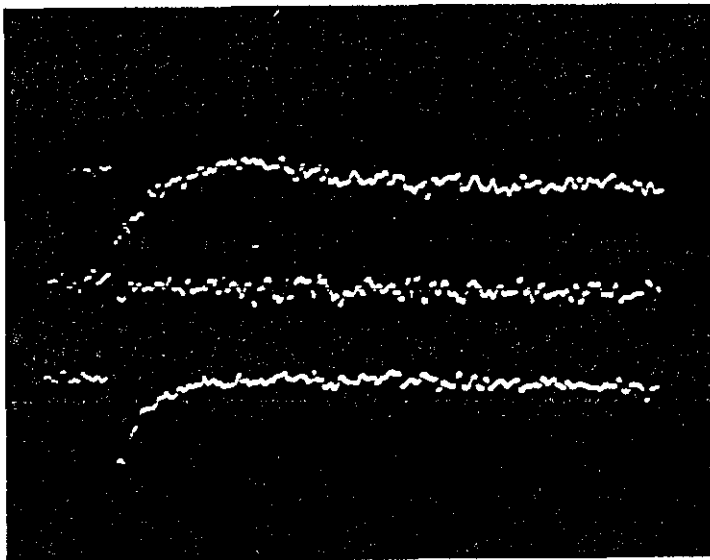


damage

Fig 4.16

no damage

Peak 1

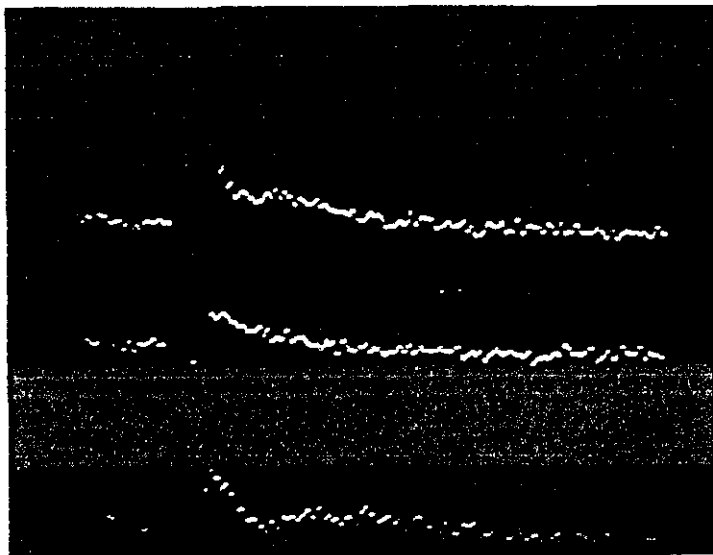


C

B

A

Fig 4.17a



A

B

C

Fig 4.17b

but there is no deflection on the trace obtained in vacuum. This confirms that at least for non damaging pulses and therefore in the absence of plasma formation the deflection of the beam is produced by the re-radiation of heat from the sample surface into the surrounding air and not by the presence of hot areas on the surface itself.

Typical of damaging fluences is fig 4.17b where traces a and c were obtained at atmospheric pressure and trace b in a vacuum. It can be seen that traces a and c have 2 peaks whilst b only has 1. In each of the 3 traces damage together with a visible spark i.e. plasma formation was observed. Thus the first peak in the traces is attributed to the deflection associated with the plasma and the second peak in a and c with the heating of the air which, as there is no air in vacuum, is absent from trace b.

These experiments confirm that for no plasma formation i.e. no damage the deflection of probe beams is due to the re-radiation of heat from the surface of a sample and not the hot areas on the surface itself whilst if damage occurs then most of the deflection can be associated with the presence of a small plasma.

4.6.3 Beam Deflection Theory

A great deal of theoretical work¹³⁻¹⁷ has been done in analysing the propagation of a light beam through heated media, mostly in connection with the photothermal effect with the consequence that the solutions obtained therein to the heat and beam deflection equations are for a periodically varying heat source i.e. a chopped cw laser beam and not for the case of heating by a single laser pulse. As commented on by Jackson et al¹⁶ it is technically possible to obtain the solution for pulsed heating from that of periodic heating by means of the Laplace transform inversion formula but because of the nature of the solutions involved is impossible to perform in practice. The authors¹⁶ derive expressions for the deflection of a beam through an infinitely weakly absorbing media that has been heated by a single pulse. Therefore in the absence of published work a very simplistic theory of the beam deflection due to a pulsed laser was developed by the author with the aim of obtaining at least qualitative agreement with experiment.

The problem in making a theoretical analysis of the beam deflection in the experiments described in this chapter is complicated by the fact that the air is heated by the surface of the sample and not directly by the laser beam. Therefore in order to solve the problem rigorously, the heat conduction equation would have to be solved in both the sample and air simultaneously for a heat source induced on the sample surface by the excimer radiation

which would be extremely difficult to perform.

The approach adopted here is to assume that the surface temperature of the sample is determined by loss to the bulk of the sample rather than by loss to the air which in view of the sample being a metal and good heat conductor whilst the air is a poor conductor is adjudged to be reasonable. This then means that the temperature of the surface can be found by a straight forward solution (i.e. one ignoring the presence of the air on one side of the surface) of the heat equation for a surface source as was discussed in chapter 3. Once the temperature on the surface is known it is used to determine the temperature distribution in the air by assuming the air has no initial temperature distribution but is heated via conduction by the known temperature distribution on the sample's surface and ignoring any heating by convection. Once the distribution in the air is found it is used in an equation that describes the propagation of light in an inhomogeneous medium to determine the beam deflection.

4.6.3a Surface Temperature of the Metal

From chapter 3 the temperature on the surface which is taken as the $z=0$ plane is given by

$$T(r,t)_{z=0} = \frac{2aE}{\pi p_m c_m t_0 \sqrt{\pi k_m}} \int_0^{t'} \frac{1}{(t-t')^{\frac{1}{2}}} \frac{1}{(8k_m(t-t')+w^2)} \times e^{-\left(\frac{2r^2}{(8k_m(t-t')+w^2)}\right)} dt' \quad 2)$$

where $t' = t_0$ $t >$ beam duration
 $= t$ $t <$ beam duration

and p_m, c_m and k_m are respectively the density, specific heat capacity and thermal diffusivity of the metal (see chapter 3 for further details).

For the times to be considered (up to $\approx 10 \mu s$) in this analysis the factor $8k_m(t-t_i)$ is always much less than w^2 and can safely be ignored e.g. for the beam used w^2 is $\approx 10^{-7} m^2$ which would require for the 2 terms to be equal t to be of the order of $10^{-3} s$. Hence ignoring this factor it is possible to perform the integration with respect to time to give the temperature on the metal surface in closed form as

$$T(r,t)_{z=0} = \frac{4aEe^{-\left(\frac{2r^2}{w^2}\right)}}{\pi p_m c_m t_0 \sqrt{\pi k_m} \cdot w^2} f(t) \quad 3)$$

where $f(t) = \sqrt{t}$ $t < \text{beam duration}$
 $= \sqrt{t} - \sqrt{t-t_0}$ $t > \text{beam duration}$

4.6.3b Determination of the Air Temperature

Carslaw and Jaeger¹⁷ show that the temperature T of a solid at a point $P(x',y',z')$ at a time t due to zero initial temperature in the volume of the solid and with the temperature on the surface $z=0$ given by $\phi(x,y,t)$ for any x,y and t is given by

$$T(x',y',z',t) = k_{a0} \int_0^t \left[\iint \phi(x,y,t') \frac{\partial U}{\partial n_i} ds \right] dt' \quad (4)$$

where $ds = dx.dy$ and $\partial n_i = \partial z$

and U is the Green's function for temperature such that $U_{z=0} = 0$ for all t , and is given by

$$U = \frac{z}{8(\pi k_a)^{5/2} (t-t')^{5/2}} \left[e^{-\left\{ \frac{(r-r')^2}{4k_a(t-t')} \right\}} - e^{-\left\{ \frac{(r+r')^2}{4k_a(t-t')} \right\}} \right] \\ \times e^{-\left\{ \frac{-z^2}{4k_a(t-t')} \right\}} \quad (5)$$

which yields the following integral for the temperature distribution in the gas

$$T(r,z,t) = \frac{4aEk_a}{\pi P_m c_m t_0 \sqrt{\pi k_m}} \int_0^\infty \int_0^{2\pi} \int_0^z z e^{-\left\{ \frac{2r'^2}{w^2} \right\}} f(t-t') e^{-\left\{ \frac{z^2}{4k_a(t-t')} \right\}} \\ \times \frac{1}{8(\pi k_a)^{5/2} (t-t')^{5/2}} \times e^{-\left\{ \frac{(r^2+r'^2 - 2rr'\cos(\theta-\theta'))}{4k_a(t-t')} \right\}} \quad (6) \\ \times 2\pi r' dr' d\theta' dt'$$

and thus the temperature in the air as

$$T_a(r, z, t) = \frac{2aE}{\pi p_m c_m t_0 \sqrt{\pi k_m k_a} (\pi)^{3/2}} z e^{-\left(\frac{2r^2}{w^2}\right)}$$

$$\times \int_0^t \frac{f(t-t')}{(t-t')^{3/2}} e^{-\left(\frac{z^2}{4k_a(t-t')}\right)} dt' \quad 7)$$

where p_a, c_a and k_a are respectively the density, specific heat capacity and thermal diffusivity of the air.

4.6.3c Equation of Optical Deflection

Theoretical work has yielded a number of essentially equivalent expressions for the equation governing the propagation of a light beam in an inhomogenous media of which the one developed by Aamodt and Murphy¹⁸ is used here. There expression for the probe beam deflection angle, is given in vector form by the line integral

$$\phi = - \int_P \frac{1}{n} \frac{dn}{dT} \text{ Grad } T \times dL \quad 8)$$

where T is the temperature, n the refractive index and P the path of the probe beam with dL being an incremental distance along P .

The co-ordinate system used was to have the excimer beam propagate in the z direction and be incident normally on the sample which is situated in the $z=0$ plane. The probe beam propagates in the xz plane and is incident on the sample at an angle of $\approx 15^\circ$ corresponding to the angle between the excimer and dye beams in fig 4.7b. For the purposes of the calculation the cross section of the probe beam will be assumed small and any effects it might produce ignored. The knife edge was mounted so that the system was sensitive to deflections in the y direction. Hence the angle of deflection in the y direction is given by

$$\phi_y = - \frac{1}{n} \frac{dn}{dT} \int_P \left(\frac{\partial T}{\partial z} dx - \frac{\partial T}{\partial x} dz \right) \quad 9)$$

Careful consideration as to the path that the beam traverses is required. To first order, following Aamodt and Murphy¹⁸ the beam is assumed to make small deviations from straight line propagation in the xz plane allowing the path P to be written in the form

$$\underline{P} = x\underline{i} + z\underline{k} \quad \text{where} \quad x = mz + c$$

Where $\arctan(m)$ is the angle between the excimer (heating beam) and the HeNe probe beam (see fig 4.7b). The sign of m must change on reflection from the surface since the beam will have crossed the $x=0$ axis as shown in fig 4.18. For propagation toward the sample...

$d\underline{L} = dx\underline{i} - dz\underline{k}$ with $x = -mz + c$ giving the deflection in the y direction as

$$\phi_{in} = -\frac{dn}{dT} \frac{1}{n} \int_0^{\infty} \left(\frac{\partial T}{\partial x} - m \frac{\partial T}{\partial z} \right) dz \quad (10)$$

and for propagation away from the sample.....

$d\underline{L} = dx\underline{i} + dz\underline{k}$ with $x = mz + c$ giving the deflection in the y direction as

$$\phi_{out} = \frac{dn}{dT} \frac{1}{n} \int_0^{\infty} \left(\frac{\partial T}{\partial x} + m \frac{\partial T}{\partial z} \right) dz \quad (11)$$

Substitution of 7) into 10) and 11) yields the following expression for the deflection of the probe beam as a function primarily of time...

$$\begin{aligned} \phi_{total} &= \phi_{in} + \phi_{out} \\ &= \phi(x = -mz+c) + \phi(x = mz+c) \end{aligned} \quad (12)$$

where

$$\begin{aligned} \phi &= \frac{-2aE}{n\pi P_m c_m t_0 \sqrt{\pi k_m k_a} \pi^{3/2}} \frac{dn}{dT} \\ & \int_0^t \int_0^{z_{end}} \left[\frac{-4xz}{w^2} - m \left(1 - \frac{z^2}{2k_a(t-t')} \right) \right] x e^{-\left(\frac{2r^2}{w^2}\right)} \cdot \frac{f(t-t')}{(t-t')^{3/2}} \\ & x e^{-\left(\frac{z^2}{4k_a(t-t')}\right)} dt' dz \end{aligned} \quad (13)$$

A computer program was written to numerically evaluate the above expression

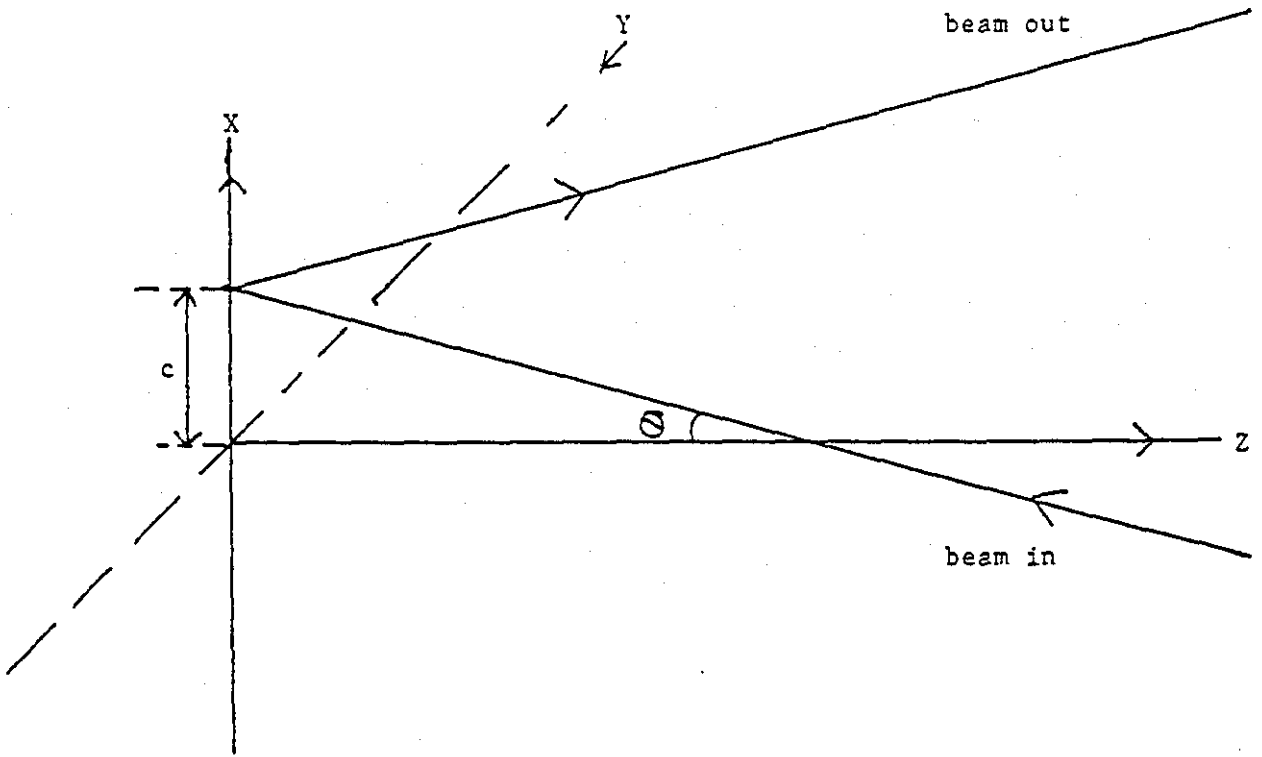


Fig 4.18

inward $X = -mZ + c$

outward $X = mZ + c$

where $|m| = |\tan \theta|$

as a function of time and the constant c which represents the point at which the excimer and probe beams intersect. The evaluation of the integration over z which should ideally be from 0 to ∞ was actually carried out over a range 0 to 100 μm as initial running of the program with various values of z showed that for the times used evaluation for $z > 100 \mu\text{m}$ produced answers identical to that of $z = 100 \mu\text{m}$. A measure of the distance that heat diffuses in a given time is

$$x = (Kt)^{0.5}$$

where t is the time and K the diffusivity.

By the end of the excimer pulse (25 ns) heat will have diffused of the order of 0.5 μm into the air. The sampling distance for the z integration was set equal to this value since the temperature distribution will not be accurately (spatially) to any better than the diffusion length.

4.6.4 Results of Calculations

Typical results of the calculation of the beam deflection for different time scales are shown in figs 4.19a and b. For ease of comparison an experimentally recorded deflection signal on a time scale similar to the one in the graph in each fig is included.

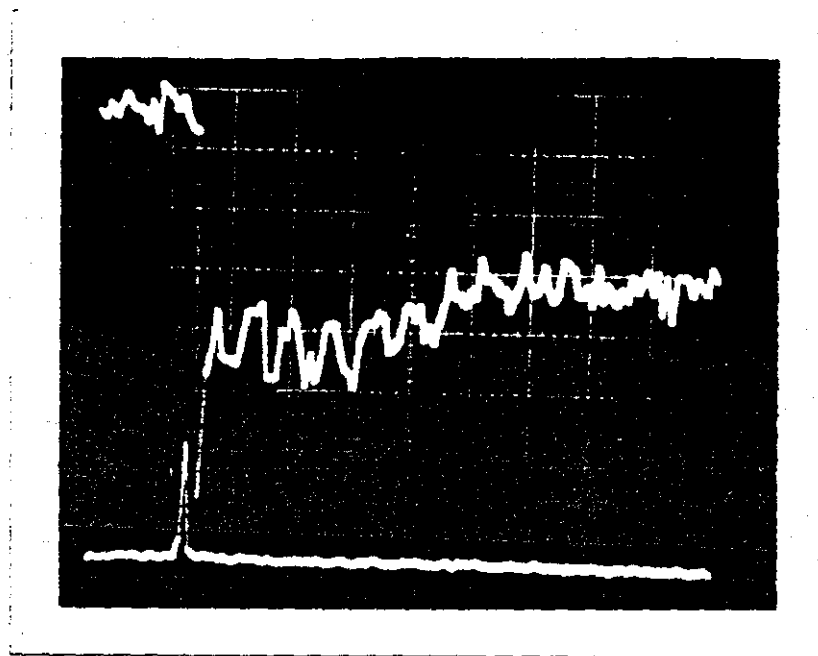
Comparison of the theoretical and experimental signals shows that the calculated signal has the correct gross features in that it has a very fast initial decay which slows down and shows that appreciable signal relative to the peak signal persists for many microseconds. In detail the agreement is very poor with the calculated signal falling off too fast initially and too slowly at later times.

The fact that the detailed agreement is very poor is not surprising since a very crude model has been used but at least it gives a basis for understanding the deflection of the probe beam due to the heating of the sample's surface and subsequently the surrounding air by the excimer laser.

4.7 Conclusions

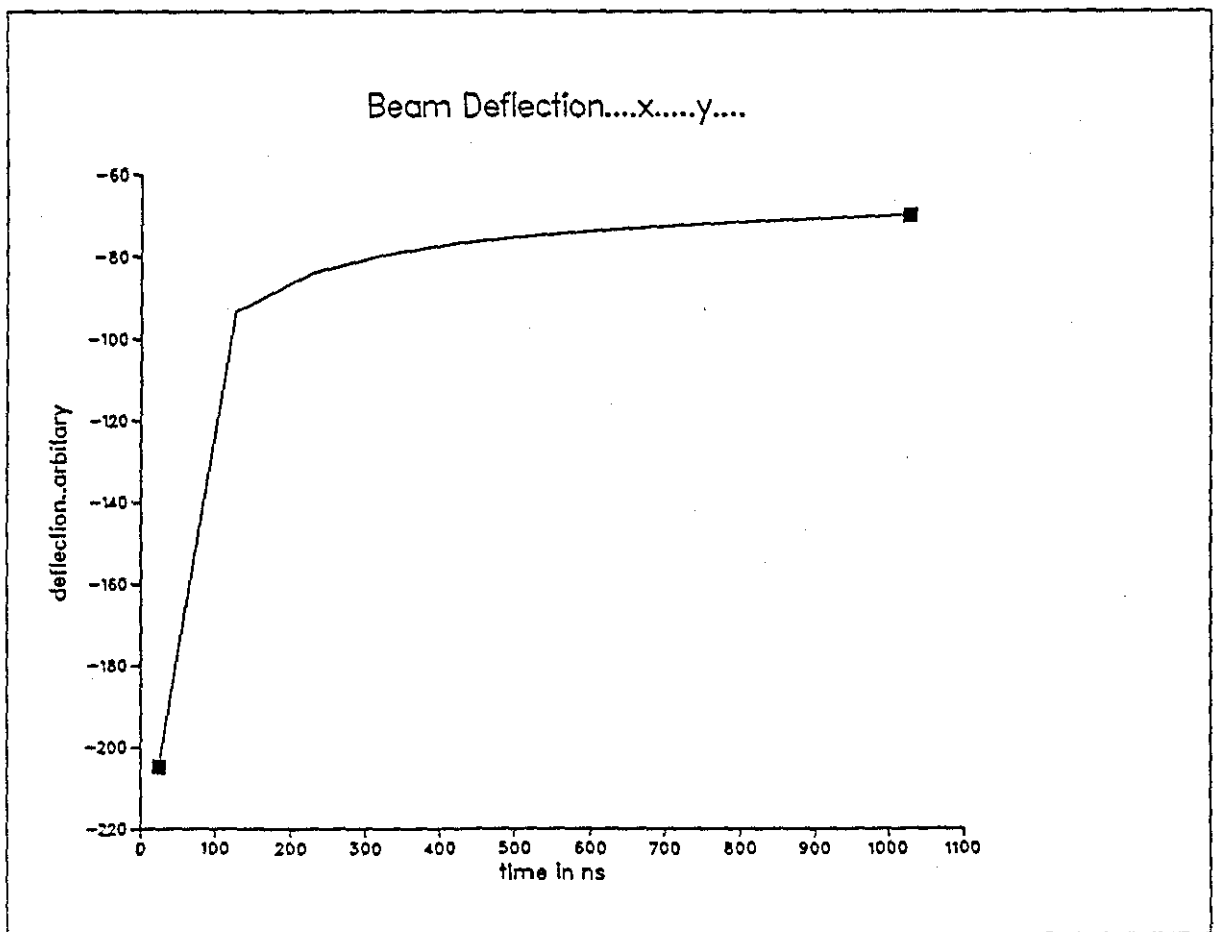
It has been shown that the transient effects and their subsequent probing that are associated with an excimer laser heating a surface can be attributed to 2 mechanisms, the formation of plasma if damage occurs and the heating of the air by the samples surface leading to a deflection of a probe beam. This deflection has at least qualitatively been explained theoretically.

A simple practical technique has been developed for EDLID which gives excellent correlation between predicted and actual damage locations. A more powerful probe (dye) laser would enable the transient effect to be recorded at a smaller fraction of the damage threshold than these results were



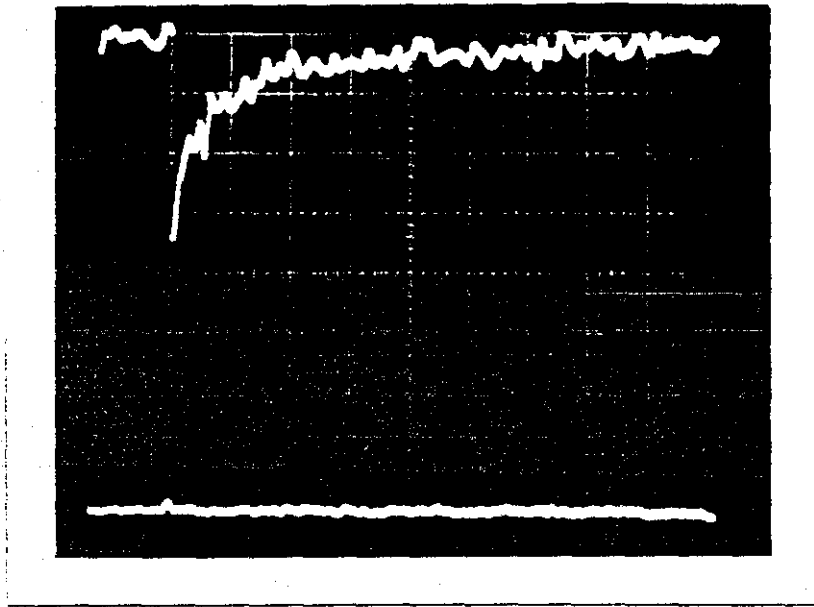
time base
200 ns/div

Experimental



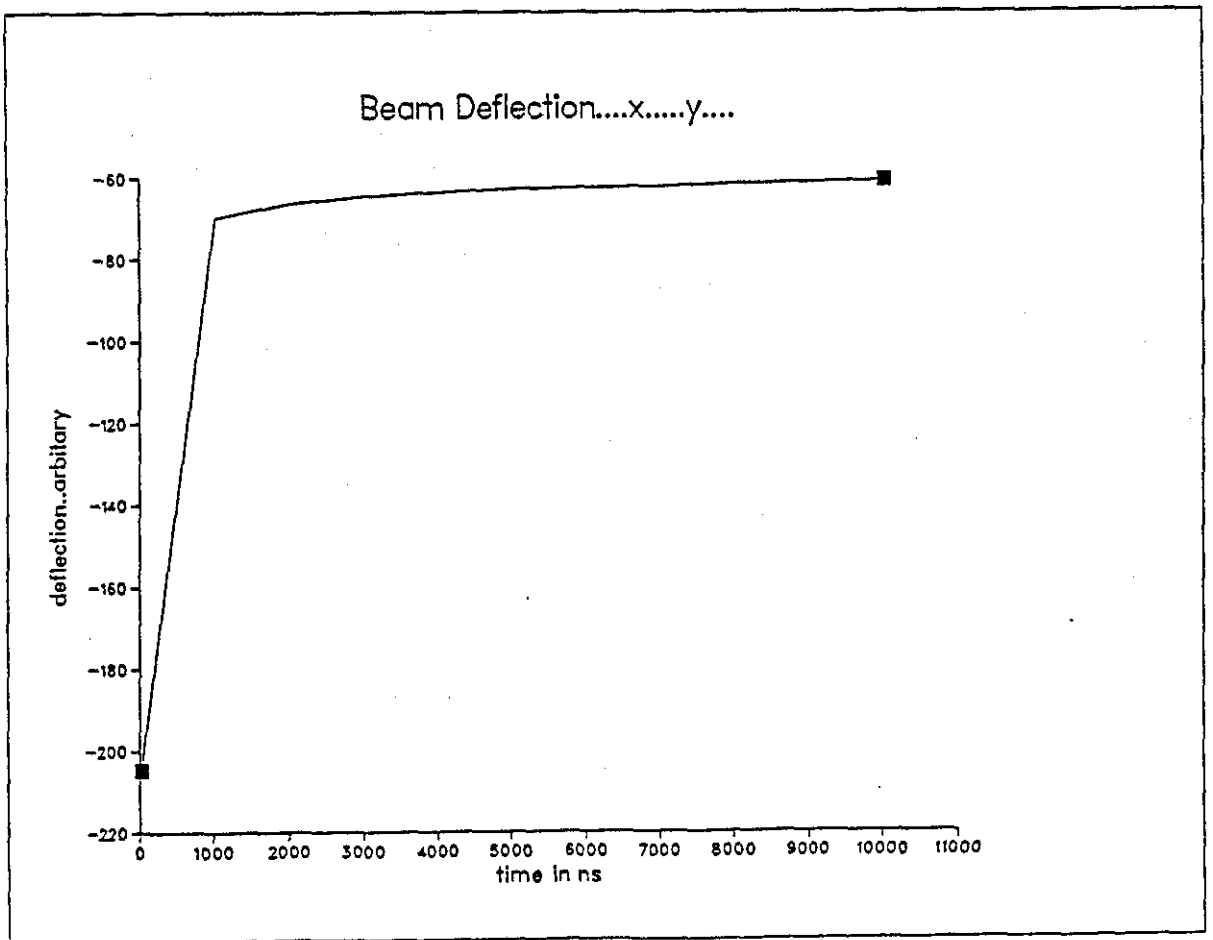
Theoretical

Fig 4.19a



timebase
10 μ s/div

Experimental



Theoretical

Fig 4.19b

obtained at thereby improving the sensitivity and usefulness of the technique. Further work is required to determine if monitoring the transient image will provide a basis for predicting how many more laser pulses a sample can withstand before damaging.

References

- 1) J.O.Porteus, M.J.Soileau and C.W.Fountain, Appl.Phys.Lett. 29 156 (1976)
- 2) J.O.Porteus et al, NBS Spec Pub 435 1975 p207
- 3) J.O.Porteus et al NWC TP 5988 Annual report no.9 Naval Weapons Centre, China Lake, Calif April 1976-May 1977 p145 (unpublished)
- 4) J.A.M.Jeen and J.M.Green, J.Phys. E:Sci Instrum. 17 191 91984)
- 5) R.P.Freese and K.J.Teegarden in Laser Induced Damage in Optical Materials NBS Spec.Pub. 568 1979 p313
- 6) J.A.Abate and R.Roides J.Phys.Colloq.(France) C-6 497-502 (1983)
- 7) C.D.Mars and J.O.Porteus J.Appl.Phys. 57 1719 (1985)
- 8) Refs 7-16 in the paper by S.C.Seitel, Lasers and Optronics October 1987 p67
- 9) P.A.Temple, Naval Weapons Centre, China Lake, California, USA.. private comm.
- 10) Chapter 9 in "Optics" by Hecht and Zajac, pub. by Addison Wesley 1974
- 11) G.A.Brooker, Oxford University, England....private comm
- 12) "Principles of Optics" (6th ed) by Born and Wolf, pub. by Pergamon 1980
- 13) F.A.McDonald, Can.J.Phys. 64 1023 (1986) and references therein
- 14) A.Roscenwaig and J.Opsal, IEEE Trans. Ultrasonics, Ferroelectrics and Frequency Control 33 516 (1986) and references therein
- 15) Further details can be found from any number of the multitude of papers on the photothermal effect
- 16) W.B.Jackson et al Appl.Opt 20 1333 (1981)
- 17) "Heat Conduction in Solids" by Carslaw and Jaeger pub. by Oxford Univ. Press 1986
- 18) I.C.Aamodt and J.C.Murphy J.Appl.Phys 52 4903 (1981)
- 19) "Fundamentals of Optics" by Jenkins and White, pub. by McGraw Hill 1976

Chapter 5

Colloidal Silica as a High Damage UV Antireflection Coating

5.1 Introduction

In large complex laser systems which have many transmitting surfaces, the $\approx 4\%$ reflection loss associated with each surface rapidly becomes totally unacceptable and thus over the years much work has been done to develop high quality antireflection (AR) coatings which can reduce the loss per surface to $< 0.1\%$. The development of colloidal silica as an AR coating has not only proved very successful but, due to the morphology of the colloidal silica coating, interaction with an intense laser beam reveals in a simple explainable manner "well known" effects, which can be considerably more complex in other materials.

5.2 Antireflection Theory

In order to determine the conditions for a coating/substrate combination to be AR, calculations to find the electric (E) field within the materials must be made. The following section applies ideas from standard electromagnetic (EM) theory to the problem of determining E fields in dielectric stacks. The basic EM ideas can be found in numerous books such as those by Jackson¹ or Lorrain and Corson². The book by Born and Wolf³ gives a very thorough discussion of the problem in hand but, for the purposes of the work herein a modified version of the analysis by Brooker⁴ is adequate.

Consider an EM wave travelling left to right incident on the interface at $x = 0$ between two dielectrics as shown in fig 5.1. Let the incident, reflected and transmitted electric fields be described respectively by...

$$E_{in} = E_1 * \exp(i(\omega_1 t - k_1 x))$$

$$E_{ref} = E_r * \exp(i(\omega_1 t + k_1 x))$$

$$E_{trans} = E_t * \exp(i(\omega_2 t - k_2 x))$$

where $\omega_{1,2}$ are angular frequencies ($2\pi f$) in media 1 and 2 respectively, $k_{1,2}$ is the wavenumber ($2\pi/\lambda$) in each media, t is time and x distance. A similar set of equations will describe the magnetic (H) fields.

In order to satisfy the boundary conditions of Maxwell's equations for all time, the time dependence of each wave must be the same, hence

$$\omega_1 = \omega_2$$

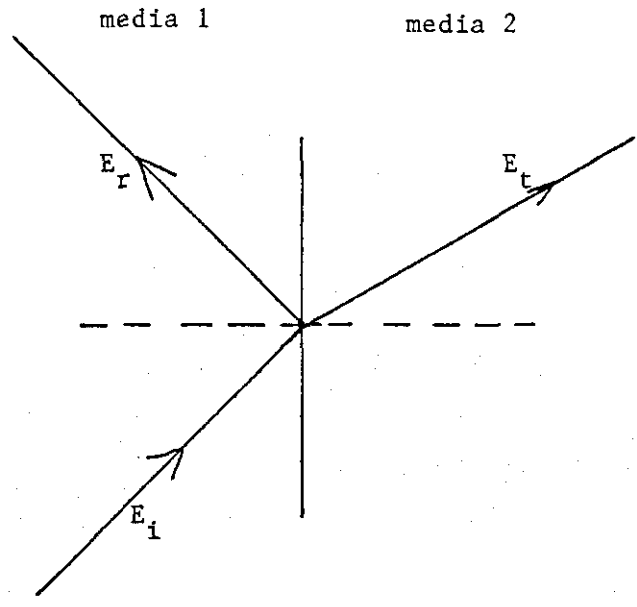


Fig 5.1

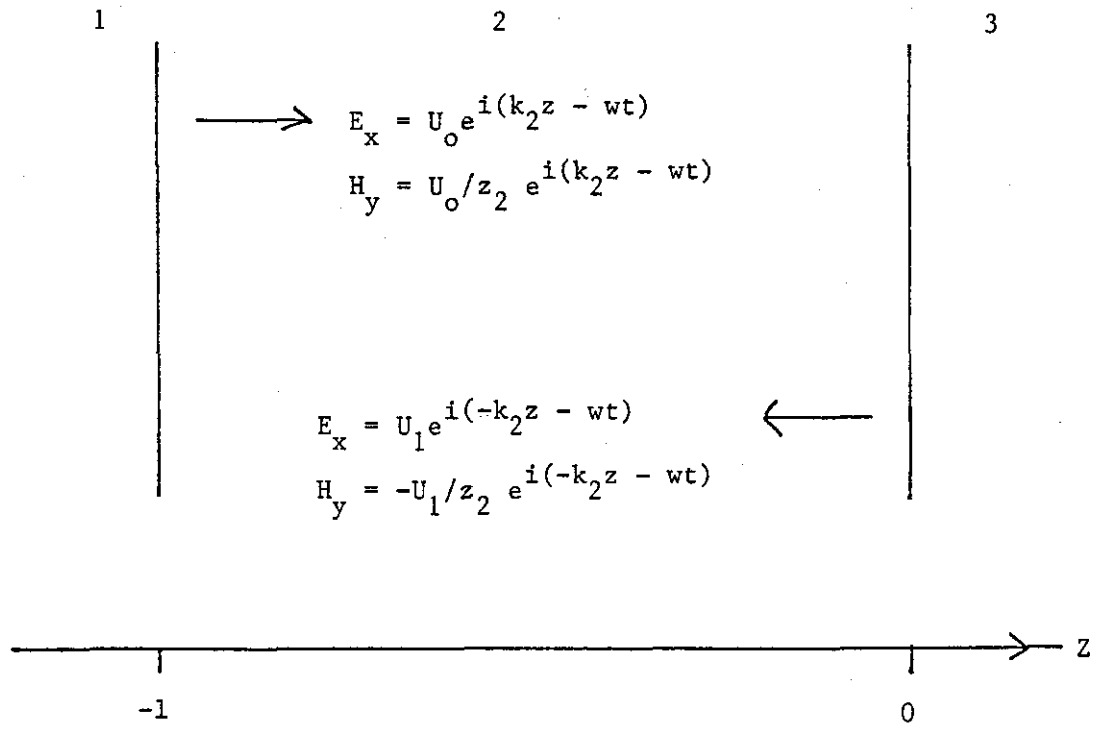


Fig 5.2

Applying the boundary conditions of continuous tangential E and H fields at $x = 0$ and eliminating the common time dependent terms gives..

$$E_i + E_r = E_t \quad 1$$

$$H_i - H_r = H_t \quad 2)$$

As all the waves are transverse electromagnetic (TEM) it is possible to write

$$E = zH \quad 3)$$

where z is known as the impedance. Since only non magnetic media are being considered

$$z = c\mu_0/n$$

where c is the vacuum speed of light, μ_0 the permeability of free space and n the refractive index for the media concerned.

Substituting 3) into 1) and 2) and by either addition or subtraction of 1) and 2) yields..

$$E_r/E_i = (z_2 - z_1)/(z_1 + z_2) = r \quad 4)$$

$$E_t/E_i = 2z_2/(z_1 + z_2) = t \quad 5)$$

where r and t are respectively the amplitude reflection and transmission coefficients.

The customary treatment⁵ of the E fields within a given layer in a multilayer dielectric stack involves representing the E field as the sum of 2 sinusoidal waves travelling in opposite directions within the layer. Consider fig 5.2 which shows a single layer coating (media 2) on a substrate (media 3) with air on the other side (media 1) where the opposite travelling waves have amplitudes U_0 and U_1 .

At the air coating interface ($z = -1$) the fields are

$$E_x(-1) = U_0 \exp(-ik_2 l) + U_1 \exp(ik_2 l) \quad 6)$$

$$H_y(-1) = U_0/z_2 \exp(-ik_2 l) - U_1/z_2 \exp(ik_2 l) \quad 7)$$

Similarly the fields at the coating substrate interface ($z = 0$) are

$$E_x(0) = U_0 + U_1 \quad 8)$$

$$H_y(0) = U_0/z_2 - U_1/z_2 \quad 9)$$

Elimination of U_0 and U_1 from 6)-9) and noting that k_2 and z_2 are real as the media are non absorbing yields

$$E_x(-1) = \cos(k_2 l) * E_x(0) - iz_2 * \sin(k_2 l) * H_y(0) \quad 10)$$

$$H_y(-1) = -i/z_2 * \sin(k_2 l) * E_x(0) + \cos(k_2 l) * H_y(0) \quad 11)$$

The load impedance of the system is defined as $z_1 = (E_x/H_y)_{z=0}$ whilst the input impedance is $z_{in} = (E_x/H_y)_{z=-1}$ and hence by dividing 10 and 11

$$z_{in} = \frac{z_1 - iz_2 * \tan(k_2 l)}{1 - (iz_1/z_2) * \tan(k_2 l)}$$

In the special case that the coating is a quarter wave layer i.e. $k_2 l = \pi/2$ then

$$z_{in} = z_2^2 / z_1 \quad 12)$$

The amplitude reflection coefficient for this quarter wave case with the light incident from the left is given by substituting 12) into 4) and denoting z_{air} by z_0 gives

$$r = (z_2^2 / z_1 - z_0) / (z_2^2 / z_1 + z_0) \quad 13)$$

Since for antireflection r must be zero i.e. $z_2^2 = z_1 z_0$ and, using the definition of z means that

$$n_2^2 = n_s \quad 14)$$

where n_s is the substrate refractive index.

5.3 Electric Field Considerations

With reference to the previous section and from fig 1 it can be seen that on the interface the electric field is different from the incident field by a factor of $1+r$. Using 4) and the definition of z ,

$$q = 1+r = 2n_1 / (n_1 + n_2) \quad 15)$$

In the case of an air glass interface $n_1 = 1$ and $n_2 = 1.5$ giving $q = 0.8$ i.e. the field is less on the interface than the incident field. However for

a glass air interface $n_1 = 1.5$ and $n_2 = 1$ giving $q = 1.2$ i.e. the interface field exceeds the incident field.

In the case of a single layer AR coating the field at the air coating interface, E_{in} must equal that of the incident wave i.e. $q = 1$, but what of the field at the coating substrate interface?. Using 6), 7) and $k_2l = \pi/2$ for a quarter wave layer U_0 and U_1 can be found..

$$U_0 = iE_{in}/2*(1+1/n_2) \quad U_1 = -iE_{in}/2*(1-1/n_2)$$

and hence the field at the coating substrate interface by 8) is

$$E_{sub} = iE_{in}/n_2 \quad 16)$$

Consider a thick glass slab of index $n_s = 1.5$ subjected to an incident field E .

The field on the incident air glass interface and thus the field in the bulk of the glass is $0.8E$. The exit surface field is on a glass air interface and thus has a field 1.2 times that of the field incident on the interface. It is the bulk field that is incident on the exit interface and hence the field at the glass air interface is $1.2*0.8E_0 = 0.96E_0$.

In order to make glass with $n_s = 1.5$ antireflecting a quarter wave coating with $n_2 = 1.22$ is required. In this case if a field E_0 is incident from air the field at the air coating interface must still be E_0 . From 15) the field at the coating substrate will be $0.82E_0$ and thus the field at the exit surface of the substrate is $0.98E_0$.

5.4 Review

There are an enormous variety of both coating materials and techniques that are used to make high reflectors (HR) and antireflectors (AR). Macleod⁶ gives a comprehensive review of the commonly used coating techniques.

Optical coating materials for the UV generally fall into two classes, oxides and fluorides. Nearly all other materials cannot be used due to their intrinsic absorption at UV wavelengths. To minimize scattering and absorption, low surface roughness and high purity materials are required. A detailed discussion of material requirements can be found in the work by Pawlewicz⁷.

The remainder of this section now concentrates on colloidal silica and the sol-gel process.

The deposition of coatings from solution is a well established method of providing good quality coatings, indeed Fraunhofer⁸ back in 1817 mentions

such a process. Coatings deposited from solution have been found to have high damage thresholds⁹ and it is because of this that much effort has been devoted in recent years to improving solution deposition techniques. The sol-gel process is one such recent development that can produce exceptionally good coatings of almost any metal oxide¹⁰⁻¹¹. Detailed descriptions of the various stages and materials involved in the sol-gel process can be found in a number of references¹²⁻¹⁴.

Briefly, the sol-gel process to deposition of oxide layers is based upon the formulation of organometallic solutions containing mainly alkoxides, which are deposited onto the substrate and can be taken to gelation in a controlled manner. Upon formation, the gel is slowly dried and then baked to form a semi-porous film. The solution is actually deposited on the substrate either by dipping or spinning techniques¹⁴.

The performance of SiO₂ coatings (silica) deposited by this process is critically dependent upon the materials and catalyst used to produce the colloidal silica. Early work¹³⁻¹⁴ not only showed that if the wrong catalyst was used the coatings would become contaminated with carbon and have low damage thresholds but that the use of thin films improved the threshold.

5.5 Coatings Tested

The coatings tested at Loughborough were provided by Dr. I. N. Ross of the Rutherford Appleton Laboratory. They consisted of colloidal silica deposited onto 3 inch square fused silica plates. The coatings consisting of colloidal silica prepared in a variety of ways were in the form of quarter wave (optical thickness) antireflectors for a number of UV wavelengths. The refractive index of the fused silica is ≈ 1.52 at 250 nm thus requiring a coating of index 1.23 to form a perfect AR combination. The refractive index of the colloidal silica (see next section) is ≈ 1.22 thus using the fused silica substrate the theoretical reflectivity is 0.001%.

The bare silica coating is very susceptible to contamination (and thus a low damage threshold) via absorption of molecules into the voids in the coating structure, in particular organic molecules pose a very significant problem. To overcome this problem overcoating with a half wave (optical thickness) layer of various barrier materials was tried.

5.6 Structure of Colloidal Silica Coatings

Colloidal silica consists of silica particles suspended in a suitable solvent, usually ethanol, with the particles being kept apart by electrostatic repulsion as the particles carry the same charge. The size of the particles depends upon the method and materials used to produce the

solution. Transmission Electron Microscopy (TEM) both here and at the Rutherford showed (fig 5.3) that the particles are near spherical with a diameter of $\approx 20\text{nm}$.

Given that the refractive index of silica at 249nm is 1.44 and using a mean polarizability model as in Born and Wolf³ to estimate the refractive index of the mixture of silica and air, the particles, in order to reach an index of 1.22 must be arranged so as to occupy $\approx 50\%$ of the available space in the quarter wave layer. For an index of 1.22 a quarter wave layer at 249 nm is $\approx 50\text{nm}$ thick i.e. a little over 2 particles thick. As evident from fig 5.3, the particles are randomly distributed within the available space; however, as will be shown later in coatings that have been exposed to the excimer beam the particles can take up an ordered distribution.

5.7 Damage Thresholds

Damage tests were performed on a variety of samples according to the method discussed in chapter 2 the results of which are summarised in tables 5.1 and 5.2.

As discussed in chapter 1 for a parallel incident beam the intensity due to the E field as a result of an in phase reflection is higher at the back (exit) surface of a sample by a factor of $4n^2/(1+n)^2$ which for glass is ≈ 1.44 . To avoid problems of exit (back) surface damage due to this increased electric field, the samples were positioned well past the focal plane of the lens such that the area of the beam at the exit surface was at least 1.5 times that at the entrance surface.

During the tests a variation of the effects associated with the so called "breath test"¹⁵ (see 5.9) was observed. TEM studies of the samples revealed not only the cause and explanation of the observed effects but the presence of LIPSS on the samples (see 5.10).

5.8 Analysis and Comments on the Damage Results

The spot nature of the damage to sample 1 indicates that it is isolated defects that are damaging. Given that prior to the development of colloidal silica it was essentially impossible to obtain AR coatings for the UV with thresholds above $\approx 4 \text{ Jcm}^{-2}$, the results for sample 2 show that colloidal silica has the potential to be a major step forward in AR coating technology.

Due to the complications involved in having to distil the colloidal silica to remove the non silica particles, samples of distilled and undistilled colloidal silica were prepared to allow a comparison of the damage thresholds and morphology. On sites with no initial scattering (clean sites) the thresholds of samples 2 and 3 were identical. However sample 3 has large

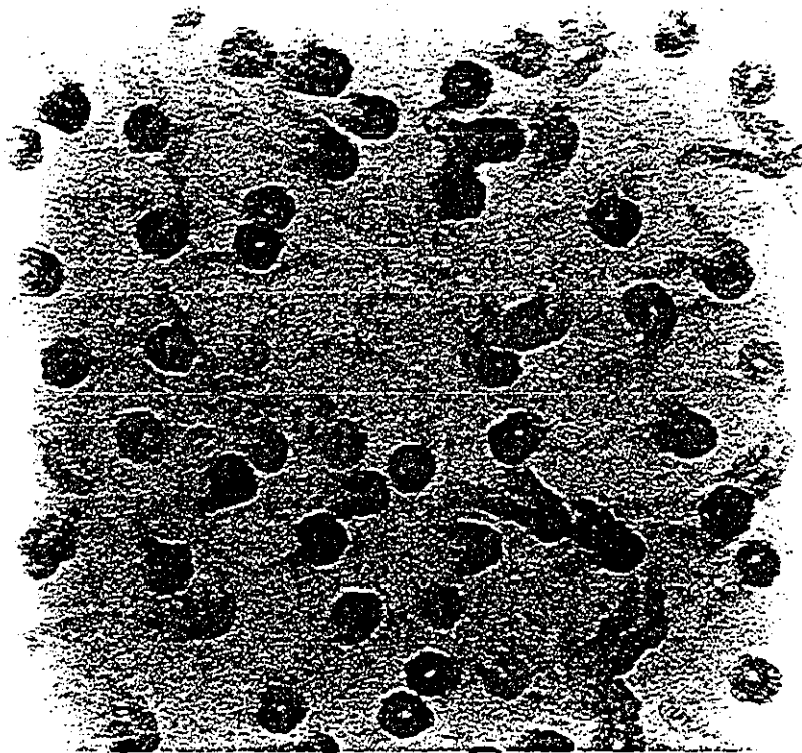


Fig 5.3

TEM picture of the Colloidal Silica coating. Magnification X250,000

Note how the colloidal silica appears as near spherical particles of ≈ 20 nm diameter.

numbers of initial scattering (dirty) sites which sample 2 does not have and on these the threshold is substantially lower. The morphology of the damage was identical on each. These results show that the distillation process is essential and must be used to produce the optimum coatings.

Further work at the Rutherford resulted in an improvement in the technology and lead to samples 4 and 5 being tested. The GR650 overcoat is a highly transmitting UV polymer which was applied to act as a barrier layer to reduce the contamination of the colloidal silica. With reference to the figures for spot/blow off type damage it is apparent that the improved coatings have a much higher threshold and that though the GR650 reduces this figure, it is to a level that is still better than that available with other coatings and coating technologies.

However all of the coatings showed a change on exposure to a single low fluence pulse which showed no further change with subsequent pulses and, which the author has named "darkening". All the damage tests were performed according to the method in chapter 2 and thus had the knife edge set to cut off 50% of the light. As a result of a low energy excimer pulse a significant fraction of the beam target interaction area went darker i.e. less light was observed to come

Table 5.1

Sample no.	Description
1	bare fused silica substrate
2*	distilled colloidal silica
3*	undistilled colloidal silica
4*	improved colloidal silica
5*	improved colloidal silica with GR650 overcoat**

* all coatings on fused silica substrates.

** overcoat in the form of a half wave layer

Table 5.2

Sample no.	Damage Threshold Jcm ⁻²	Morphology of Damage
1	16.5	spot ⁺
2	7.7 - 7.9 * <3.0	" " darkening ⁺⁺
3	7.7 - 7.9 ^{**} <6.5 ^{***} <3.0	spot blowoff ⁺⁺⁺ darkening
4	10.8 - 11.2 <3.0	spot/blow off darkening
5	4.0 - 4.4 <3.4	spot/blow off darkening

* figures preceded by a < sign mean that the sample was not tested at lower values.

** figures for a "clean" site i.e. one where there was no detectable scatter prior to the laser pulse.

*** figures for a "dirty" site i.e. one where there was detectable scatter prior to the laser pulse.

+ "spot" refers to the creation of isolated small (spot) scattering sites as a result of the damage on what were initially "clean" sites.

++ "darkening" refers to a uniform reduction over the a significant fraction of the laser target interaction area in the detected light (see 5.8).

+++ "blow off" refers to increased scatter as a result of the laser pulse on an initially dirty site. Damage on these dirty sites when viewed, appears to result in the expulsion of material from the surface (blow off).

from this area. There was no damage in the sense that there was no blow off from existing scattering sites and no new sites were created, nor did the darkening appear to effect the optical properties of the samples.

Ellipsometry measurements were performed to try and determine if there was any change in either the refractive index or optical thickness of the coating in the areas that darkened. These measurements were severely restricted by both the difficulty of aligning the probe laser with the areas that had darkened and the size of the probe laser spot in relation to that of the darkened areas. The changes found if any were extremely small and well within the variation found across the undamaged areas of the sample and hence these measurements were inconclusive.

The cause of this darkening which is similar to that observed in laser annealing (chapters 1 and 3) is not clear, but in the course of discussions with Dr. Ross there was the suggestion that it is caused by photo-ablation/photo-decomposition of the coating. Further detailed work is obviously required to fully explain this effect.

On the assumption that the fused silica and the colloidal silica have the same intrinsic damage threshold then, from the analysis in 5.3 the ratio of the electric fields at the damage threshold of the bare substrate to that of the coated sample should be $1/0.8$. Since in any actual experiment what is measured is the energy of the beam which is proportional to E^2 , the ratio of the damage fluences should be $1^2/0.8^2 = 1.56$. The ratio of the experimentally determined thresholds (using the best coating) is ≈ 1.5 . This analysis although very crude shows that the coatings using the improved colloidal silica are performing in a near optimal manner and it is the author's contention that further development (based on quarter wave layers) will not significantly improve the damage threshold. However the use of an overcoat material is a very recent innovation and further work on this may result in the overcoated samples approaching the non overcoated damage threshold.

5.9 The "Breath Test"

5.9.1 Introduction

The so called breath test¹⁵ refers to the deliberate fogging of a surface in order to detect any changes in that surface. The fogging is usually achieved by literally breathing heavily near the sample so as to raise the local water vapour content of the air which subsequently condenses on the sample's surface forming a very thin layer of moisture.

The size of water molecules on the surface and thus its appearance when viewed (particularly at near grazing incidence) will depend critically on

local surface conditions. Thus the breath test (when the sample is suitably imaged) can be a very sensitive test for changes in the surface of a sample. However, the breath test is in practice, of little use in damage testing for two reasons

1) Since the nucleation rate depends so critically on the first atomic layer i.e. first few nm whereas even in metals a light wave can propagate 40 nm or more (especially in the UV) it is highly questionable as to whether a small change in the first few nm would significantly change the optical properties of the sample. Hence it is quite possible that the breath test is responding to changes that are of no consequence optically.

2) Since this method requires water i.e. a contaminant to be regularly condensed onto the surface it is not suitable for a large number of the coatings and samples used since they can be hygroscopic.

To understand the sensitivity of this method a review of some basic nucleation theory must first be undertaken.

5.9.2 Nucleation Theory

A comprehensive introduction to nucleation theory can be found in the book edited by Zettlemoyer¹⁶ whilst the use of lasers in modifying nucleation barriers has been documented by several authors, in particular by Tsao and Ehrlich¹⁷. The theory now presented is based upon the analysis in ref 16.

Homogeneous nucleation i.e. the condensation of a substance from a vapour onto itself occurs via the random growth of small condensed aggregates or clusters of vapour molecules. The barrier to this process is the increase in the free energy (ΔG) associated with the cluster as compared to the vapour phase. It can be shown that this increase ΔG can be written as the sum of two terms one positive and associated with the excess free energy needed to form the surface of a cluster, the other negative and associated with the difference in free energy of the vapour and condensed phases. Since both of these terms depend upon r the radius of the cluster, ΔG exhibits a maximum ΔG^* at what is termed the critical cluster size r^* . Fig 5.4 is a sketch of the typical variation of ΔG with r . Clusters smaller than r^* are unstable relative to the vapour phase and tend to evaporate whilst clusters larger than r^* become progressively more stable and thus tend to grow. Under these conditions the nucleation rate N for critically sized clusters can be written as

$$N = A \exp(-\Delta G^*/kT)$$

17)

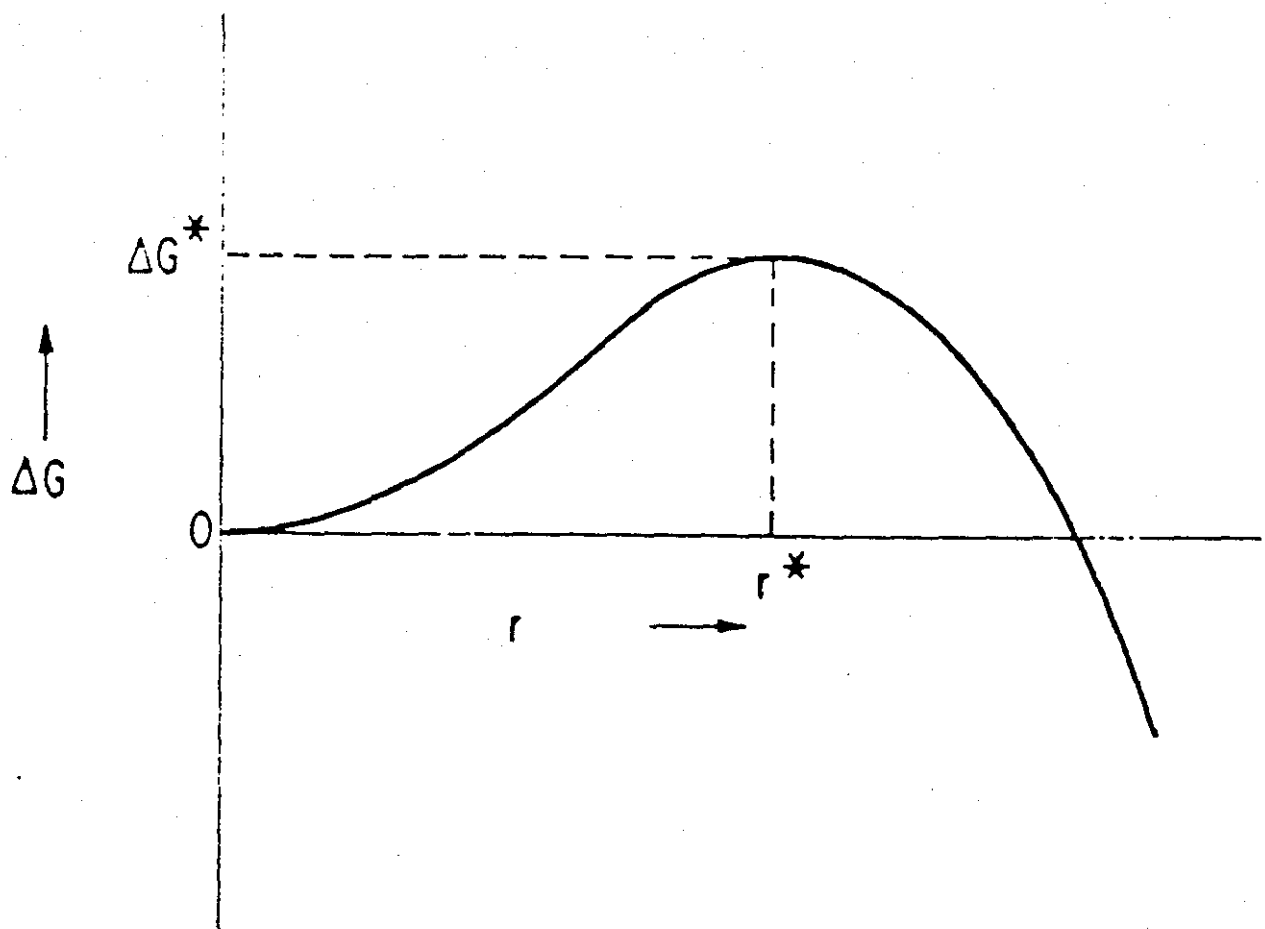


Fig 5.4

Free energy of embryo formation versus size

Note peak at r^* (from ref 16)

where A is a normalization constant, k the Boltmann's constant and T the absolute temperature.

Heterogeneous nucleation i.e. the condensation of a vapour onto a different substrate material yields a similar expression to 17); however, the ΔG term is considerably more complicated as the vapour phase and substrate material are no longer the same but again has a maximum value. Denoting the coefficient of surface tension between liquid and substrate by γ , the contact angle between liquid and substrate by θ , the volume of one condensed molecule by v , the local pressure by p and the saturation pressure at the local temperature T by p_s the maximum in ΔG is given by

$$\Delta G^* = 16\pi/3 * (\gamma^3 \phi(\theta) / \Delta G_v^2) \quad 17$$

where

$$\phi(\theta) = (2 - 3\cos\theta + \cos^3\theta) / 4 \quad 18)$$

and

$$\Delta G_v = -kT/v * \ln(p/p_s) \quad 19)$$

The nucleation rate of these critically sized clusters is then given (to within a constant) by substitution of 17) into 16).

5.9.3 Effects of surface Properties on the Nucleation Rate

Considering the nucleation of water on glass at a temperature of around 25°C, with $v = 3 \text{ E-}29 \text{ m}^3$ and $p/p_s = 2$ table 5.3 shows the effect on the nucleation rate of varying γ or θ .

Table 5.3

$\gamma \text{ Nm}^{-1}$	$\theta \text{ Deg}$	$\Delta G^* \text{ xE-}19$	$\exp(-\Delta G^*/kT) \text{ xE-}66$
0.072	130	6.26	2.2
0.072	132	6.34	0.34
0.071	130	6.00	1100

From table 5.3 it is evident that small changes in the surface properties particularly in the coefficient of surface tension γ can enormously effect the nucleation rate.

5.9.4 Results

5.9.4a Observed Effects

It was found that if an area was irradiated by a non damaging fluence and then fogged by breathing on it, a clear imprint of the laser beam could always be seen.

Fig 5.5a shows the surface of sample 2 after irradiation by a single pulse of fluence 3 Jcm^{-2} . Fig 5.5b shows the same area as 5.5a but after it has been fogged just sufficiently for the difference in the areas to show up and before the moisture has evaporated. The laser irradiated area shows up quite clearly as a dark patch with the characteristic size and shape of the laser beam. Since a Schlieren imaging system was used this means that, either at the time the picture was taken no condensation (nucleation) had occurred on the irradiated area or if it had, it was insufficient to cause detectable deviation of the light. In order to try and elucidate the mechanism responsible for this effect TEM and Laser Induced Mass Analysis (LIMA) studies were carried out. For the purposes of these studies an area on sample 2 was irradiated by a single excimer pulse at 4 Jcm^{-2} .

5.9.4b LIMA Results

LIMA, being a technique for determining the elemental/molecular composition of a substance, was used to determine if there was any difference in the composition of irradiated and non irradiated areas. LIMA involves using a laser (in this case a quadrupled Nd:YAG) to vaporize and ionize a small area of the test sample. By suitably applied electric fields this plasma is directed into a time of flight mass spectrometer for analysis.

Figs 5.6a-d show the results of the LIMA studies. The labelling "damage zone" refers to an irradiated area and "background" to a non irradiated area close to the irradiated area. LIMA can be set to detect either positive ions (+ve labelling) or negative ions (-ve labelling). It is apparent from these graphs that there is no difference in composition between the areas and hence the observed effects are not compositional dependent.

On a coating/substrate combination that is nominally just silicon and oxygen it is interesting to note that there is a whole range of impurities present of which the main peaks in the graphs are appropriately labelled.

5.9.4c TEM Results

Replicas of the surface (irradiated and non irradiated areas) using a standard acetate/carbon coating technique were made for use in the TEM studies. Figs 5.7a-b show typical TEM pictures for non irradiated (5.7a) and irradiated (5.7b) areas.

In 5.7a it can be seen that the particles occur singularly and

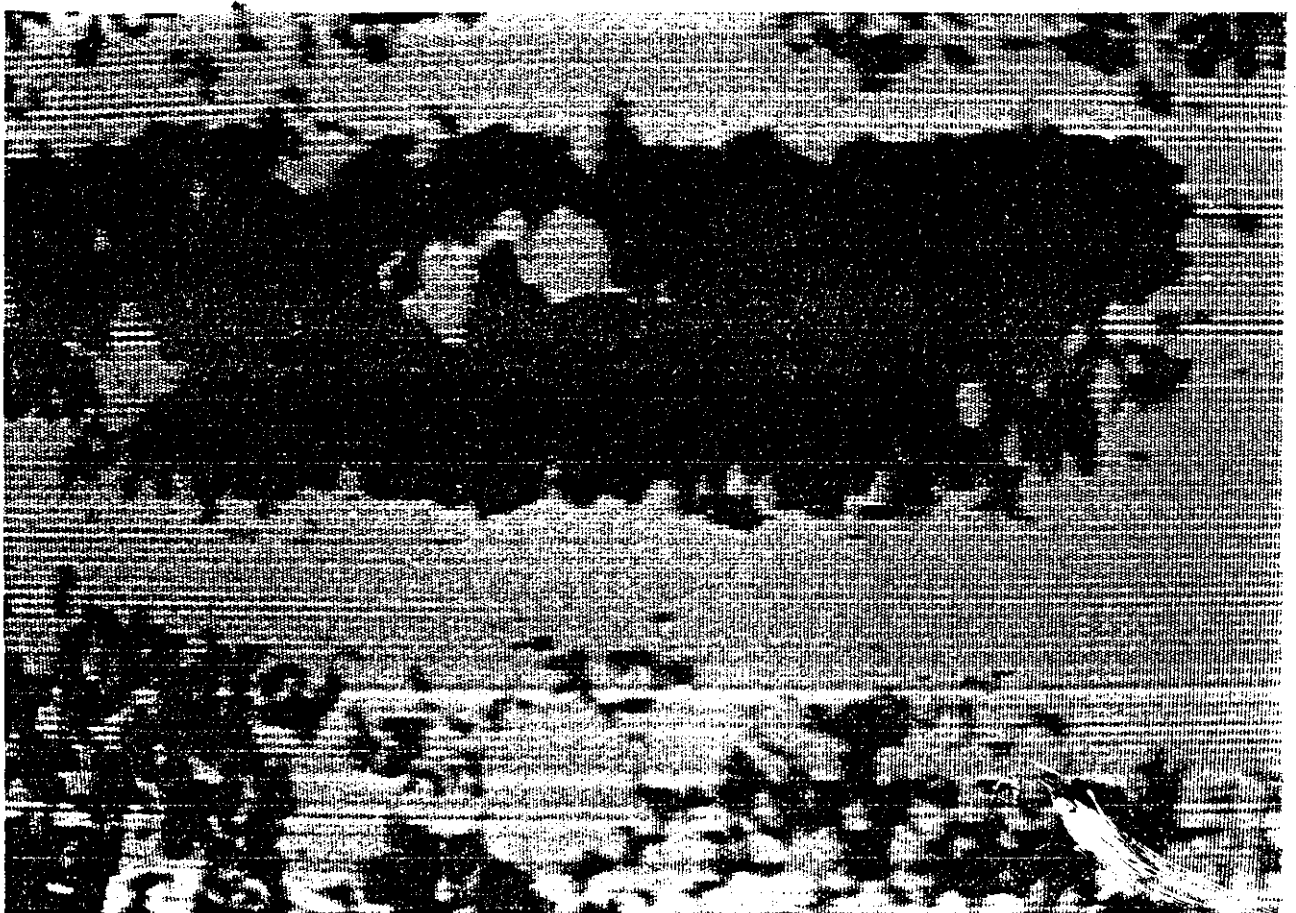
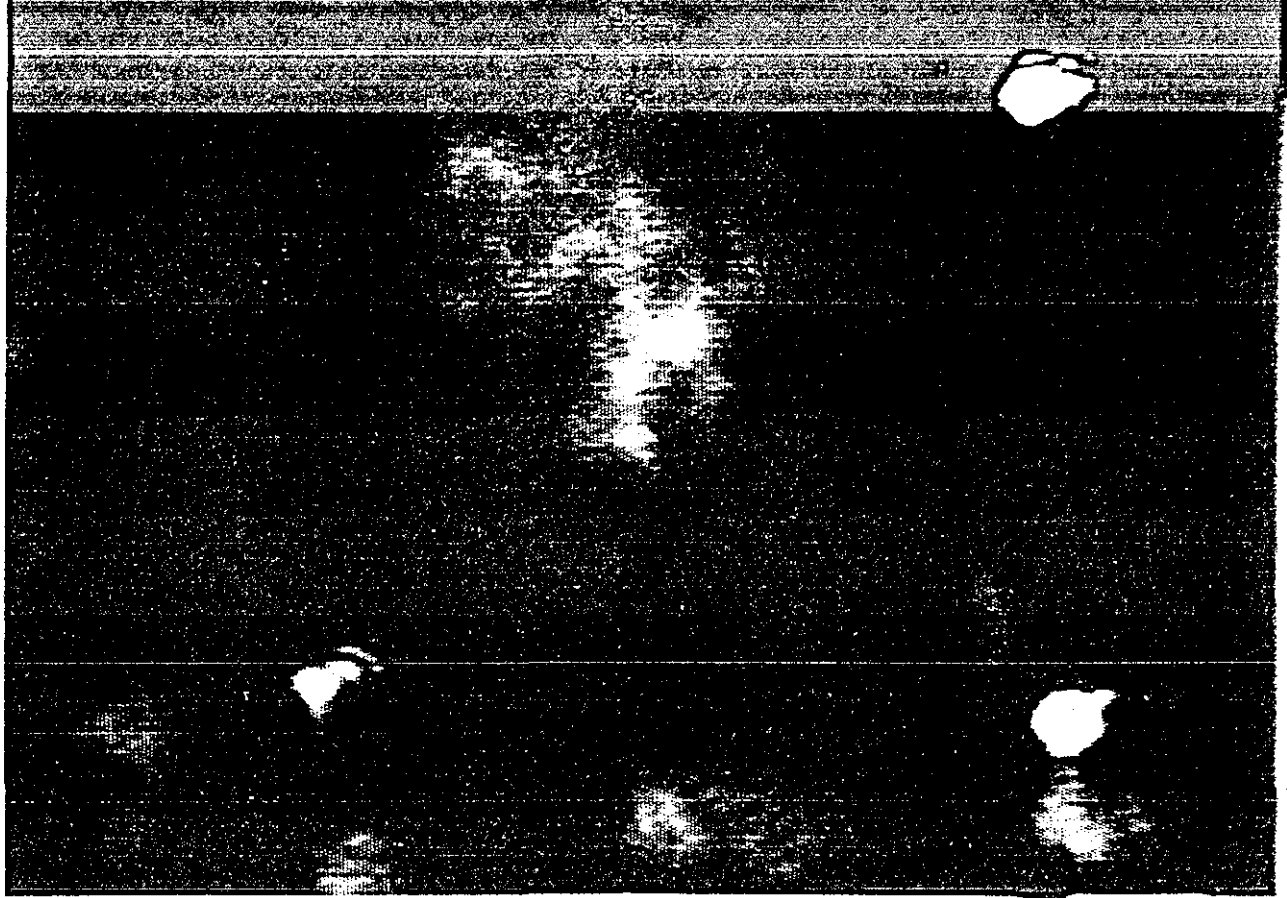
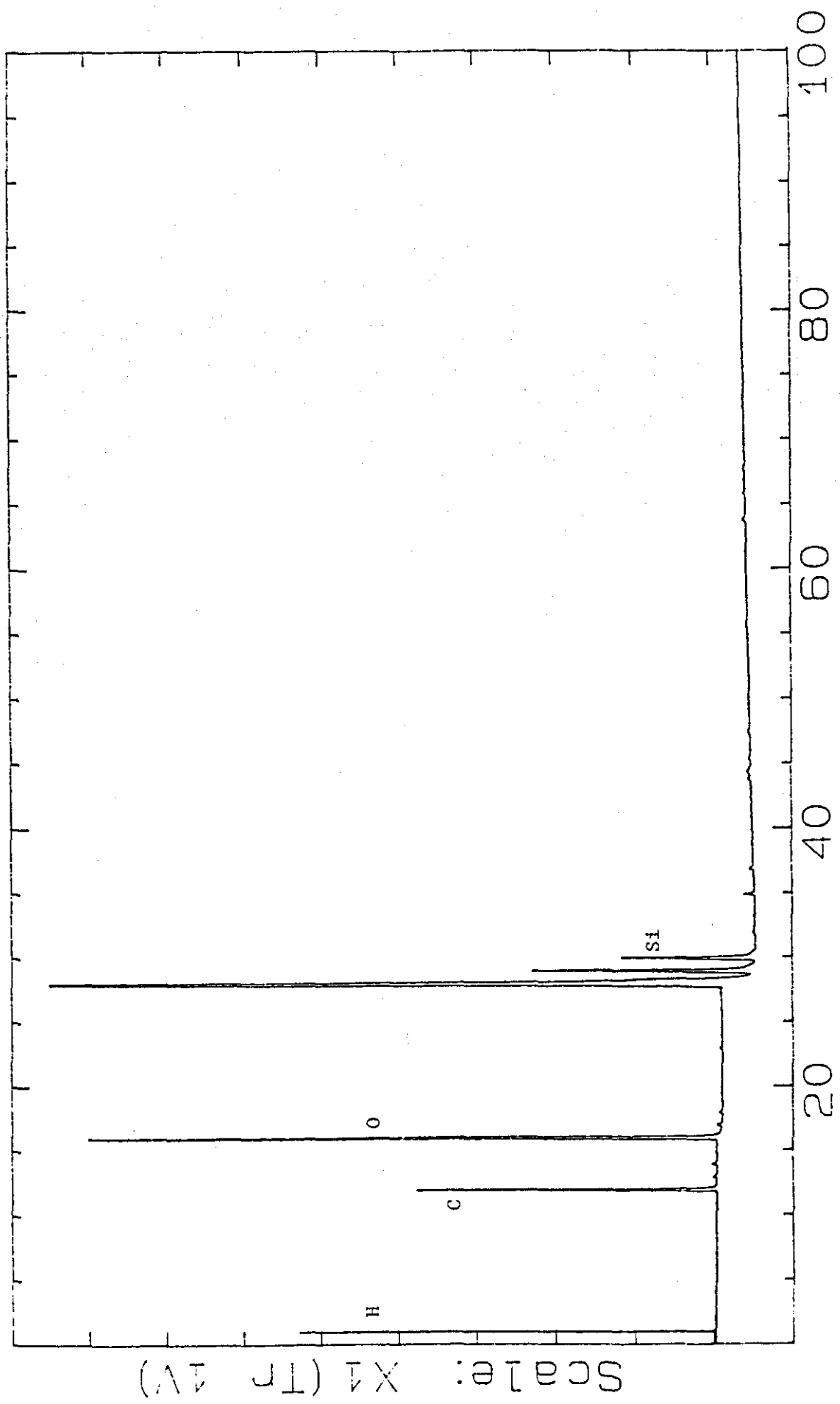


Fig 5.5

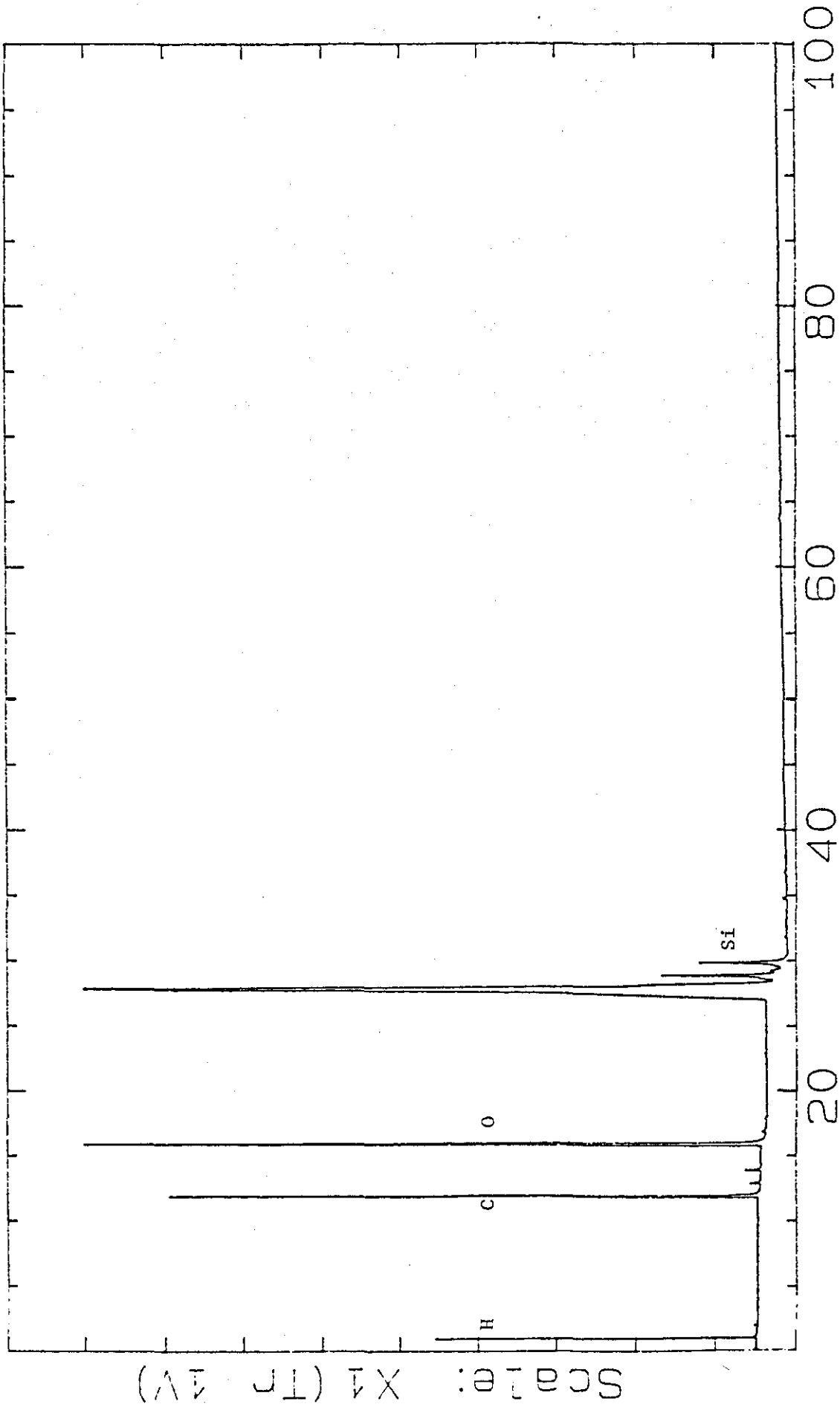
Top a) colloidal silica coated surface exposed to non damaging laser pulse
Bottom b) breath test effect on the same surface as in a)



File name: 86931.P
23/12/86

Fig 5.6a

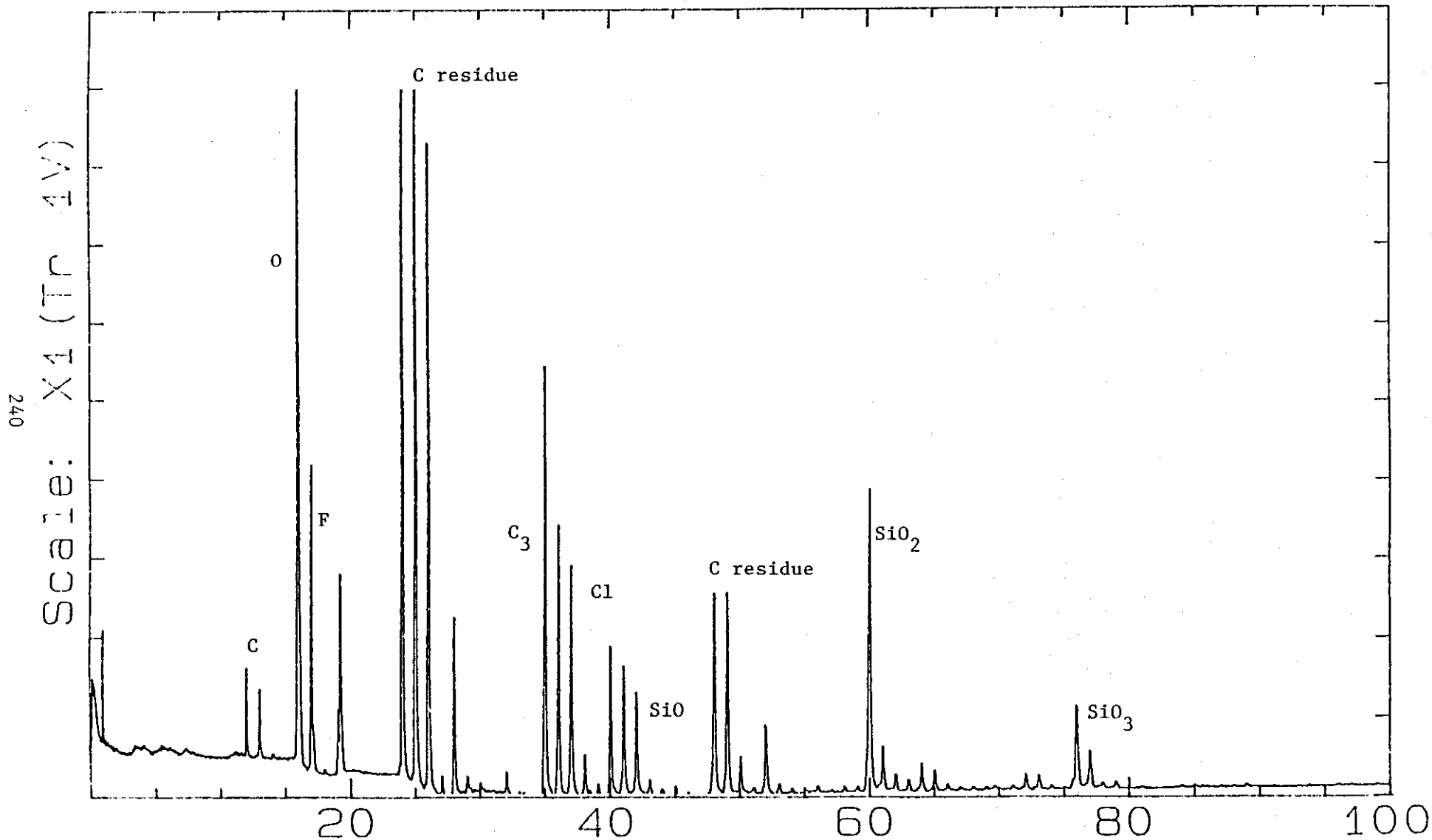
DAMAGE ZONE +VE



File name: BGP31_3
23/12/86

Fig 5.6b

BACKGROUND +VE



DAMAGE ZONE -VE

Fig 5.6c

File name: B6P31_5

23/12/86

Scale: X1 (Tr 1V)

241

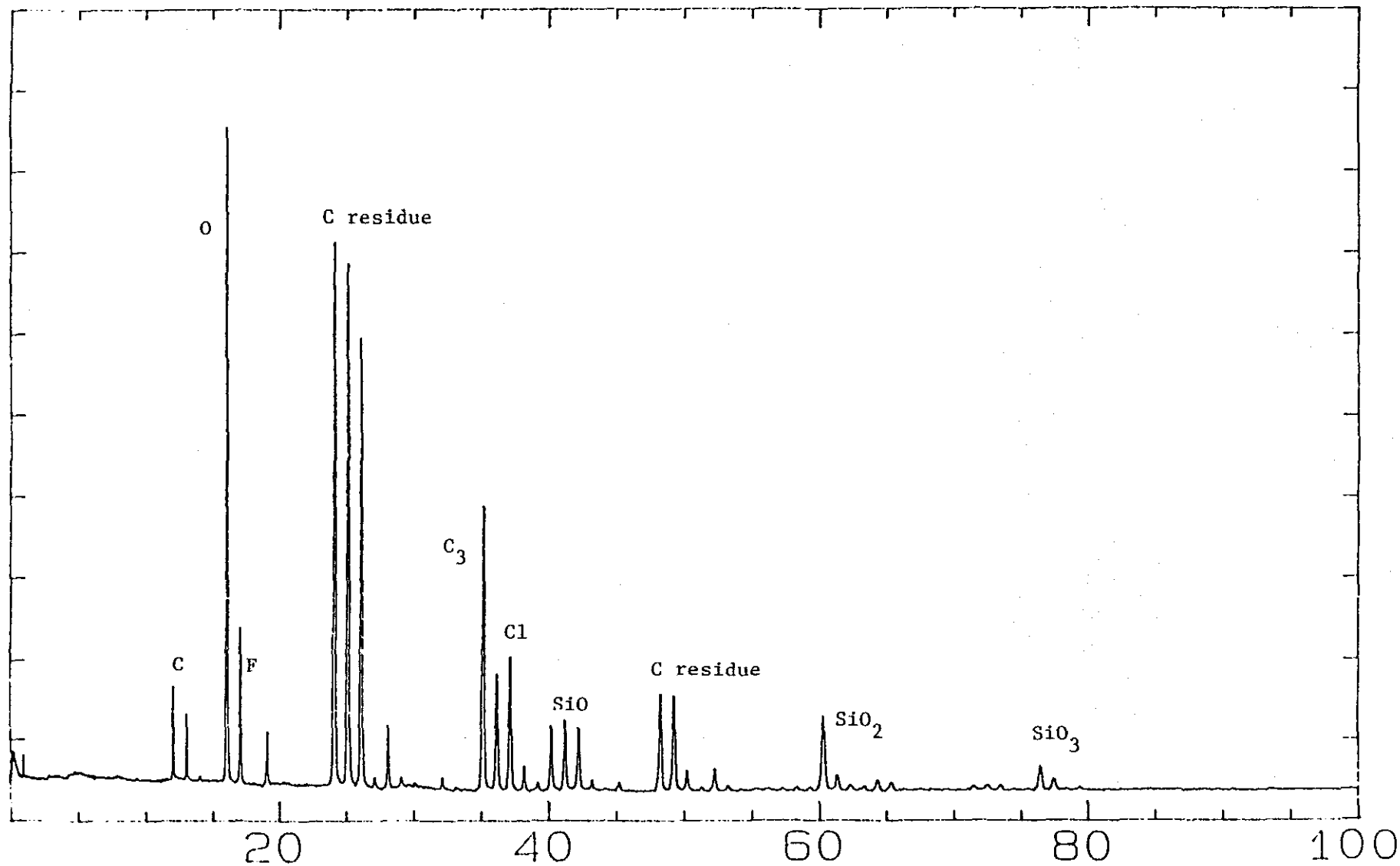


Fig 5.6d

File name: B6P31_4

BACKGROUND -VE

23/12/86

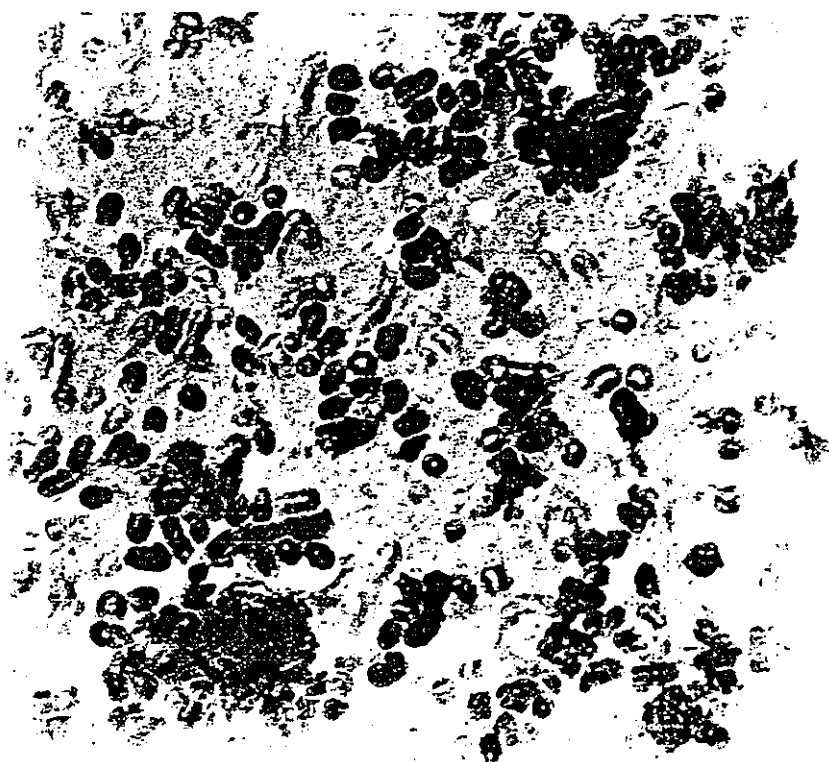
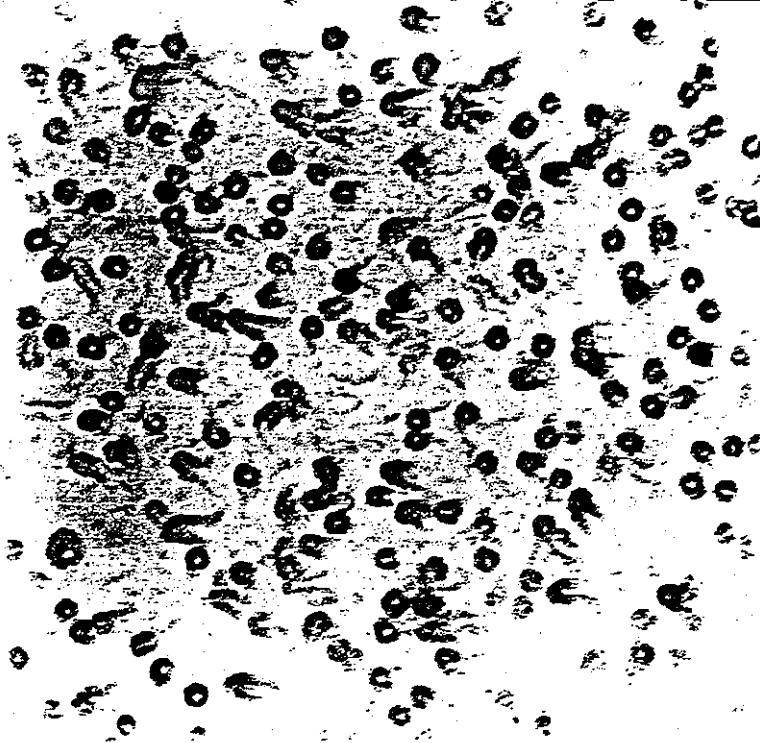


Fig 5.7

TEM pictures of colloidal silica surfaces, magnification X125,000

Top a) non irradiated

Bottom b) irradiated by non damaging pulse

essentially randomly distributed across the observed area, whereas those in 5.7b occur predominantly in clusters each of which contains several particles. This predominance of clusters is probably due to the rapid heating and cooling that the surface undergoes as a result of excimer irradiation.

5.9.5 The "breath test" - Explanation

In view of the LIMA and TEM results it is the author's contention that the observed breath test effects are due to a difference in nucleation conditions between irradiated and non irradiated areas resulting from the change in physical structure of the coating. From the limited results obtained it is not possible to determine if the change in nucleation conditions simply results in less particles being nucleated and or a significant change in the physical size of the nucleated particles.

5.10 Dielectric UV LIPSS

In the course of the TEM studies it was observed that some of the irradiated areas had what appeared to be ripple patterns (LIPSS) on them and is to the author's knowledge the first time that UV generated LIPSS have been detected on dielectric materials.

Since these structures were generated by only a single pulse it is perhaps not surprising that they are poorly defined in comparison to those in chapter 3 where 20 -40 pulses were used. Figs 5.8a is typical of the majority of observed patterns where it can be seen that each ripple which occurs as a vertical line is formed by the ordered coalescence of several particles and that the ripple spacing is $\approx 240\text{nm}$.

Since the colloidal silica is a dielectric and the irradiation was performed at normal incidence, calculation of the efficacy factor (chapter 3) yields graphs such as 5.8b and predicts spacings of 210 or 240 nm depending upon the orientation of the fringes, in broad agreement with that observed experimentally.

Occasionally well defined ripples such as those in figs 5.9a-b but with substantially different spacing were observed. It is clear particularly from fig 5.9b that the ripples are formed by the alignment of small clusters each of which comprises several particles and have a spacing of $\approx 350 - 400\text{nm}$ i.e. almost a factor of 2 larger than the above theory predicts.

The difference in the morphology of the two types of ripples (heavily damaged areas to areas of small change) suggests that the patterns were formed at significantly different temperatures and thus electric fields. At the time of writing the author has no detailed explanation as to the cause of these

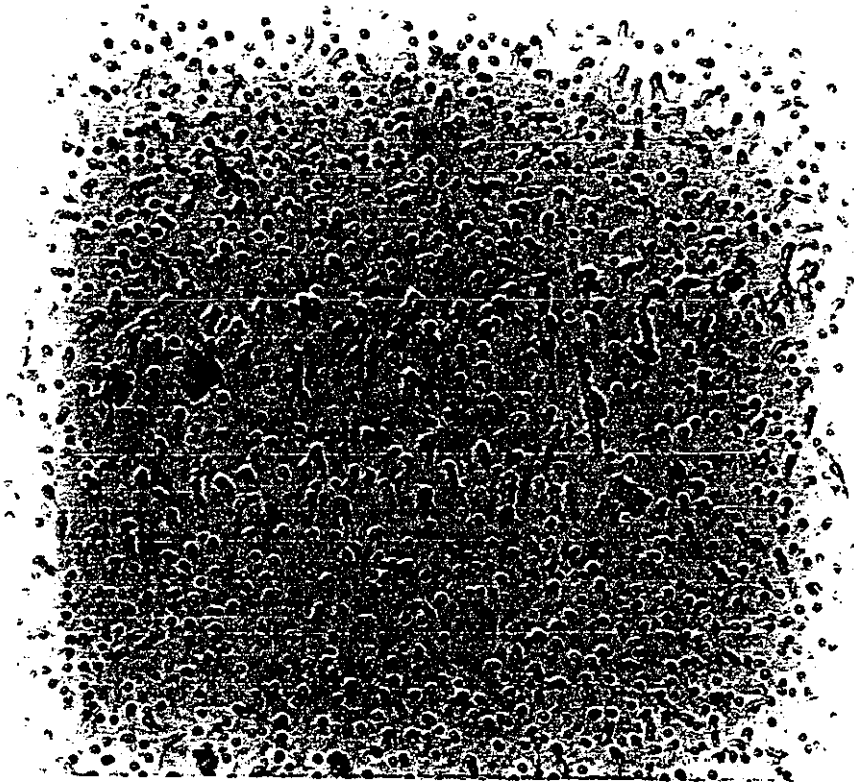


Fig 5.8a

TEM picture of LIPSS on colloidal silica magnification X50,000. Fringes are of poor quality and form in vertical lines.

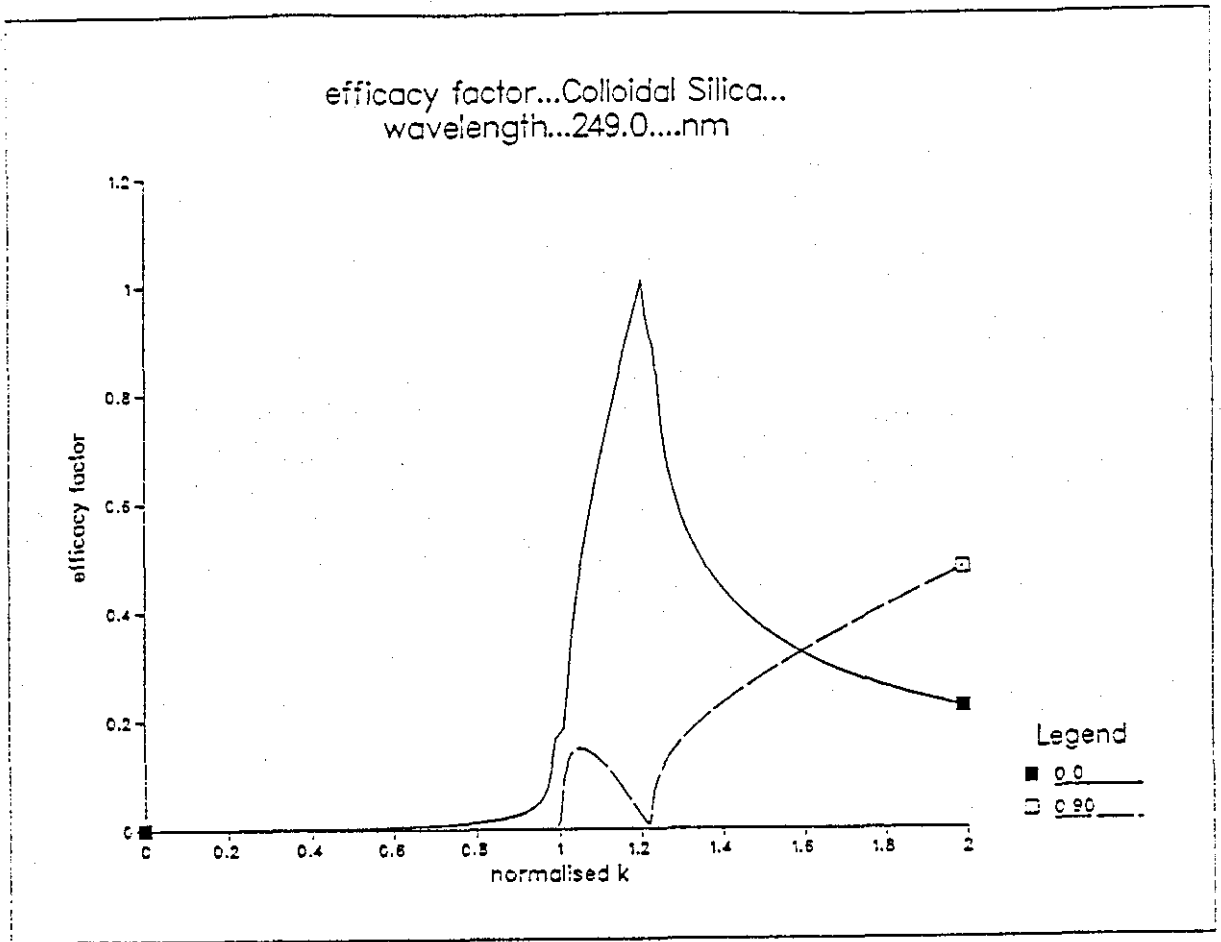


Fig 5.8b

Efficacy factor for colloidal silica irradiated at normal incidence. Obtained by using $s = 0.4$ and $f = 0.5$ (see chapter 3).

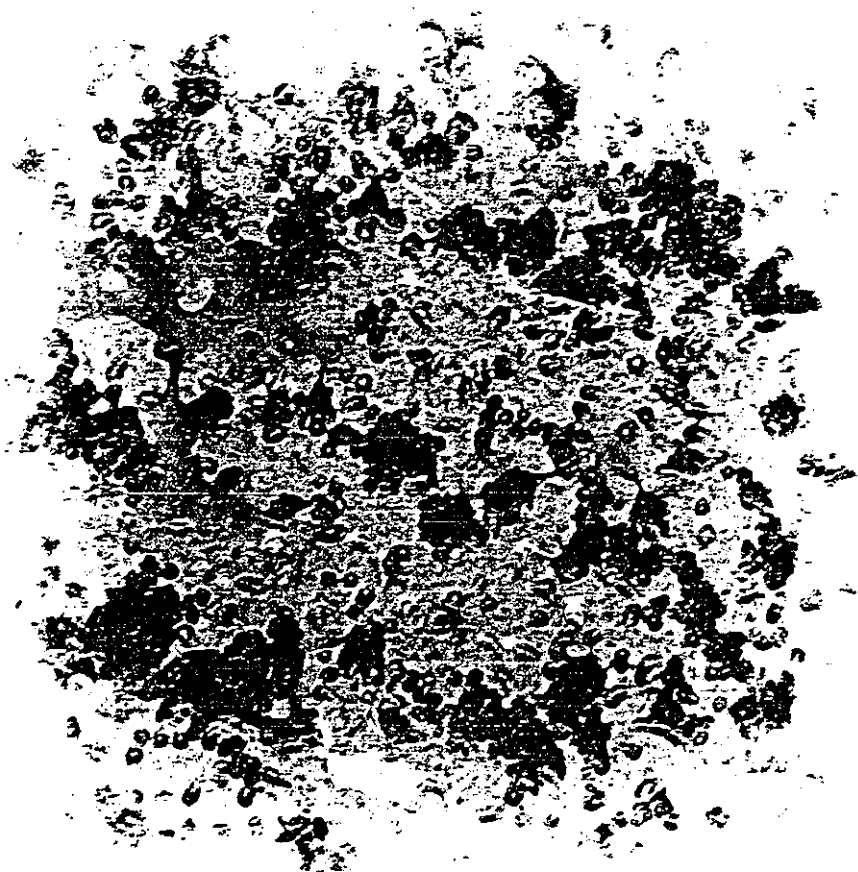
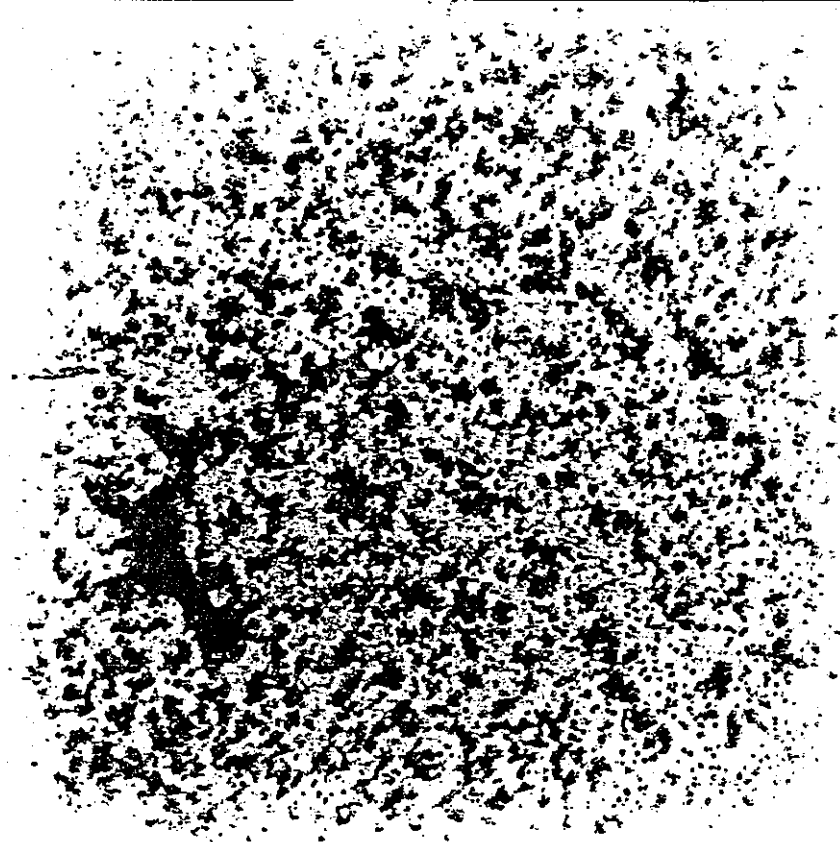


Fig 5.9

TEM pictures of anomalously spaced LIPSS on colloidal silica
Top a)magnification X25,000
Bottom b)magnification X83,500

anomalous ripples.

5.11 Conclusion

The results herein show that colloidal silica is potentially a very good AR coating material and at the time of writing, has been used to provide a large fraction of the required AR coatings for the UV laser work at the Rutherford Appleton Laboratory.

A simple demonstration and explanation of the so called "breath test" based on the coalescence of the silica particles has been given together with the first report of UV generated LIPSS on dielectrics.

References

- 1)"Classical Electrodynamics" by J.D.Jackson ,2nd ed. pub. by Wiley 1975.
- 2)"Electromagnetic Fields and Waves" by P.Lorrain and D.Corson ,2nd ed. pub. by Freeman 1970.
- 3)"Principles of Optics" by M.Born and E.Wolf 6th ed. pub. by Pergamon 1983.
- 4)G.A.Brooker ,Oxford University...private comm.
- 5)P.H.Berning,Phys.Thin Films 1 69 (1963)
- 6)H.A.Macleod ,Proc SPIE 476 128 (1984)
- 7)W.T.Pawlewicz, Proc SPIE 476 163 (1984)
- 8)J.von Fraunhofer "Versuche uber die Ursachen des Anlaufens und Mattwerdens des Glasses und die Mittel,denselben zu vorzukommen (1817)," p 35. Joseph von Fraunhofer Gesammelte Schriften, Munich, 1888.
- 9) W.H.Lowdermilk,Proc SPIE 387 55 (1983)
- 10)B.E.Yoldas, Appl.Opt. 21 2690 (1982)
- 11)H.Schroeder, Phys.Thin Films 5 87 (1969)
- 12)B.S.Bilan ,Newcastle Polytechnic 1986,industrial report on "The development of colloidal silica AR coatings at the rutherford Appleton Laboratory".
- 13)I.M.Thomas ,Appl.Opt. 25 1481 (1986)
- 14)W.H.Lowdermilk et al, NBS Spec.Pub 688 p372 (1985)
- 15)D.Milam, Proc. SPIE 140 52 (1978)
- 16)"Nucleation" ed. by A.C.Zettlemoyer, pub. by Dekker inc 1969.
- 17)J.Y.Taso and D.J.Ehrlich, J.Crst.Growth 68 176 (1984)

Chapter 6

Conclusions and Recommendations

The work in this thesis can be divided into 2 categories, that primarily of academic interest i.e.the LIPSS work (chapter 3) and that of primarily practical interest (chapters 2,4 and 5), in particular the beam profiling work (chapter 2) and the EDLID work (chapter 4).

The LIPSS work consisted of a detailed comparison of theory and experiment under conditions that severely tested the theory.The results showed in general the validity of the theory to at least first order if not beyond.

An investigation of LIPSS for different samples,laser wavelengths and pulse durations would provide further tests for the current theories.The use of ultra short (\ll lps) pulses may well reveal new types of structures as both the coherence time and the number of optical cycles that the pulse duration corresponds to decrease.

Unfortunately it was not possible with the equipment available at the time to obtain time dependent information on the formation and subsequent growth of the ripples which almost certainly would yield valuable information.With the current availability of a short pulse dye laser with relatively low trigger jitter (as used in chapter 4) and as a part of the ongoing program of LIPSS research at Loughborough, the author together with colleagues plan to investigate the growth of these ripple structures.By varying the delay between the excimer and dye lasers it will be possible to built up over a number of shots and test areas data on the formation of the ripples as a function of time.

One of the basic assumptions in all of the proposed ripple theories is that the interference effect leads to some degree of periodic melting which, to the author's knowledge has never been directly experimentally verified.Use of a suitably powerful dye laser timed to arrive at the end of the excimer pulse coupled with Schlieren imaging of a suitably angled sample should allow direct probing of the surface temperature and thus test this hypothesis.

The beam profiling work has shown the advantages of a system that enables a complete spatial profile of a pulsed laser beam to be obtained in a single shot although the use of a more sensitive camera would reduce the problem of saturating the fluorescence and still produce a large enough video signal

whilst one that responded directly to the UV light would eliminate the need for a fluorescer. Essentially finance controls the former whilst material and technological problems dominate the latter.

The use of video framestores has been shown to improve damage detection but, from a practical viewpoint it is somewhat pointless to have a system so sensitive to damage that it responds to changes much smaller than those of significance when the sample is used in reality. The "breath test" is a potential prime example of this.

In the author's view the most important work in this thesis is that on EDLID, where a technique has been developed which for the first time enables a single non damaging laser pulse to provide all the information required to make accurate predictions as to the spatial location of damage sites.

The most important question touched on briefly in the text concerns the magnitude of the transient heating effect as the number of shots incident on the target approaches that required to cause damage. The work showed that as damage was approached the transient effect increased and that immediately after damage it was reduced, but was not able to resolve the main question which is...

"does the magnitude of the transient effect decrease before permanent damage occurs?."

If the answer to this is "yes" then monitoring the transient effect will not only yield the spatial locations of the damage sites but, with processing subject to suitable criteria will yield the extra number of shots that a sample can withstand before damaging i.e. essentially tell the operator how far away damage is. If the answer is "no" then the technique still gives the spatial locations of damage sites. Detailed investigation of this is obviously required.

There are several ways of improving the technique such as simply using more sensitive cameras and or the use of a more powerful dye laser. However based upon all of the work performed and techniques used, one method of improving the technique clearly stands out as now detailed.

The technique developed here allows the location of defects to be found but is limited to optical resolutions and provides no information on the nature i.e. physical imperfection or impurity of the defect. If the above technique could be combined with in-situ LIMA and SEM equipment then an extremely powerful setup would result since the SEM would enable the defect area (optically identified) to be inspected at high magnification and so reveal if it is a physical imperfection whilst the LIMA would enable the

composition of any defect to be determined. This would have to be all done in situ due to the problems of finding the defect if the sample was transferred between mounts and machines.

At the time of writing the author together with his supervisor ^{is} actively seeking funds for equipment with the aim of trying to answer some of the above questions. Ideally the funds would be adequate for all of the above equipment, but even funds only sufficient for improved cameras and or dye laser would enable the technique to be made substantially more sensitive.

The results of chapter 5 clearly show that colloidal silica is a major step forward in AR coating technology but that further work particularly on half wave barrier layers is urgently required to fully realize the potential of this technique.

