

UNIVERSITY OF LONDON
Imperial College of Science and Technology
The Blackett Laboratory
Optics Section

FABRICATION OF MULTISCALE RANDOM
ROUGH SURFACES WITH PRESCRIBED
SURFACE STATISTICS

Elizabeth Berry

Thesis submitted for the degree of
Doctor of Philosophy
of the University of London (1986)

ABSTRACT

Scattering of coherent radiation by a random phase screen results in the formation of speckle. The statistics of such speckle can be predicted by using a simple scattering model to describe the phase screen and its effect. The particular phase screen of interest here is a random rough surface.

The thesis begins with a general introduction outlining the historical development both of the characterisation of rough surfaces and associated theoretical models, in order to show how the current investigation has evolved from earlier work. This is followed by an overview of the contents of the thesis.

Statistical methods used to characterise a rough surface and an explanation of two models used to describe a surface are presented. The multiscale (or fractal) model which has a power law spectrum, is shown to correspond to some natural surfaces. It is thus of particular interest. As natural surfaces may not satisfy all the requirements of a sample for laser light scattering experiments, such as in the range of spatial frequencies or the index of the power law, it is more satisfactory to fabricate a surface whose properties are suitable.

A method of exposing photoresist to laser speckle to produce rough surfaces is outlined. Speckle theory is used to show how speckle could be manipulated to produce the required relief. Two practical methods of producing surfaces are described. In the first technique the resulting relief has the required statistics in only one direction, but the second results in surfaces which obey the rules in two dimensions.

Measurements made on surfaces using a stylus instrument are presented, showing that the surfaces have the expected form. Finally, results from the scattering of helium neon radiation from the surfaces are compared with the theoretical predictions of the multiscale model. Surfaces produced by the second technique are found to scatter in the manner predicted.

CONTENTS

Fabrication of multiscale random rough surfaces
with prescribed surface statistics

	<u>Page no</u>
Abstract	2
Contents	3
List of tables and plates	8
List of figures	9
Introduction	10
1. Methods for the statistical description of a continuous random (or stochastic) process	14
1.1 Introduction	15
1.2 Stochastic processes	15
i) Definition of a stochastic process	15
ii) Probability distribution function	15
iii) Moments of a stochastic process	17
iv) Stationary processes	21
v) Ergodic processes	21
1.3 Autocorrelation function	21
i) Definition	21
ii) Properties of the autocorrelation function	23
1.4 Autocovariance function	26
i) Definition	26
1.5 Correlation length	27
1.6 Structure function	27
i) Definition	27
ii) Reasons for using the structure function	29
iii) Derivation of the structure function	29
iv) What the structure function represents	30

v)	Structure function for a stationary process with zero mean	30
1.7	Power spectrum	31
i)	Definition	32
ii)	Origin of the word power in the context	34
2.	An overview of scattering by random phase screens	35
2.1	The single scale phase screen	35
2.2	The multiscale phase screen	36
2.3	Kolmogorov spectrum	37
2.4	Forms of autocorrelation function	37
2.5	Fractal geometry	40
i)	The importance of dimension (number of dimensions) in fractal geometry	41
ii)	The coastline example of a fractal set	43
iii)	Hyperbolic probability distribution	44
iv)	Some fractal parameters	45
2.6	Use of structure function instead of autocorrelation function when working with power law spectra	46
2.7	Examples of phase screens with power law spectra	48
2.8	Work described in this thesis	49
3.	Comments on rough surface fabrication	51
3.1	Reasons for surface fabrication	51
3.2	Previous techniques used	52
3.3	The coating and exposure of photoresist	55
i)	Preparation of plates	55
ii)	Application of photoresist to plates	55
iii)	Exposure of plates	61
iv)	Development	62

4.	Some statistical properties of laser speckle patterns.....	63
4.1	Types of speckle pattern.....	63
4.2	Some useful statistical properties of a single fully developed speckle pattern.....	64
	i) First order statistics.....	64
	ii) Second order statistics.....	69
4.3	Speckle pattern resulting from the incoherent addition of a number of fully developed speckle patterns.....	72
	i) Probability density function.....	72
	ii) Contrast.....	73
4.4	Speckle pattern resulting from the addition of a fully developed speckle pattern and a coherent background.....	75
	i) Probability density function.....	76
	ii) Contrast.....	78
4.5	Power spectrum resulting from the incoherent addition of a number of fully developed speckle patterns.....	78
4.6	Power spectrum resulting from the addition of a fully developed speckle pattern and a coherent background.....	78
5.	Methods used to fabricate multiscale random rough surfaces.....	82
5.1	Introduction.....	82
5A	One dimensional surface.....	82
	5A.1 Principles.....	82
	5A.2 Practice.....	90
5B	Two dimensional surfaces.....	100
5B.a	Summation of Gaussians.....	100
	5Ba.1 Principles.....	100
	5Ba.2 Practice.....	110
5B.b	Speckle and coherent background.....	113
	5Bb.1 Principles.....	113
	5Bb.2 Practice.....	115

6. The production of surfaces; mechanical measurements of surfaces	121
6.1 Parameters used in surface production	121
i) Surface 141	121
ii) Surface 200	121
iii) Surface 201	124
iv) Surface 198	124
6.2 Talysurf measurements	124
i) Probability density function	126
ii) Autocorrelation function	131
iii) Variance	134
iv) Structure function	136
v) Power spectrum	138
vi) Overall comments	140
6.3 Notes on the experimental determination of correlation functions	144
7. Estimation of surface roughness statistics by light scattering	147
7.1 Scattering theory	147
7.2 Estimating statistics from light scattering	148
7.3 Prediction of angular distribution of mean scattered intensity	151
i) Thin phase screen model	151
ii) Assumptions	152
iii) The Chandley equations	154
iv) Comparison with other approaches	157
v) Applications in this work	159
vi) Incorporation of the multiscale model into Chandley's equations	160
vii) Applications in this work	164

7.4	Experimental far field, coherent light scattering measurements	165
i)	Experimental arrangements	165
ii)	Equations to be applied	170
iii)	Measurements from one dimensional multiscale surface. Surface 141	171
iv)	Measurements from two dimensional multiscale surface. Surface 200	173
v)	Measurements from single scale surfaces. Surfaces 201, 198	175
7.5	Prediction of angular distribution of mean scattered intensity from knowledge of autocorrelation function of surface heights	179
	Conclusion	183
	Appendix. Derivation of the structure function corresponding to an unmodified power law spectrum	185
	Acknowledgements	188
	References	189
	Published papers	197

LIST OF TABLES

Number	Page
3.1	56
5.1	107
6.1	122
6.2	123

LIST OF PLATES

3.1 Linnik interferogram of an example of ridging on a flow coated plate.	60
5.1 Speckle from Lloyd's mirror arrangement of figure 5.9. a) without mirror, b) with mirror.	93
5.2 Speckle from cube beamsplitter arrangement of figure 5.12. a) slit separation 5mm, b) slit separation 28mm.	96

LIST OF FIGURES

Number	Page	Number	Page
1.1	18	5.17	116
1.2	28	5.18	116
1.3	28	5.19	116
1.4	33	5.20	119
2.1	48	5.21	119
3.1	53	6.1	128
4.1	68	6.2	129
4.2	68	6.3	130
4.3	70	6.4	133
4.4	70	6.5	133
4.5	74	6.6	133, 135
4.6	74	6.7	137
4.7	77	6.8	137
4.8	77	6.9	137
5.1	84	6.10	139
5.2	84	6.11	141
5.3	85	6.12	142
5.4	85	6.13	146
5.5	87	7.1	157
5.6	89	7.2	166
5.7	89	7.3	172
5.8	89	7.4	172
5.9	91	7.5	174
5.10	91	7.6	174
5.11	95	7.7	176
5.12	95	7.8	176
5.13	99	7.9	178
5.14	108	7.10	178
5.15	111	7.11	181
5.16	111	7.12	181

INTRODUCTION

Investigation of any rough surface, or more generally, inhomogeneous medium, can involve scattering of electromagnetic radiation of a wavelength appropriate to the scale of the sample. In order to predict how the radiation is scattered some sort of theoretical model is generally invoked to describe the scatterer. The medium may be adequately described by its probability density function and autocorrelation function. The probability density function is nearly always taken to be Gaussian both because it allows considerable simplification of the analysis, and as it is observed in practice. It is the choice of a suitable form for the autocorrelation function which provides more difficulty.

The earliest work (eg Beckmann & Spizzichino 1963) assumed an autocorrelation function which was Gaussian in form. This model is called a single scale model, the characteristic scale size of the screen being given by the width of the autocorrelation function. Such a model applies well to certain cases; for example scattering by a liquid crystal (Pusey and Jakeman 1975), a thermal plume (Parry et al 1977), and a ruffled water surface (Walker and Jakeman 1982); but it is not universally applicable. An alternative form of autocorrelation function, where the surface cannot be characterised by a single length scale is required.

The roots of this multiscale approach lie in the theory of turbulence. Richardson (1922) suggested a description for the structure of turbulent flow. The flow can be considered to consist of a series of separate volumes of fluid, the sizes of which cover all possible scales between an inner and an outer scale size. Kolmogorov (1941) showed that such a flow would have a structure function of velocity components which followed a power law with index $-2/3$. The connection with wave propagation was made by Tatarski (1961) who showed that the structure functions of both refractive index and phase fluctuations obeyed the same power law. In addition, the power spectrum of the phase fluctuations also follows a power law. Measurements confirming the validity of such a model for the scattering of large wavelength electromagnetic radiation are fairly common

(Intriligator & Wolfe 1970; Clifford et al 1971). Data published by Sayles and Thomas (1978) showed that many rough surfaces too have power law spectra, albeit over a fairly short range of spatial frequency. Surfaces they featured include a motorway, a lunar sea, a grit blasted surface, and a ground metal surface.

Interest in phase screens containing a variety of scale sizes was heightened with the publication of Mandelbrot's work on fractals (1975, 1982). Mandelbrot expounded a new geometry which allowed objects to have a non-integral dimension, and described natural phenomena more closely than classical geometry was able. A feature of a typical fractal set is that it contains an infinite range of scale sizes, the fluctuations of the larger ones being greater than those of the smaller. This corresponds to a power law spectrum without bounds.

Describing a phase screen using a power law spectrum then seems to be a very promising approach. Much has been written (Berry 1979; Jakeman & McWhirter 1977) predicting how electromagnetic radiation will be scattered from such a screen, but experimental confirmation is scarce, certainly on a laboratory scale. What is needed is a surface whose statistics are known, and preferably of a form which allows the use of simplifying assumptions in the scattering prediction. In addition, it should be possible to use the surface in both reflection and transmission using laser radiation. The purpose of this work has been to prepare such a surface.

In the first chapter of this thesis the statistical methods for describing a continuous random process are outlined, with particular emphasis on those statistics of relevance to the work which follows. These are the probability distribution function, the autocorrelation function, the structure function and the power spectrum.

Chapter two gives a general overview of scattering by random phase screens. The historical origins of the power law spectrum are described, and corresponding forms for the autocorrelation function suggested. The use of the structure function in place of the autocorrelation function is explained. In this chapter too, a simple introduction to the principles of fractal geometry is provided. The link between fractals and power law spectra is pointed out.

The third chapter consists of a brief section outlining the reasons for producing a surface, followed by a description of methods used by other workers. The decision to use a method involving the exposure of photoresist to speckle patterns from an argon ion laser is justified. The chapter ends with a description of the techniques involved in using photoresist.

Chapter four expounds the theory of speckle which will be applied when making surfaces. Two alternative means of obtaining a Gaussian probability density function of intensities are suggested. The first involves the incoherent summation of ten or more independent fully developed speckle patterns. The second is by the coherent addition of a speckle pattern with a plane background wave, of intensity about ten times the mean speckle intensity. The corresponding power spectra are given.

Chapter five is subdivided into parts A and B. Part A shows how a surface which has a power law spectrum of a desired index in one direction of spatial frequency could be obtained in principle. Such a surface is described as one dimensional. The practical arrangement used to produce such a surface is then described. In part B two methods are described, in principle and practice, for producing an isotropic surface. Only the first of these methods, which utilises the summation of fully developed speckle patterns was successful in practice.

The surfaces were produced to be used for testing theories of the scattering of electromagnetic radiation, and this is reflected in chapter seven. An outline is given of how light scattering can be used to estimate surface statistics such as the standard deviation and probability density function. The relation of the angular distribution of mean scattered intensity to the typical scale size and standard deviation of the surface is derived for both the single and multiscale case. The experimental arrangement for verifying these equations using fabricated surfaces is described, and measurements presented. Finally, a method of predicting the angular distribution of mean scattered intensity from a knowledge of the autocorrelation function of surface heights is discussed.

Mechanical measurements made using a Talysurf stylus instrument are presented in chapter six, following a listing of the values chosen for the parameters of the methods adumbrated in the previous chapter.

Chapter seven is followed by concluding remarks on the efficacy of the methods tried for fabricating rough surfaces.

Included at the back of this thesis are two published papers concerning the work described herein. The first (Berry 1985a) was presented at Optics '84 at Keele University. The second, (Berry 1985b), which contains Talysurf measurements from the one dimensional surface, was presented at Image Science '85 in Helsinki.

CHAPTER ONE

METHODS FOR THE STATISTICAL DESCRIPTION OF A CONTINUOUS
RANDOM (OR STOCHASTIC) PROCESS

1.1 Introduction

In this chapter the statistics used to describe any continuous random process are defined, and some of their properties are investigated. The particular random process encountered throughout this thesis is the variation of surface height with position on a surface.

The statistics considered in this chapter are

Probability distribution function	$P(h)$	15
Probability density function	$p(h)$	16
Mean	μ	18
Second moment	μ_2	19
Variance	σ^2	19
Standard deviation	σ	20
Autocorrelation	$C(X)$	21
Normalised autocorrelation function	$C_{\text{NORM}}(X)$	22
Autocovariance	$\gamma(X)$	26
Correlation length	ℓ	27
Structure function	$D(X)$	27
Power spectrum	$G(u)$	31

1.2 Stochastic processes

1.2 i Definition of a stochastic process

The word 'stochastic' is derived from the Greek, and means 'pertaining to chance'. It is used interchangeably with the word random.

The random process with which we shall be particularly concerned in this thesis is the variation of surface height with position on the surface. For continuity the random process shall be $H(x)$ throughout.

A stochastic process can be defined as a collection of random variables, (Papoulis 1985)

$$\left\{ H(x), x \in X_1 \right\} \quad (1.1)$$

where X_1 denotes the set of values of position at which the process is defined. For a single outcome of the process, there is only one observation of each random variable. Each of these values changes for different outcomes according to probabilistic laws.

Statistical problems usually involve estimating the properties of a population from a sample. There is an infinite set of possible random functions which could be observed. This infinite set is called the ensemble. Each member of the ensemble is a possible realisation of the stochastic process. A measurement of height fluctuations can thus be thought of as one realisation. This series is denoted by $h(x)$. In the case of such a measurement there is only a notional population. The analysis is then concerned with evaluating the properties of the probability model which generated the series observed.

1.2 ii Probability distribution function

The n 'th order joint probability distribution function of $h(x_1), \dots, h(x_n)$ is defined as (Papoulis 1985 p208)

$$P(h_1, \dots, h_n; x_1, \dots, x_n) = \text{prob} \left[h(x_1) < h_1, \dots, h(x_n) < h_n \right] \quad (1.2)$$

However, as the process $h(x)$ is usually stationary (section 1.2 iv), the expression is written

$$P(h_1, \dots, h_n)$$

with the dependence on x implicit.

The n 'th order joint probability density function is defined as the n 'th derivative of P ,

$$p(h_1, \dots, h_n) = \frac{\partial^n}{\partial h_1 \dots \partial h_n} P(h_1, \dots, h_n) \quad (1.3)$$

Also,

$$P(h_1, \dots, h_n) = \int_{-\infty}^{h_1} \dots \int_{-\infty}^{h_n} p(h_1', \dots, h_n') dh_1' \dots dh_n' \quad (1.4)$$

If $P(h_1, \dots, h_n; x_1, \dots, x_n)$ is known for every h_i, x_i , and n then the statistical properties of the stochastic process can be determined. In other words, equation (1.3) completely specifies a stochastic process. (Papoulis 1985 p215). However, it is often the case that only the first and second order distribution functions are required. The mean, variance, and autocorrelation of $h(x)$ can all be expressed in terms of the distribution function (sections 1.2 iii, 1.3 i).

First order probability density function

The first order probability density function (pdf), $p(h)$, gives the probability of the value of $h(x)$ being within a certain range. The total area under the curve described by $p(h)$ is unity i.e.

$$\int_{-\infty}^{\infty} p(h) dh = 1 \quad (1.5)$$

The probability density function is often Gaussian in form

$$p(h) = (1/\sqrt{2\pi}\sigma) \exp \left\{ -(h-\mu)^2 / 2\sigma^2 \right\} \quad (1.6)$$

where μ is the mean, and σ the standard deviation of $h(x)$. Two realisations of this curve are shown in figure 1.1. The curves are symmetric about $h=\mu$, the mean of the function (zero in this case). The shape of the curve depends on the value of the standard deviation σ . The larger σ , the more spread out the curve. Whatever values μ and σ take, the Gaussian distribution has the following property. About one observation in three lies more than one standard deviation from the mean. One in twenty lies more than 2σ from the mean. Less than one in 300 lies more than 3σ from the mean. Gaussian curves presented in this work will always be plotted normalised, with zero mean and unit standard deviation to allow easy comparison with measurements.

1.2 iii Moments of a stochastic process

Moments

A useful way of describing a stochastic process is to give the moments of the process. The n 'th moment of a probability distribution of the random variable H is defined to be the statistical average of the n 'th power of H .

$$\langle H^n(x) \rangle = \int_{-\infty}^{\infty} H^n(x) p(H_1, \dots, H_N) dH_1, \dots, dH_N \quad (1.7)$$

$p(H_1, \dots, H_N)$ is the N 'th order probability density function.

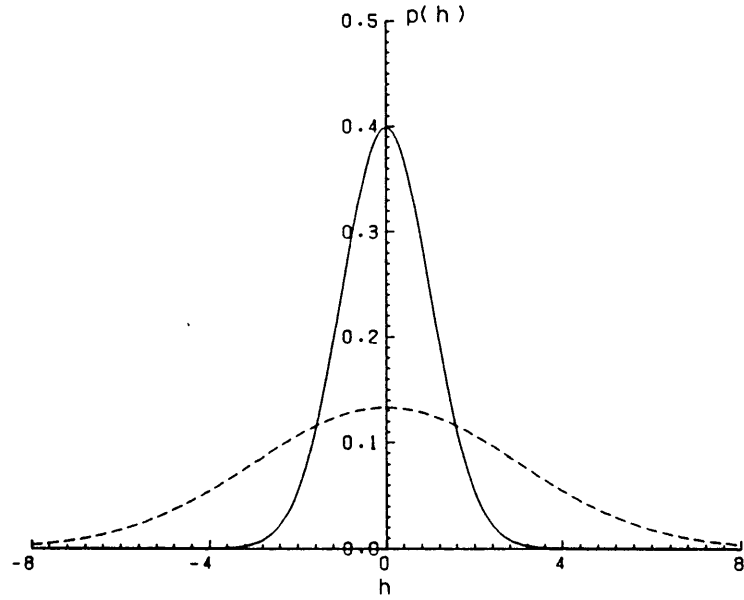


Figure 1.1. Comparison of two Gaussian probability distributions, with different standard deviations. Solid curve $\sigma = 1$; dashed curve $\sigma = 3$.

The Mean

The mean is a measure of location, which means that it gives a quantity which is typical of the data.

The population mean is given by the first moment of the first order probability density function, i.e.

$$\mu = \int_{-\infty}^{\infty} H(x) p(H) dH \quad (1.8)$$

therefore

$$\mu = \overline{H(x)} \quad (1.9)$$

The sample mean, $\overline{h(x)}$, is calculated using

$$\overline{h(x)} = \sum_{i=1}^N \frac{h(x_i)}{N} \quad (1.10)$$

Note that the set of data used is a sample taken from the population. The sample mean $\overline{h(x)}$ is an estimate of the population mean μ .

A zero mean process is one where the values of $h(x)$ are both positive and negative, and the mean has a value of zero. Any process may be converted to one with a zero mean, simply by subtracting the sample mean from each value.

The second moment of the first order probability density function, μ_2 , is given by

$$\mu_2 = \int_{-\infty}^{\infty} H^2(x) p(h) dH \quad (1.11)$$

$$\mu_2 = \overline{H^2(x)} \quad (1.12)$$

Higher order moments are sometimes used.

Central moments

The n'th central moment of the probability distribution of a random variable H is defined as

$$\langle (H(x) - \overline{H(x)})^n \rangle = \int_{-\infty}^{\infty} (H(x) - \overline{H(x)})^n p(H_1, \dots, H_N) dH_1 \dots dH_N \quad (1.13)$$

The variance

The variance is the second central moment of the first order probability function. The variance is a measure of spread. This means that it indicates the variability of the data, which is information not given by the mean.

$$\sigma^2 = \int_{-\infty}^{\infty} (H(x) - \overline{H(x)})^2 p(h) dh \quad (1.14)$$

$$= \overline{(H(x) - \overline{H(x)})^2} \quad (1.15)$$

For a set of data, with sample mean $\bar{h}$

$$s^2 = \frac{\sum_{i=1}^N (h(x_i) - \bar{h})^2}{(N-1)} \quad (1.16)$$

$(N-1)$ is used rather than N in this expression because $\bar{h}$ is not the true population mean μ . Equation (1.16) expresses s^2 as an unbiased estimator of σ^2 . In practice the difference between using N and $(N-1)$ in the denominator is minimal.

Note that

$$\begin{aligned} \sigma^2 &= \overline{(H(x) - \overline{H(x)})^2} \\ &= \mu_2 - \mu^2 \end{aligned} \quad (1.17)$$

The standard deviation

The standard deviation σ has the same units as $h(x)$, and is the square root of the variance.

In this work the population mean is never known. The estimators of the mean and variance, i.e. $\bar{h}$ and s^2 will be used, but will be given using the population symbols, i.e. mean μ and variance σ^2 .

It is convenient at this point to introduce the concepts of stationarity and ergodicity.

1.2 iv Stationary processes

This means that the statistics are independent of the particular region on the surface chosen to make measurements, i.e. the joint density of $H(x)$ and $H(x+X)$ is independent of x . Thus the mean and the variance are both constants independent of the value of x . (Papoulis 1985 p220).

1.2 v Ergodic processes

If a process is ergodic, then statistical properties derived from an observed series of height fluctuations are identical to those derived from the ensemble. Ensemble averages are denoted by triangular brackets $\langle \rangle$. Thus for an ergodic process

$$\sigma^2 = \overline{(h(x) - \overline{h(x)})^2} \quad (1.18)$$

$$= \langle (H(x) - \langle H(x) \rangle)^2 \rangle \quad (1.19)$$

All ergodic processes are also stationary. Stationary processes are not necessarily ergodic.

From this point, processes will be assumed to be ergodic, and all averages will be denoted using $\langle \rangle$, rather than overbars.

1.3 Autocorrelation function

1.3 i Definition

The autocorrelation function may be defined directly from the second order probability density function. It is the expected value (mean) of the product $h(x_1)h(x_2)$

$$C(x_1, x_2) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} h_1 h_2 p(h_1, h_2; x_1, x_2) dh_1 dh_2 \quad (1.20)$$

i.e.

$$C(X) = \langle h(x) \cdot h(x+X) \rangle \quad (1.21)$$

When the lag $X=0$

$$C(0) = \langle h(x)^2 \rangle = \mu_2 \quad (1.22)$$

For a zero mean process

$$C(0) = \sigma^2 \quad (1.23)$$

For a complex process

$$C(x_1, x_2) = \langle h(x_1) \cdot h^*(x_2) \rangle \quad (1.24)$$

i.e.

$$C(X) = \langle h(x) \cdot h^*(x+X) \rangle \quad (1.25)$$

Often a normalised autocorrelation is used, where

$$C_{\text{NORM}}(X) = C(X)/C(0) \quad (1.26)$$

1.3 ii Properties of the autocorrelation function

Consider a stationary process $h(x)$, with a mean μ , variance σ^2 , and an autocorrelation function $C(X)$.

Property one

The autocorrelation function is an even function of the lag X .
See property three for complex processes.

$$C(X) = C(-X) \quad (1.27)$$

This can be shown as follows

$$C(X) = \langle h(x) \cdot h(x+X) \rangle \quad (1.28)$$

but $h(x)$ is a stationary function, therefore

$$\begin{aligned} C(X) &= \langle h(x-X) \cdot h(x) \rangle \\ &= C(-X) \end{aligned} \quad (1.29)$$

Property two

$$\left| C_{\text{NORM}}(X) \right| \leq 1 \quad (1.30)$$

This is proved by noting that

$$\left\langle \left\{ \lambda_1 h(x) + \lambda_2 h(x+X) \right\}^2 \right\rangle \geq 0 \quad (1.31)$$

for any constants λ_1, λ_2

$$\begin{aligned} & \left\langle \left\{ \lambda_1 h(x) + \lambda_2 h(x+X) \right\}^2 \right\rangle = \\ & \lambda_1^2 \langle h(x)^2 \rangle + \lambda_2^2 \langle h(x+X)^2 \rangle + 2\lambda_1\lambda_2 \langle h(x)h(x+X) \rangle \end{aligned}$$

therefore

$$(\lambda_1^2 + \lambda_2^2) C(0) + 2\lambda_1\lambda_2 C(X) \geq 0 \quad (1.32)$$

$$\text{If } \lambda_1 = \lambda_2 = 1$$

$$\begin{aligned} 2C(0) + 2C(X) &\geq 0 \\ C(0) &\geq -C(X) \\ C(X) &\geq -C(0) \end{aligned} \quad (1.33)$$

But

$$C_{\text{NORM}}(X) = C(X)/C(0) \quad (1.26)$$

therefore

$$C_{\text{NORM}}(X) \geq -1 \quad (1.34)$$

$$\text{If } \lambda_1 = 1, \lambda_2 = -1$$

$$\begin{aligned} 2C(0) - 2C(X) &\geq 0 \\ C(0) &\geq C(X) \end{aligned} \quad (1.35)$$

But

$$C_{\text{NORM}}(X) = C(X)/C(0) \quad (1.26)$$

therefore

$$C_{\text{NORM}}(X) \leq 1 \quad (1.36)$$

i.e.

$$\left| C_{\text{NORM}}(X) \right| \leq 1 \quad (1.30)$$

Property three

The autocorrelation is Hermitian for complex processes, from (1.24)

$$C(x_2, x_1) = \langle h(x_2) \cdot h^*(x_1) \rangle \quad (1.37)$$

$$C(x_2, x_1) = C^*(x_1, x_2) \quad (1.38)$$

i.e.

$$C(-X) = C^*(X) \quad (1.39)$$

Property four

A stationary joint Gaussian process is completely determined by its mean, variance and autocorrelation function. However, a process is not uniquely defined by its autocorrelation function, many non-Gaussian processes can have the same autocorrelation function. (Jenkins & Watts, 1968 p170).

1.4 Autocovariance function

The autocovariance function gives a measure of the distribution of variations of $h(x)$ in the x direction.

1.4 i Definition

The autocovariance function $\gamma(x_1, x_2)$, is defined as

$$\gamma(x_1, x_2) = \langle \left[(h(x_1) - \langle h(x_1) \rangle) (h(x_2) - \langle h(x_2) \rangle) \right] \rangle \quad (1.40)$$

For a stationary process the statistics do not vary over the surface so only the difference $x_1 - x_2$ matters, and

$$\langle h(x_1) \rangle = \langle h(x_2) \rangle \quad (1.41)$$

If $x_1 - x_2 = X$, then

$$\langle h(x) \rangle = \langle h(x+X) \rangle \quad (1.42)$$

and

$$\gamma(X) = \langle \left[(h(x) - \langle h(x) \rangle) (h(x+X) - \langle h(x) \rangle) \right] \rangle \quad (1.43)$$

$$\gamma(X) = \langle h(x) \cdot h(x+X) \rangle - \langle h(x) \rangle^2 \quad (1.44)$$

Thus, from (1.21)

$$\gamma(X) = C(X) - \langle h(x) \rangle^2 \quad (1.45)$$

The variance, σ^2 , is the autocovariance when $X = 0$, i.e.

$$\gamma(0) = \langle \left[h(x) - \langle h(x) \rangle \right]^2 \rangle \quad (1.46)$$

$$= \langle h(x)^2 \rangle - \langle h(x) \rangle^2 = \sigma^2 \quad (1.47)$$

For a zero mean process, the autocovariance

$$\gamma(X) = \langle h(x) \cdot h(x+X) \rangle \quad (1.48)$$

So, for a zero mean process the autocorrelation function and the autocovariance are the same.

1.5 Correlation length

A correlation length may be defined from the autocorrelation function as being a typical scale size in the x direction of variations in the original function $h(x)$. If the autocorrelation function is Gaussian, then the correlation length, ℓ , may be defined as the $1/e$ width of the function. ℓ is the value of the lag when the autocorrelation function has fallen to $1/e$ of its value at lag zero. This is illustrated in figure 1.2.

For other functions the correlation length may be defined differently, but qualitative conclusions can always be drawn from the shape of the autocorrelation function. Two types of profile are illustrated in figure 1.3 together with corresponding autocorrelation functions. The autocorrelation function is narrower when the scale lengths of the fluctuations in the profile are shorter.

1.6 The structure function

1.6 i Definition

The structure function is closely related to the autocorrelation function. It has, however, a number of features which make it a useful supplement to the autocorrelation function, and it can sometimes

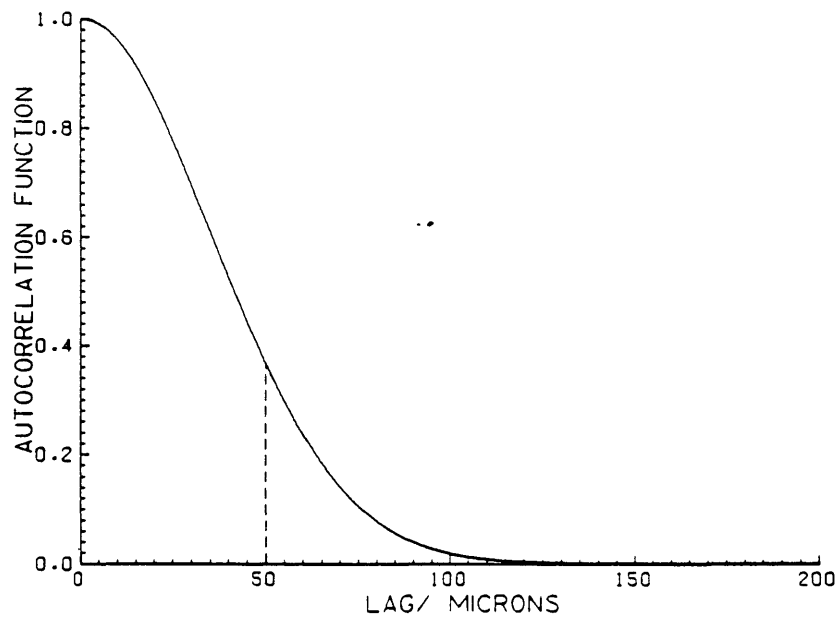


Figure 1.2. Gaussian autocorrelation function, with correlation length 50 μm .

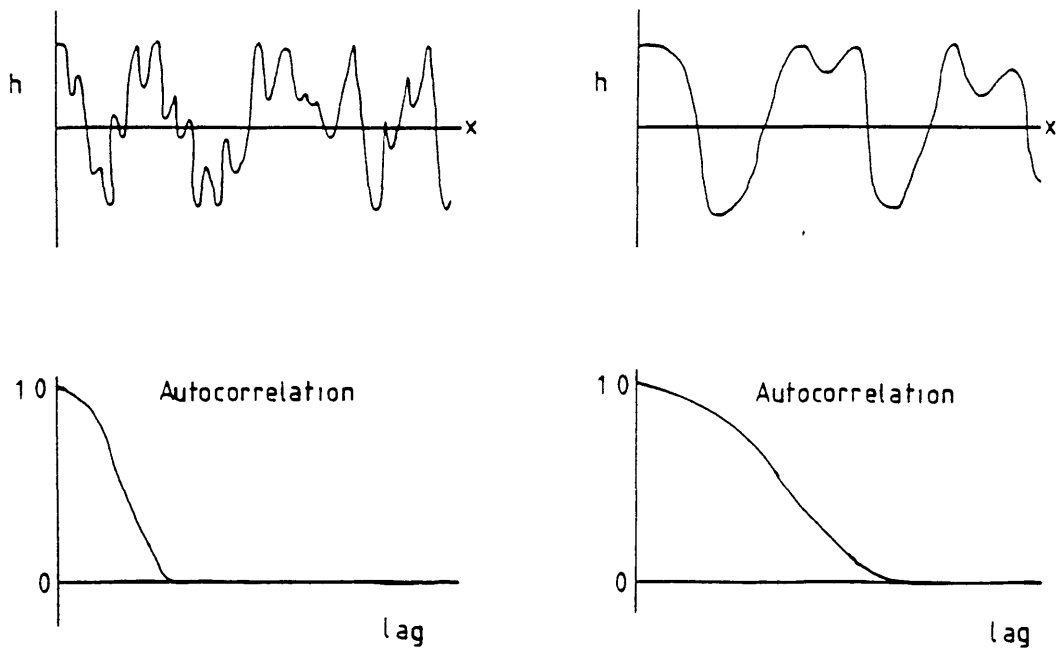


Figure 1.3. Two example surface profiles, shown with their corresponding autocorrelation functions, to illustrate the link between the width of the autocorrelation function and the correlation length.

replace it.

The structure function is defined as

$$D(X) = \langle \left[h(x) - h(x+X) \right]^2 \rangle \quad (1.49)$$

1.6 ii Reasons for using the structure function

The structure function was first introduced in turbulence theory, in the papers of Kolmogorov (1941a, b); it was found that the autocorrelation function was not particularly applicable because the processes to be described were not stationary random processes. Some meteorological variables of the atmosphere undergo smooth, slow changes. Such slow changes, e.g. of humidity or temperature could be considered to be either slow random fluctuations, or changes in the mean. The expression (1.43) for the autocovariance (and hence the autocorrelation) involves the means $\langle h(x) \rangle$, $\langle h(x+X) \rangle$, but for a stationary process they are assumed to be the same. For a non-stationary process it is helpful to consider

$$h'_X(x) = h(x) - h(x+X) \quad (1.50)$$

For not too large X , slow changes in $h(x)$ do not affect the value of this difference. Even if $h(x)$ is not stationary, then $h'_X(x)$ can be, though possibly only approximately. This means that $h(x)$ can be described as a random function with stationary increments.

1.6 iii Derivation of the structure function

The autocorrelation function of the increment function $h'_X(x)$

$$C_{h'}(X) = \langle h'_X(x) h'_X(x+X) \rangle \quad (1.51)$$

$$= < \left[h(x) - h(x+X) \right] \left[h(x+X) - h(x+X+X) \right] > \quad (1.52)$$

Now using $(a-b) \cdot (c-d) =$

$$(a-d)^2 + (b-c)^2 - (a-c)^2 - (b-d)^2 \quad (1.53)$$

$$\begin{aligned} C_{h^2}(X) &= \frac{1}{2} < (h(x) - h(x+2X))^2 > \\ &+ \frac{1}{2} < (h(x+X) - h(x+X))^2 > \\ &- \frac{1}{2} < (h(x) - h(x+X))^2 > \\ &- \frac{1}{2} < (h(x+X) - h(x+2X))^2 > \end{aligned} \quad (1.54)$$

i.e. the autocorrelation function $C_{h^2}(X)$ is expressed as a linear combination of functions

$$D = < \left[h(x+nX) - h(x+(n+1)X) \right]^2 > \quad (1.55)$$

For $C_{h^2}(X)$ to depend only on X , then D must also be a function of X . This function is the structure function

$$D(X) = < \left[h(x) - h(x+X) \right]^2 > \quad (1.56)$$

It is a basic characteristic of a random function with increments whose distribution is a stationary random function.

1.6 iv What the structure function represents

The value of $D(X)$ measures the intensity of the fluctuations of $h(x)$ of length $\leq X$.

1.6 v Structure function for a stationary process with zero mean

$$D(X) = < \left[h(x) - h(x+X) \right]^2 > \quad (1.57)$$

$$D(X) = \langle \left[h(x) \right]^2 \rangle - 2 \langle h(x)h(x+X) \rangle + \langle \left[h(x+X) \right]^2 \rangle \quad (1.58)$$

But

$$\langle \left[h(x) \right]^2 \rangle = \langle \left[h(x+X) \right]^2 \rangle = c(0) \quad (1.59)$$

due to stationarity, therefore

$$D(X) = 2 \left\{ c(0) - c(X) \right\} \quad (1.60)$$

$$= 2c(0) \left\{ 1 - c(X)/c(0) \right\} \quad (1.62)$$

$$= 2c(0) \left\{ 1 - c_{\text{NORM}}(X) \right\} \quad (1.63)$$

For large X , $c_{\text{NORM}} \rightarrow 0$, therefore

$$D(X) = 2c(0) \quad (1.64)$$

From (1.47), for a zero mean process

$$c(0) = \gamma(0) = \sigma^2$$

therefore

$$D(X) = 2\sigma^2 \quad (1.65)$$

The usefulness of the structure function, even for a stationary process will be considered in section 2.6.

1.7 Power Spectrum

This function is also known as the power spectral density function. It is complementary to the autocorrelation function, and is useful for considering the distribution of spatial frequencies in the process.

It should be emphasised that the autocorrelation function and the power spectrum are equivalent ways of describing a stationary random function. The autocorrelation function applies in the spatial domain, the power spectrum in the frequency domain.

Figure 1.4 shows how the power spectra from two similar traces are quite distinct. There is a deterministic signal at 15 cycles per mm in trace b).

1.7 i Definition

The power spectrum of a process $h(x)$ is the Fourier transform of its autocorrelation function (Papoulis 1985 p265)

$$G(u) = \int_{-\infty}^{\infty} C(X) \exp (-2\pi i u X) dX \quad (1.66)$$

Thus, we may write

$$C(X) = \int_{-\infty}^{\infty} G(u) \exp (2\pi i u X) du \quad (1.67)$$

Now,

$$C(-X) = C^*(X) \quad (1.39)$$

So the power spectrum is always a real function of u . Also, when $h(x)$ is real, then $C(X)$ is real and even, and so $G(u)$ is also real and even. In this special case equations (1.66) and (1.67) may be written

$$G(u) = \int_{-\infty}^{\infty} C(X) \cos 2\pi u X dX \quad (1.68)$$

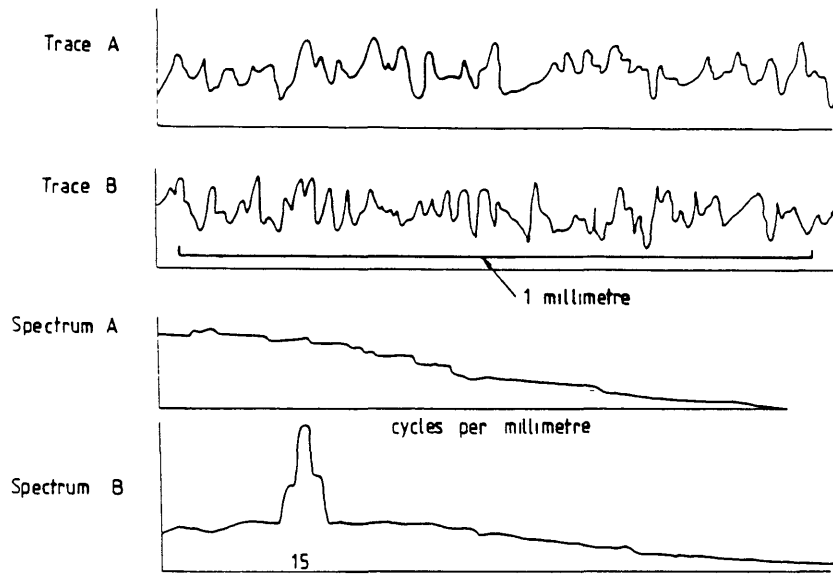


Figure 1.4. Two example surface profiles with their corresponding power spectra, to illustrate how apparently similar profiles can have very different spectra.

$$C(X) = \int_{-\infty}^{\infty} G(u) \cos 2\pi uX \, du \quad (1.69)$$

The functions $C(X)$ and $G(u)$ are Fourier transforms of each other. The Fourier transform of any autocorrelation function $C(X)$ must be non-negative. If the Fourier transform is negative at any point, then $C(X)$ cannot be the autocorrelation function of a stationary random process. It is also true that if there is a function $C(X)$ whose Fourier transform is non-negative, then $C(X)$ is the autocorrelation function of an existent stationary random process. Thus any model proposed for a surface must have a non-negative power spectrum, or the corresponding autocorrelation function cannot represent a stationary random process.

1.7 ii Origin of the word "power" in the context

The word "power" arose from the use of this function for series which vary in time and not in space, for example the flow of current. The power spectrum is then a direct measure of the power. (given by the time average of the energy) of each component frequency in the signal. In Optics where the variations occur in space the function is often known as the Wiener spectrum.

CHAPTER TWO

AN OVERVIEW OF SCATTERING BY RANDOM PHASE SCREENS

In this chapter the development of theory to explain the scattering of electromagnetic waves by scintillation of atmospheric, ionospheric, and interplanetary media is described. The Kolmogorov theory of turbulence leads to the prediction of a power law spectrum of refractive index fluctuations, and it is noted that such a phase screen is quite widely applicable. The basic concepts of fractal geometry are presented to demonstrate the equivalence of the approach of Berry (1979) with those previously described. The frequent use of the structure function, rather than the autocorrelation function with which it is often considered interchangeable using equation (1.63), is explained.

A survey of phase screens with power law spectra precedes a description of the properties of the surfaces whose production is described in the remainder of this thesis.

2.1 The single scale phase screen

A model frequently used to describe a rough surface or phase screen from which light is scattered, is a single scale one (Jakeman 1980). Such a model assumes that the screen has a Gaussian autocorrelation function. The width of this curve corresponds to the characteristic scale size of the screen. The size of irregularities in the emerging wave front is related to the size of those of the screen with a dependence on the standard deviation of the phase variations (Goodman 1984 & section 7.3 iv). The power spectrum of the phase screen is also Gaussian. If the $1/e$ width of the autocorrelation function is L/π , then the width of the power spectrum is $1/L$. The probability distribution of phase variations is Gaussian.

Early examples of the application of such a model were in the field of ionospheric scattering and interplanetary scintillation. Ratcliffe (1956) considered the Fresnel approximation for a thin phase screen and using scalar diffraction theory (section 7.3).

The Fresnel region is most appropriate for the investigation of ionospheric scattering where the parameters are such that Fraunhofer patterns are not observed on Earth. For phase fluctuations less than 1 radian, it can be assumed that the autocorrelation function of the amplitudes at the ground is the same as that of the electron densities in the ionosphere. The model was extended by Mercier (1962) to include strong scattering.

Early work on light scattering from rough surfaces, where measurements are made in the far field, did not use the thin phase screen model but scalar diffraction theory is still applied. Davies (1954) and Beckmann (and Spizzichino 1963; 1967) developed the first predictions about the distribution of ensemble-averaged far-field scattered coherent light. They considered rough surfaces rather than thin phase screens, but the surface was described in much the same way as in the larger scale approaches. It had a Gaussian probability distribution, autocorrelation function and power spectrum.

The thin phase screen model was later used by Chandley (1979), Goodman (1984), Jakeman and Pusey (1973) and Jakeman and McWhirter (1977) to predict far field statistics. The similarities between these approaches is described in chapter seven.

2.2 The multiscale phase screen

Up to this point the single scale model has been applied to situations ranging from the measurement of near field radio wave scintillation produced by the atmospheric, interplanetary (Ratcliffe 1956) or interstellar (Rickett 1969) media, to far field fluctuations produced in laser radiation scattered by, for example, ground glass (Chandley 1979).

It was in the field of radio wave diffraction that the weaknesses of the single scale model first became apparent. The theory of turbulence developed in the early 1940's by Kolmogorov (1941a, b) showed that the spectrum of the phase variations would have a power law form, between certain inner and outer scale sizes, rather than the Gaussian one described earlier. Although ionospheric results

at first appeared to be well described by a Gaussian spectrum (Hewish et al 1964), later measurements suggested that a power law spectrum would be more appropriate (Intriligator and Wolfe 1970). The power law spectrum was already widely applied to atmospheric turbulence (Clifford et al 1971).

2.3 Kolmogorov spectrum

Richardson (1922) suggested a description of the structure of turbulent flow. The flow was considered to consist of a series of volumes of fluid, each separate from one another. The sizes of these vortices covers all possible scales between an inner and an outer scale size. The largest vortices absorb energy, which is then transferred sequentially through vortices of decreasing size to the smallest. When the energy reaches the smallest variations it is dissipated as heat due to viscosity.

Kolmogorov (1941) and Obukhov (1941) made an analytical interpretation of this model. Their reasoning resulted in a relation known as the law of two thirds, which gives the structure function of the velocity components as a power law of index $-2/3$. Tatarski (1961) extended the argument to consider the structure function of refractive index fluctuations in the flow. He showed that the structure functions of refractive index and phase fluctuations obey the law of two thirds. He also showed that if the spectrum is worked out from a one dimensional representation of the structure function then it follows a power law of index $-5/3$. If the three dimensional structure function is used, then the power spectrum has a power law index of $-11/3$.

2.4 Forms of autocorrelation function

Clearly if the spectrum of refractive index fluctuations is not Gaussian, then a different form must also be assumed for the autocorrelation function.

If no limiting inner or outer scale sizes are specified, then the resulting expressions for the power spectrum and structure function are analytically convenient, if not physically applicable. The power spectrum is written

$$G(u) \propto u^{-(\nu+1)} \quad (2.1)$$

It is shown in the appendix that the corresponding structure function is

$$D(X) \propto |X|^\nu \quad (2.2)$$

This case was considered by Rumsey (1975). For a two dimensional screen $1 < \nu < 5$; for a one dimensional screen $0 < \nu < 2$. This approach is an improvement on that of Buckley (1971) as it allows the power spectrum to fall off as a power law at high frequency.

Jakeman and McWhirter (1977) suggest forms for the phase autocorrelation function which are truncated, and give an outer scale to the power spectrum. These correlation functions are given by

1). Truncated parabolic

$$C(X) = \begin{cases} 1 - X^2/\Gamma^2 & |X| < \Gamma \\ 0 & \text{otherwise} \end{cases} \quad (2.3)$$

2). Truncated linear

$$C(X) = \begin{cases} 1 - |X|/\Gamma & |X| < \Gamma \\ 0 & \text{otherwise} \end{cases} \quad (2.4)$$

Were these expressions not truncated they would correspond to the general expression

$$C(X) = 1 - |X/\Gamma|^\nu \quad (2.5)$$

From equations (1.63) and (2.5)

$$D(X) \propto |X|^{-\nu} \quad (2.6)$$

i.e. the structure function (2.2) for an unbounded power spectrum.

Jakeman and McWhirter (1977) show that, for a deep phase screen, i.e. when the mean square phase deviation $\gg 1$ radian, then truncation of the linear phase autocorrelation function has no effect on the scattered mean intensity with angle.

Note that this model is still somewhat unphysical. There is no inner scale incorporated in the model and scale sizes continue down to values less than the wavelength of the scattered wave. Uscinski, Booker and Mariani (1981) consider a phase spectrum given by

$$G(u) \propto (1 + \Gamma^2 u^2)^{-m/2} \quad (2.7)$$

$$1 < m < 4$$

where Γ is the outer scale size.

The inner scale size is considered so small as to be negligible.

For the case

$$G(u) = \frac{\Gamma}{\pi(1 + \Gamma^2 u^2)} \quad (2.8)$$

where $m = 2$, the corresponding correlation function can be found by inverse Fourier transformation and is given by

$$C(X) = \exp(-|X|/\Gamma) \quad (2.9)$$

where $m = 4$, and

$$G(u) = \frac{2\Gamma}{\pi(1 + \Gamma^2 u^2)^2} \quad (2.10)$$

$$C(X) = (1 + |X|/\Gamma) \exp(-|X|/\Gamma) \quad (2.11)$$

Note that for $\Gamma \gg |X|$, the exponential term may be expanded, and the correlation functions reduce to the linear and parabolic expressions (2.4) and (2.3) given by Jakeman and McWhirter (1977).

Berry (1979) tackled the problem of scattering by a random phase screen using a completely different terminology - that of fractal geometry.

2.5 Fractal Geometry

The word fractal was coined by B B Mandelbrot in his first essay on the topic (1975). The word is based on the Latin adjective fractus, and its corresponding verb frangere meaning to break. Frangere has the additional meaning to create irregular fragments. The word fractal suggests both breaking up and irregularity.

A fractal set generally has the following properties:-
(a qualitative description of such a set is given in section 2.5 ii)

- 1) Infinite structure, either regular or random, which results in the set being non-rectifiable and non-differentiable.
- 2) Scaling i.e. invariance under certain transformations of scale.
- 3) Self-similarity i.e. invariance under displacement.
- 4) Non-integral Hausdorff-Besicovitch dimension (section 2.5 i).
- 5) Hyperbolic probability distribution (section 2.5 iii).

Already it is clear that fractal sets are likely to be related with the phase screens previously described. For example the phase screens have infinite structure and self similarity. The importance of the Hausdorff-Besicovitch dimension must be explained.

2.5 i) The importance of dimension (number of dimensions)
in fractal geometry

A period known as the Crisis in Mathematics arose when a class of function was produced which was continuous and non-differentiable. The first of these was the Weierstrass function, others included the curves of Peano and Koch. The functions could not be explained and became known as pathological, or monsters. In 1925, a change in the interpretation of dimensionality ended the crisis. Four types of dimension will now be considered.

a) Topological dimension

Topology is the branch of mathematics which studies form. Topologically, any polygon outline has the same form, that of a circle. Thus the topological dimension for any function of this type, for example an island coastline, is 1. This is the topological dimension D_T . D_T is always an integer.

b) Fractal, or Hausdorff-Besicovitch, dimension

In contrast the Hausdorff-Besicovitch dimension, D , need not be an integer. D results from work by Carathéodory in 1924 and Hausdorff in 1919.

Classically, in order to evaluate the area of a planar shape, the area is approximated to many small squares, and the area found by summing the squares of the side lengths. Carathéodory extended this approach, and used spheres instead of squares in order to avoid the use of a known dimension. In general, a d -dimensional shape requires the addition of expressions of the form (Mandelbrot 1982)

$$h(\rho) = \gamma(d)\rho^d \quad (2.12)$$

$$\text{where } \gamma(d) = \frac{\Gamma[(1/2)]^d}{\Gamma(1 + d/2)} \quad (2.13)$$

and is equal to the volume of a sphere of unit radius.

Hausdorff extended this by allowing d to be fractional. In this case, $h(\rho)$ may be any test function which tends to zero with ρ . Equation (2.12) is then called the h -measure, and it is d -dimensional.

Besicovitch showed that, when d is not integral, for every set representing a d -dimensional shape there exists a real value D . If $d < D$ then the h -measure is infinite, if $d > D$ the h measure is zero. D is therefore the critical dimension of the set.

The two dimensions D_T and D must satisfy the Szpilrajn inequality (Hurewicz and Wallman 1941). Note that a fractal is by definition a set for which the Hausdorff-Besicovitch dimension strictly exceeds the topological dimension

$$D > D_T \quad (2.14)$$

c) Effective dimension

The effective dimension is not a quantitative measure, but is included because it helps give an intuitive feel for the concept of the fractal dimension. In addition, the effective dimension has links with other ideas of modern Physics, as it suggests that the numerical dimension depends on the relation of the object to the observer.

An example commonly used is that of a ball of string. Observed from very far away it appears as a point i.e. zero dimensional. As closer observation is made the object appears in turn as a three dimensional ball; a collection of one dimensional threads; and then reassumes three dimensionality as each thread becomes a column. This sequence is repeated on further magnification, until zero dimensionality is attained by the point-like appearance of an atom. Many objects and sets may be considered in this way. The transitions between zones with well defined dimension can be considered fractal, with $D > D_T$.

d) Similarity dimension

The similarity dimension is not applicable to random fractal sets, with which we are most concerned. It is included for completeness, as Hausdorff showed that in many cases it is equal to the fractal dimension.

Consider some non-fractal self-similar sets such as a straight line, or a unit square. A straight line of unit length may be divided into $N=b$ subintervals, each of length $r=1/b$. Similarly, a unit square can be divided into $N=b^2$ squares, each of side $r=1/b$. The similarity dimension is defined as

$$D_{SIM} = \log N / \log 1/r \quad (2.15)$$

For the examples given the value reduces to the topological dimension. The value is non-integral however for geometrically constructed self-similar fractal sets such as the Koch curve (Mandelbrot 1982).

2.5 ii) The coastline example of a fractal set

This example has been chosen to demonstrate the features of fractal geometry for two reasons. Firstly it is a statistical rather than a geometrically constructed example, and so ties in more realistically with the random functions under consideration. Secondly, there is a historical link with L F Richardson, who has already been mentioned in connection with turbulence theory.

A typical coastline exhibits the properties of self similarity and scaling. The statistical form of a coastline is the same whatever the scale of the map chosen to take measurements from. A stretch which appears smooth at low magnification will display irregularity on closer inspection. This property can also demonstrate the non-differentiability of fractal sets. If the curve always displays structure if the magnification is large enough, then where can a tangent be drawn to the curve?

A coastline is also non-rectifiable, this property is again linked to its infinite structure. Consider measuring a line or a circle with a set of dividers, open to a width ϵ . The length of the line or curve is given by the number of steps taken multiplied by ϵ . If ϵ is too large, then the value obtained for the circumference of the circle will be too small. The result will converge quickly, though, as ϵ is reduced. This is not the case for a coastline as finer detail is always resolved with reduction of ϵ .

Richardson (Mandelbrot 1982) found that the length of the common frontier between Spain and Portugal as measured by each country differed by 20%. Clearly the value chosen for ϵ was different in each case. Richardson derived an empirical relationship

$$L(\epsilon) = F \epsilon^{1-D} \quad (2.16)$$

where L is the total length, measured using an interval ϵ .

Mandelbrot (1982) re-interpreted these results and attached a significance to D which Richardson did not consider. He showed, using a geometrically constructed model for a coastline for which the similarity dimension could be evaluated, that if D is the fractal dimension, then

$$L(\epsilon) = \epsilon^{1-D} \quad (2.17)$$

describes the length exactly.

2.5 iii) Hyperbolic probability distribution

According to the central limit theorem the probability distribution of the sum of a large number of independent random variables is Gaussian. The variance of the random variables is assumed to be finite. In the case of a fractal set, where there is infinite structure, the variance is infinite (this is shown in section 2.6), and the central limit theorem does not predict a Gaussian distribution. A possible solution is hyperbolic, i.e.

$$P(h) \propto h^{-D} \quad (2.18)$$

Clearly this is not of great practical applicability, as, when $h=0$, assuming a finite variance, $P(h) \rightarrow \infty$. A Gaussian probability distribution is assumed when outer and inner scale sizes are present.

2.5 iv) Some fractal parameters

The link between the fractal dimension and the index of a power law spectrum is given by Berry (1979), in an argument following the original of Orey.

$$D = (5 - \alpha)/2 \quad (2.19)$$

where $\alpha = \nu + 1$, and is the index of the power law spectrum.

The case $D=1.5$ ($\alpha=2$) corresponds to the Brownian fractal, so named because the value arises for Brownian motion. $D=1$ is the marginal fractal; $D=2$ is the extreme fractal, the graph of the function almost fills a finite area.

For a spectrum given by

$$G(u) \propto u^{-\alpha} \quad (2.20)$$

the structure function is given by (Berry 1979)

$$D(X) = T^{3-\alpha} |X|_{\infty}^{\alpha-1} \quad (2.21)$$

This expression is derived from the power spectrum in a similar manner to that shown in the appendix for the non-fractal representation.

T is the topothesy, and is defined as the distance over which the chord drawn between two points on a fractal curve has an rms slope of one radian, i.e.

$$\left\langle \left[h(x+T) - h(x) \right]^2 \right\rangle / T^2 = 1 \quad (2.22)$$

It should be noted that in cases where an unmodified power law spectrum is proposed, the autocorrelation function is not defined and the structure function is used instead. This point is considered in the next section.

2.6 Use of the structure function instead of the autocorrelation function when working with power law spectra

The use of the unmodified power law spectrum, one which implies the absence of both inner and outer scales, has been mentioned in sections 2.4 and 2.5 iv). Uscinski et al (1981) show that the unmodified power law spectrum can only be used if the observer is very close to the screen, and irregularities the size of the outer scale and larger produce negligible fluctuations in the scattered amplitude. Such a restriction is unrealistic for far-field laser light scattering measurements, and even for ionospheric applications.

The unmodified power law spectrum leads to a very simple form of structure function (equation 2.2). It will now be shown that this function must be used rather than the autocorrelation function, since the latter does not exist for this unmodified form of spectrum.

Consider the unmodified power law spectrum

$$G(u) = A / |u|^2 \quad (2.23)$$

The corresponding structure function is given by (Tatarski 1961; appendix)

$$D(X) = 2A\pi |X| \quad (2.24)$$

Consider also a screen with an outer scale, whose spectrum is described by

$$G_o(u) = \frac{\Gamma}{\pi(1 + \Gamma^2 u^2)} \quad (2.25)$$

with the corresponding correlation function (Uscinski, Booker & Mariani 1981)

$$C_o(X) = \exp(-|X|/\Gamma) \quad (2.26)$$

Now, from equation (1.63)

$$D_o(X) = 2\sigma^2 \left[1 - C_o(X) \right] \quad (2.27)$$

so the corresponding structure function is

$$D_o(X) = 2\sigma^2 \left[1 - \exp(-|X|/\Gamma) \right] \quad (2.28)$$

The structure functions (2.24) and (2.28) for the unmodified and modified spectra respectively are compared in figure 2.1.

Now, for $X \ll \Gamma$

$$\exp(-|X|/\Gamma) = 1 - |X|/\Gamma$$

and

$$D_o(X) = 2\sigma^2 |X|/\Gamma \quad (2.29)$$

$$\text{If, } A\pi = \sigma^2 / \Gamma \quad (2.30)$$

then $D_o(X) = D(X)$ for small lags.

Now, as $\Gamma \rightarrow \infty$, from (2.28)

$$D_o(X) \rightarrow 2\sigma^2$$

But, in (2.24) $A\pi$ is finite, thus in (2.30) as $\Gamma \rightarrow \infty$ σ^2 must also be infinite. So, if an unmodified power spectrum is used, the variance of phase is infinite, and the second order statistic used must be the structure function.

In practical cases where there is always an outer scale the autocorrelation function can be used. It is calculated from the structure function using equation (1.63).

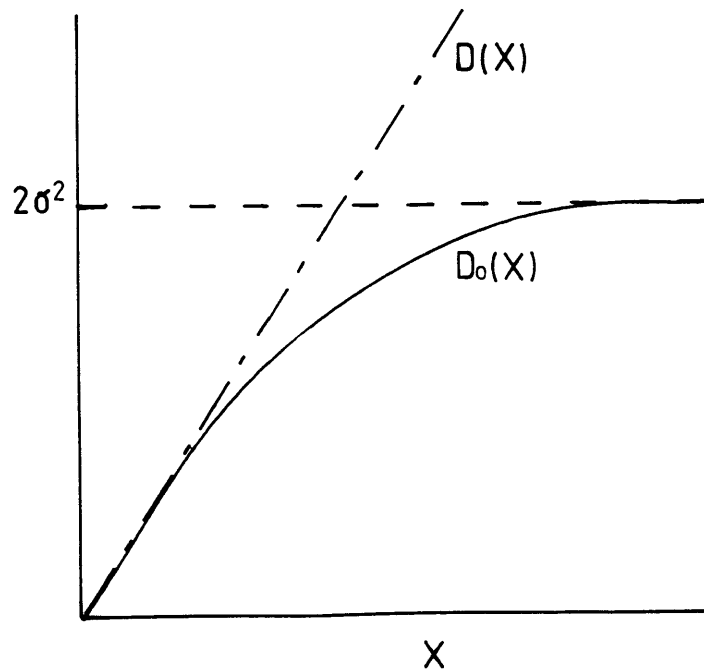


Figure 2.1. Comparison of structure functions for a modified (solid) and an unmodified (dashed) power law spectrum.

2.7 Example of phase screens with power law spectra

Examples are quite common in the large scale field of radio wave scattering (Tatarski 1961; Kerr 1972; Intrilligator & Wolfe 1970; Clifford et al 1971).

For the case of rough surfaces, Sayles & Thomas (1978) collated a vast amount of data about the surface profiles of natural and artificial surfaces, including a lunar sea, a ship hull plate, and a ring-lapped steel ball. In each case the power spectrum obeyed a power law over a short range of spatial frequency.

Light scattering measurements generally assumed a Gaussian autocorrelation function (and consequently a Gaussian spectrum) (Bennett & Porteus 1961; Beckmann & Spizzichini 1963; Beckmann 1967; Houchens & Hering 1967). The single scale, Gaussian autocorrelation function, model was found to give good agreement in certain cases. These include scattering by a liquid crystal (Pusey & Jakeman 1975), a thermal plume (Parry et al 1977), and a ruffled water surface

(Walker & Jakeman 1982). Chandley (1976) showed that the assumption of a Gaussian autocorrelation function was not valid in the case of ground glass. He found that the shape of the autocorrelation function of surface height measured with a stylus instrument was radically different from a Gaussian autocorrelation function. Chandley (1979) tried some alternative forms for the autocorrelation function, including negative exponential, and Lorentzian, but neither provided a good fit.

Thus it is clear that large scale phenomena seem to be well described by power law spectra, and theory is available to predict the scattering by this sort of phase screen (Jakeman & McWhirter 1977; Berry 1979). However, the theory has not been tested in the laboratory using a sample with known statistics. The first surface produced to test the theory was described by Jordan et al (1983). It was produced by randomly engraving a piece of germanium. Measurements were made using both $10.6\mu\text{m}$ radiation and stylus methods. The surface was shown to have an inverse square law power spectrum over several decades. An advantage of using radiation of wavelength $10.6\mu\text{m}$ as the incident radiation is that, as mechanical profiling has a resolution of about $1\mu\text{m}$, scale sizes less than the wavelength of the illumination will be noted, and their effect on the scattering investigated. With visible radiation, no measurement can be made of scale sizes smaller than the wavelength using the Talysurf.

2.8 Work described in this thesis

It is proposed to produce a laboratory scale multiscale scatterer, which will be used to test the scattering theories of Jakeman (& McWhirter 1977) and Berry (1979), using helium neon illumination. The scatterer should have the following statistical properties

- 1) Gaussian probability density of surface heights.

Note that although this assumption is necessary in the development of scattering predictions (see chapter 7), including those of Berry (1979), Mandelbrot defines a fractal as having a hyperbolic distribution (see section 2.5 iii). Jakeman (1981) also discusses how a Gaussian process is in general non-scaling.

2) Inverse power law spectrum of surface heights, of index -2 , over a range of spatial frequencies covering at least two decades. Such a spectrum will be associated with a truncated linear autocorrelation function.

The index -2 has been chosen because it simplifies the prediction of the scattered mean intensity with angle (see chapter seven), and corresponds to the Brownian fractal (section 2.5 iv).

In addition, in order to be of practical use the scale sizes, and the standard deviation of surface height should be appropriate to scattering of helium neon light.

The methods chosen to produce the surfaces are described in chapters 3 and 5. Mechanical measurements made on the surfaces are presented in chapter 6. The scattering theories mentioned briefly in this chapter are considered in more detail in chapter 7, followed by scattering measurements.

CHAPTER THREE

COMMENTS ON ROUGH SURFACE FABRICATION

In this chapter reasons for producing a multiscale random rough surface are presented. This is followed by a review of techniques which have been used previously to produce surfaces (mainly diffusers) with specified characteristics. The use of photoresist emerges as being the most promising method; its advantages over two existing techniques of multiscale surface production are suggested. The chapter is completed by a review of the problems encountered using photoresist in practice, and the ways that these difficulties might be overcome.

3.1 Reasons for surface fabrication

It has been shown in section 2.7 that many naturally occurring surfaces have an inverse power law spectrum. The question arises: Why make a special surface, when there may be one with the correct properties already in existence? The answer to this question is five fold. By making a surface one can ensure that

Point 1

The index of the inverse power law is exactly that required. For example $\alpha=2$ considerably simplifies the arithmetic expounded in section 7.3.

Point 2

The range of spatial frequencies over which the inverse power law is obeyed is known, without the need to make measurements.

Point 3

The Gaussian probability density required can be ensured.

Point 4

The scale sizes (i.e. spatial frequencies) and the depth of roughness can be chosen so that they are appropriate to the wavelength of the scattering light. There are two aspects to this

i) Although a landscape may have an inverse power law spectrum, the actual scale sizes mean that scattering experiments would require radar. The investigations would be impractically large. Reducing the scale in both horizontal and vertical directions allows scattering experiments using helium-neon laser light to be done conveniently in the laboratory.

ii) If radiation of a slightly longer wavelength, for example 10.6 microns is used on such surfaces the effects of features of scale lengths less than the wavelength of the scattered radiation can be investigated.

Point 5

A fabricated surface can be produced such that it is of use in both reflection and transmission. Clearly many natural surfaces are not.

3.2 Previous techniques used

The idea of fabricating rough surfaces is not new. Much work has been done in the area of design and synthesis of random phase diffusers. The requirement in such cases is a diffuser whose light scattering profile is controlled. For example, consider a diffuser used as a viewing screen. A 'hot spot' caused by the decrease in scattered light in directions away from the specular direction usually occurs. It could be eliminated by using a diffuser with a uniform power spectrum across the aperture (Chandler and DePalma 1968). Alternatively, such a diffuser would be useful in Fourier transform holography. Such holograms, recorded using a diffuser to illuminate the object (figure 3.1), have high redundancy. This property means that the hologram may be scratched, dirty or even broken into pieces, and there is no

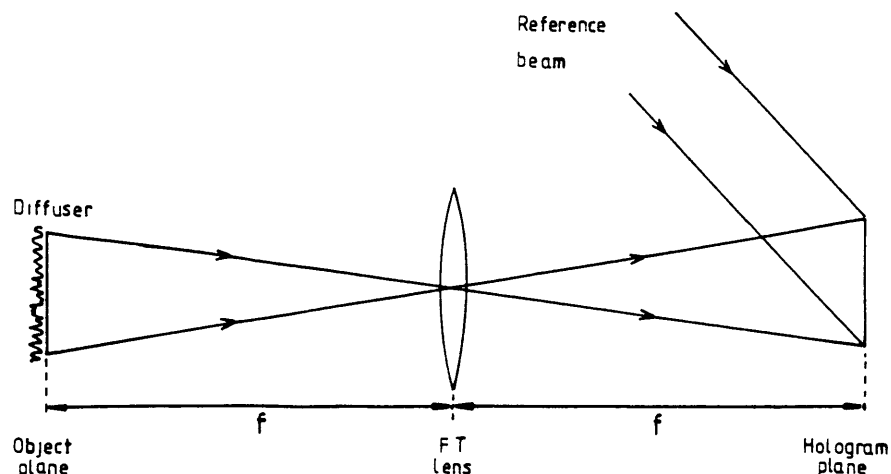


Figure 3.1. The use of a diffuser in Fourier transform holography.

great loss of information. This occurs because the diffuser is considered to spread the information over the whole hologram. Using a band limited diffuser, with a uniform power spectrum within this limit could maximise both redundancy and the signal to noise ratio (Iwamoto 1980; Kurtz, Hoadley and DePalma 1973).

Other applications include coherent imaging with reduced speckle (Lahart and Marathay 1975) and reflecting filters for infra red (Porteus 1963; Bennett 1963).

In all these cases conventional diffusers had been rejected. Such random phase diffusers are produced by etching or grinding. In the former case the time, temperature or chemical used may be altered, but these changes cannot be used to produce a diffuser with a desired property. The type of abrasive used in grinding has little effect on the resulting scattering profile (Kurtz 1972; Chandley 1979). So, alternative methods were developed, in particular photographic techniques (Waag and Knox 1972). Gelatin plates were exposed to speckle and then bleached to produce a relief. The speckle pattern used may have been spatially filtered in order to affect the scattering characteristics of the diffuser (von Bestenreiner

and Weiershausen 1971)). However, the grain of photographic emulsions introduced unwanted noise into light scattered by a diffuser. Thermoplastic films (Lee 1974) tend to exhibit 'frosting' which is a quasi-random structure containing a preferred spatial frequency.

Gray (1978) overcame these problems by using photoresist, which is a high resolution material free from noise. Gray produced random rough surfaces with both a Gaussian autocorrelation function of surface height, and a Gaussian probability density of surface height. The technique has since been used successfully by Levine and Dainty (1983) who were able to produce diffusers with known correlation lengths. The correlation lengths of the diffusers were all in the region of 80 microns, with the standard deviation of surface height about 1 micron. Such diffusers would be described by the single scale model of section 2.1.

Clearly, some refinement of the technique is required to produce multiscale surfaces, and this will be discussed in chapter four. Before returning to the use of photoresist let us look briefly at other methods of producing multiscale random rough surfaces. The first (Jordan, Hollins and Jakeman 1983) was a relatively uncontrolled technique. Their phase screen was produced by randomly scratching the surface of a piece of germanium using an engraving tool. The surface so produced could, however, be described as fractal and was the first laboratory scale phase screen made. The second method (Hollins and Jordan 1983) was more controlled. This surface was made to correspond with theoretical scattering predictions made by Jakeman and Hoenders (1982). The surface consisted of an arrangement of randomly spaced rectangular grooves, and thus had only two possible height levels. Ion-beam milling was used, the grooves produced were mostly in the range 20 to 30 microns wide to maximise the scattering of 10 micron CO_2 laser radiation. The first method thus does not satisfy points 1) and 2) of section 3.1. The second method is not sufficiently general. It was therefore decided that a technique using photoresist to record specially manipulated speckle patterns would be the best way of satisfying points 1 to 5. The use of speckle is discussed in chapter four. The remainder of this chapter is devoted to the use of photoresist in practice.

3.3 The coating and exposure of photoresist

Preparing photoresist plates for exposure should have been one of the more minor features of this work. However, it transpired that applying thick layers of photoresist to glass plates is an art, which must be learned by each user, and practiced until he finds a method which works for him. This should be borne in mind by any reader hoping that information in this section will provide instant perfect results.

The technique used is basically that of MacAndrew (1980) with a few modifications. All thicknesses quoted were measured using the Linnik interference microscope.

3.3 i . Preparation of plates

Two techniques were used successfully, they are summarised in Table 3.1.

3.3ii Application of photoresist to plates

MacAndrew (1980) lists five methods of applying photoresist to surfaces. These are:

- spinning
- flowing
- dipping
- pouring
- doctor blade or roller.

Flow coating and spinning seemed to be the most likely to result in smooth surfaces up to about ten microns thick, and these techniques were tried. The resist used for the bulk of the work is Shipley Microposit 1375, which is the most viscous resist produced by Shipley.

TABLE 3.1

Method 1

Method 2

1. Wash in tap water and washing up liquid at 50°C. Rinse thoroughly in running water at same temperature.

2. Wet thoroughly with acetone $\{(\text{CH}_3)_2\text{CO}\}$. Rub with lintless lens cleaning tissue to remove greasy marks and finger prints.

3. Working in dust free atmosphere of laminar flow cabinet apply first of following solvents to plate, lay lens tissue across surface to absorb solvent, then remove it by pulling it across in plane of surface. Repeat using next solvent.

- a) Acetone
- b) Methanol $\{\text{CH}_3\text{OH}\}$
- c) Ethanol $\{\text{C}_2\text{H}_5\text{OH}\}$
- d) Isopropanol $\{(\text{CH}_3)_2\text{CHOH}\}$

3. Clean in isopropanol $\{(\text{CH}_3)_2\text{CHOH}\}$ in ultrasonic tank. Remove, and stand almost vertically against upright support to dry in laminar flow cabinet.

4. Spray dry canned air across surface immediately before coating.

Shipley 1375 Microposit arrives from the manufacturers doubly sealed from clean room conditions. It should contain no impurities greater than .2 microns in size. However, to ensure that it is still pure when it is deposited on the surface it is fed from a glass syringe through a .5 micron filter. The filter is paper, and is held in place by a metal attachment fixed to the syringe by a luer lock.

Absolute cleanliness of all the tools used with photoresist is essential. The metal filter assembly and the syringes are cleaned before use with acetone followed by the Shipley photoresist thinner (Microposit thinner; ethyl glycol acetate and xylene). After use the syringe and filter are dismantled and soaked in a bath of thinner for several minutes before a final wash with acetone. This cleaning procedure revealed a problem with the first type of syringe used. These were made of plastic, and the snug fit of the plunger in the barrel was ensured by the provision of a rubber ring around the tip of the plunger. The repeated use of solvents caused the rubber ring to stretch, and the plunger would become stuck in the barrel. It was also likely that the rubber contaminated the resist. Such syringes were replaced by some made entirely of glass. Several of these are required for a single coating session, as the plungers tend to become firmly stuck in the barrel and require the application of solvents and brute force to free them. Coating is done in the dust free environment of a laminar flow cabinet.

3.3 ii a) Flow coating

Cleaned glass plates are stood in a large funnel at an angle of about 30° to the vertical. The photoresist is applied in a continuous broad sweep from one lower corner, around the top, to the other lower corner. The plates are left in the funnel for ten minutes, then allowed to dry at the same inclination for three hours.

Once the difficulty of achieving an even continuous flow from the syringe had been overcome, a further drawback arose. Smooth coatings could be obtained using a dilution of 2 parts resist to 1 part thinner. Such layers were about three microns thick. If the photoresist was used undiluted, or at any proportion more viscous

than 2:1, then the coating would be covered in ridges, running parallel to the direction of flow. These ridges were apparent before exposure and development, and could not be removed by baking. An example is shown in plate 3.1. Similar features occurred if the inclination of the plate to the vertical was increased in an attempt to produce thicker coatings.

3.3 ii b) Spin coating

Spin coating is the method used for applications in the electronics industry. In such cases the requirement is for a very thin layer, typically around one micron thick. Spin speeds of up to 5000 r.p.m. are recommended (Shipley). Shipley also suggest that the resist is applied before the plate is spun, and that the spin times should be longer than twenty seconds. Experimentation showed that the most suitable plates for my application were obtained by ignoring these recommendations.

The spinner used in the earlier work is the one built by MacAndrew (1980) and consists of a d.c. motor with its spindle held vertically. A mount for the plate to be coated is attached to the spindle. The motor is enclosed separately from the coating area. The coating area itself is shielded to prevent the escape of resist, but is open at the top to allow access and evaporation. A cleaned plate, attached to the mount using double sided tape, is spun at the required speed. Resist is applied to the spinning plate using a 2ml syringe and the filter arrangement described in section 3.3 ii. Plates are allowed to dry for about one hour lying flat in the laminar flow cabinet. After a further 24 hours drying flat in sealed light tight boxes, the plates may be baked for 30 minutes at 90°C. The rear surfaces of the plates are then painted with matt black paint to prevent unwanted fringes appearing in the resist, on exposure, from multiple reflections.

Initially the spinner was giving speeds of about 2400 to 3800 r.p.m. (determined using stroboscope). Plates coated at these speeds always displayed striations. These have been investigated at some length by MacAndrew (1980; Gale and Knop 1980), so a very brief

Plate 3.1. Linnik interferogram of an example of
ridging on a flow coated plate.

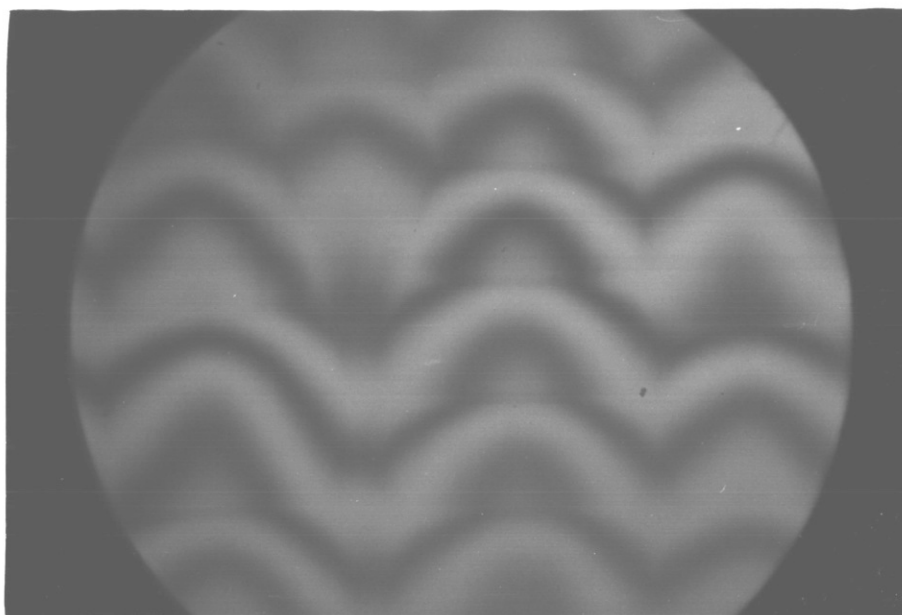


PLATE 3.1

description will be given here. Striations are radial furrows (typically .2 microns deep and 0.1mm wide) which are only apparent after development. They occur even when no exposure has taken place. Shipley recognises the problem and manufactures a striation-free photoresist. However, a major contribution to this property of the product would appear to be its low viscosity. Since this results in thinner coatings the new style resist was not used. The occurrence of striations was reduced when slower spin speeds were used, so the spinner was adjusted to spin in a range of speeds between 900 and 1830 r.p.m. It was found that at speeds in the approximate range 1000 to 1800 r.p.m. striations were absent when using Microposit 1375. Such coatings are about five microns thick. The spin time is typically ten seconds after all the resist has been applied. Latterly, a new spinner built by P M Channon has been used successfully at similar speeds.

Attempts have been made to obtain much thicker overall coatings by layering. The previous layer is baked to ensure complete evaporation, and to attempt to reduce 'melting' by the solvent in the newly applied resist. Dust is a particular problem, as plates coated with one or more layers have been transferred from box to oven and back outside the laminar flow cabinet. A blast from a can of air is usually sufficient to reduce the effects. Although other authors (Levine and Dainty 1983) have had success in coating multiple layers, I have not. The resultant coatings have tended to be uniform over small, about 1cm diameter, regions of the plate. Overall the appearance is uneven; run marks with thickened edges and thin centres are a typical imperfection.

3.3 iii Exposure of plates

Shipley Microposit 1375 increases in sensitivity with decrease in wavelength below about 460nm, peaking in the UV (363.8 nm). Argon-ion lasers have an output line at 457.9 nm. The output power is less at this wavelength than in the strong 488 nm line, but the photoresist is significantly more sensitive to it. The 457.9 nm line has been used throughout. The minimum exposure time for an approximate power density of .03 mw per square centimetre was five minutes.

3.3 iv Development

Plates were developed in a solution of one part Microposit 303 developer to six parts distilled water, at 20°C for 20 seconds. This dilution was suggested by Norman and Singh (1975) as suitable for obtaining a linear response, and has subsequently been used by MacAndrew (1980).

The developed plates were then rinsed in distilled water for about ten seconds to stop the developer action. They were then rinsed in a strong spray of tap water for about two minutes, to remove all the loosened resist. It was also found that the spray of water was a good way of removing the black paint on the rear surfaces of the plates. After a final rinse in distilled water, the plates were left standing vertically to dry naturally.

CHAPTER FOUR

SOME STATISTICAL PROPERTIES OF LASER SPECKLE PATTERNS

In this chapter a brief description of the various types of speckle is given. Interest is then focused on fully developed speckle patterns; some of their basic first and second order statistical properties are listed. Two alternative methods of obtaining a Gaussian probability density function in a composite speckle pattern are described, and the corresponding power spectrum is investigated in each case.

4.1 Types of speckle pattern

Laser speckle is the random, grainy pattern which results from the scattering of coherent illumination. The properties of fully developed coherent speckle are well known, and are set out in several reviews (Dainty 1976; Goodman 1984).

A fully developed speckle pattern is one which arises from the interference between coherent light scattered from a large number of scatterers which introduce path differences greater than one wavelength, such that phases are distributed uniformly from $-\pi$ to π . In such a case the distribution of complex amplitude in the pattern is Gaussian. This property arises from the application of the central limit theorem (Papoulis 1985 p194). The properties of the surface itself will have no effect on the fluctuations of the pattern about the mean intensity.

Non-Gaussian speckle arises when the above conditions are not satisfied. That is, when only a small number of scatterers are illuminated; the scatterer is weak (path differences are less than one wavelength); or the illumination is not coherent.

4.2 Some useful statistical properties of a single fully developed speckle pattern

Results are presented here, in the main, without derivation. Fuller treatments may be found in Goodman (1984) and Dainty (1976).
• The speckle under consideration is formed by perfectly monochromatic and polarized waves.

4.2 i) First order statistics

4.2 i) a) Complex Amplitude

If the real and imaginary parts of the complex amplitude of a speckle field are given by A_R and A_I , the following equations apply

$$\langle A_R \rangle = \langle A_I \rangle = 0 \quad (4.1)$$

$$\langle A_R^2 \rangle = \langle A_I^2 \rangle = \sigma^2 \quad (4.2)$$

where σ^2 is the variance of each part of the field.

$$\langle A_R A_I \rangle = 0 \quad (4.3)$$

The joint probability density function for the real (A_R) and imaginary (A_I) parts of the complex amplitude of a speckle field is given by

$$p(A_R, A_I) = \frac{1}{2\pi\sigma^2} \exp \left[- \frac{(A_R^2 + A_I^2)}{2\sigma^2} \right] \quad (4.4)$$

where σ^2 is the variance of both real and imaginary parts of the field.

This density function is known as a circular Gaussian density function. It applies when

- i) the phases of the scattered field are uniformly distributed between $-\pi$ and π
- ii) there is a large number of scatterers
- iii) the elementary scattering areas are unrelated to one another
- iv) the strength of a scattered component is not related to its phase.

Note also,

$$A = A_R + iA_I \quad (4.5)$$

so,

$$\langle A^2 \rangle = \langle A_R^2 \rangle + 2i\langle A_I A_R \rangle - \langle A_I^2 \rangle \quad (4.6)$$

but, from equation (4.2)

$$\langle A_R^2 \rangle = \langle A_I^2 \rangle$$

thus

$$\langle A^2 \rangle = 2i\langle A_I \rangle \langle A_R \rangle \quad (4.7)$$

but, $\langle A_I \rangle = \langle A_R \rangle = 0$ (equation 4.1)

so,

$$\langle A^2 \rangle = 0 \quad (4.8)$$

similarly,

$$\langle A^* A^* \rangle = 0 \quad (4.9)$$

4.2 i) b) Phase

The probability density function of phase is uniform

$$p(\phi) = \begin{cases} 1/2\pi & -\pi \leq \phi \leq \pi \\ 0 & \text{otherwise} \end{cases} \quad (4.10)$$

4.2 i) c) Intensity

The probability density function of intensity is of negative exponential form

$$p(I) = \begin{cases} \frac{1}{2\sigma^2} \exp(-I/2\sigma^2) & I \geq 0 \\ 0 & \text{otherwise} \end{cases} \quad (4.11)$$

The n'th moment $\langle I^n \rangle$ of intensity is

$$\langle I^n \rangle = n! (2\sigma^2)^n \quad (4.12)$$

When $n = 1$

$$\langle I \rangle = 2\sigma^2 \quad (4.13)$$

therefore

$$\langle I^n \rangle = n! \langle I \rangle^n \quad (4.14)$$

From equation (4.14)

$$\langle I^2 \rangle = 2 \langle I \rangle^2 \quad (4.15)$$

Now,

$$\sigma^2 = \langle I^2 \rangle - \langle I \rangle^2 \quad (4.15)$$

so

$$\sigma^2 = \langle I \rangle^2 \quad (4.16)$$

Using (4.13), equation (4.10) may be written

$$p(I) = \begin{cases} \frac{1}{\langle I \rangle} \exp(-I/\langle I \rangle) & I \geq 0 \\ 0 & \text{otherwise} \end{cases} \quad (4.17)$$

The above expressions apply to far field speckle (see figure 4.1). They may also be applied to patterns formed in the image plane of a rough diffuser (figure 4.2) providing that the area of the point spread function of the system contains a large number of scatterers. This condition is usually satisfied for numerical apertures less than 0.1.

4.2 i) d) Speckle size

The speckle size can be visualised as the average distance between adjacent regions of maximum and minimum brightness. This is inversely related to the effective diameter of the scatterer, being given by the approximate diameter of the first dark ring of the Airy pattern for the scattering area. For a circular diffuser the expression (4.23) for the autocorrelation function (section 4.2 ii) is the Airy disc formula. So the width of the autocorrelation function is also a useful approximation for the speckle size. For objective speckle (figure 4.1)

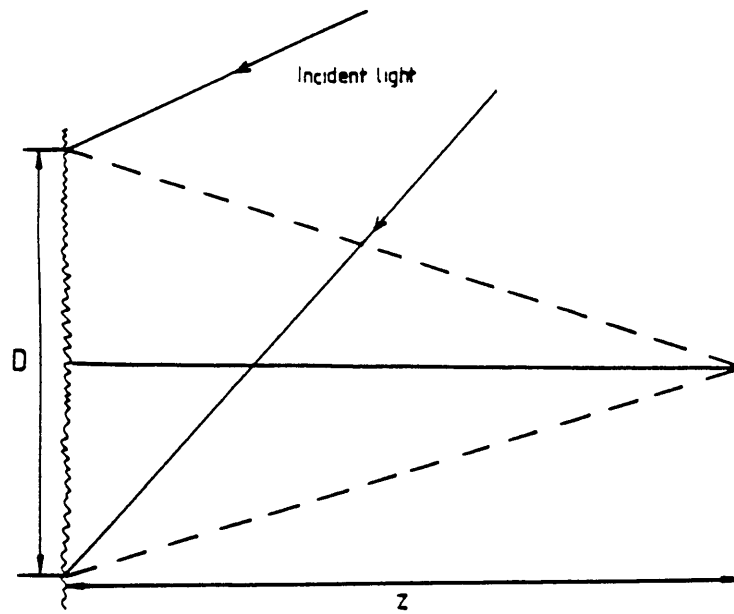


Figure 4.1. Formation of objective speckle.

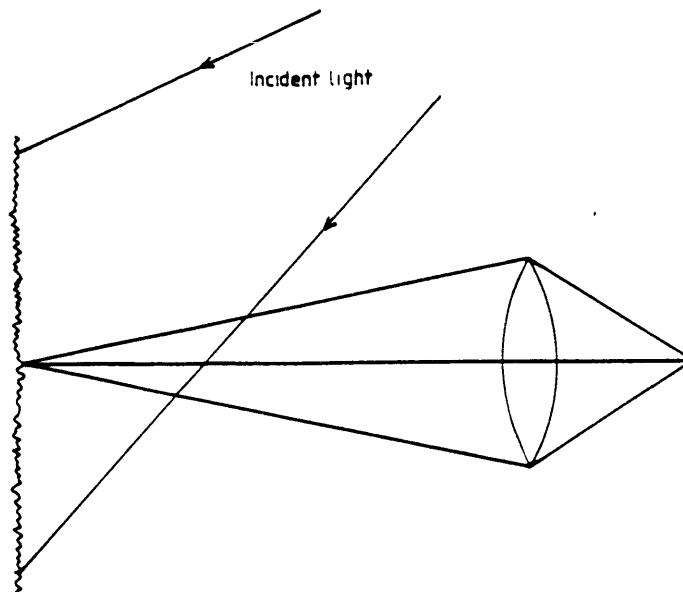


Figure 4.2. Formation of subjective speckle.

$$\Sigma \sim 1.2\lambda z/D \quad (4.18)$$

where D is the diameter of an illuminated circular region, λ is the illuminating wavelength, and z the distance to the speckle plane. For subjective speckle (figure 4.2) i.e. that formed by an imaging system

$$\Sigma \sim 0.6\lambda/N.A. \quad (4.19)$$

where $N.A.$ is the effective numerical aperture of the lens, and λ the illuminating wavelength. This may be written

$$\Sigma \sim 1.2\lambda f/no. \quad (4.20)$$

4.2 i) e) Contrast

$$C = \sigma_I / \langle I \rangle = 1 \quad (4.21)$$

4.2 ii) Second order statistics

The autocorrelation function of intensity for circular, complex Gaussian fields may be written (Goodman 1984 p36)

$$\langle I_1 I_2 \rangle = \langle I \rangle^2 + |\langle A_1 A_2^* \rangle|^2 \quad (4.22)$$

Now, assuming that the microstructure of the scattering surface is too fine to be resolved by a lens in the observation plane the size of the observation region, the autocorrelation function of intensity in the Fraunhofer plane, for the geometry shown in figure 4.3 is given by (Goodman 1984 p38)

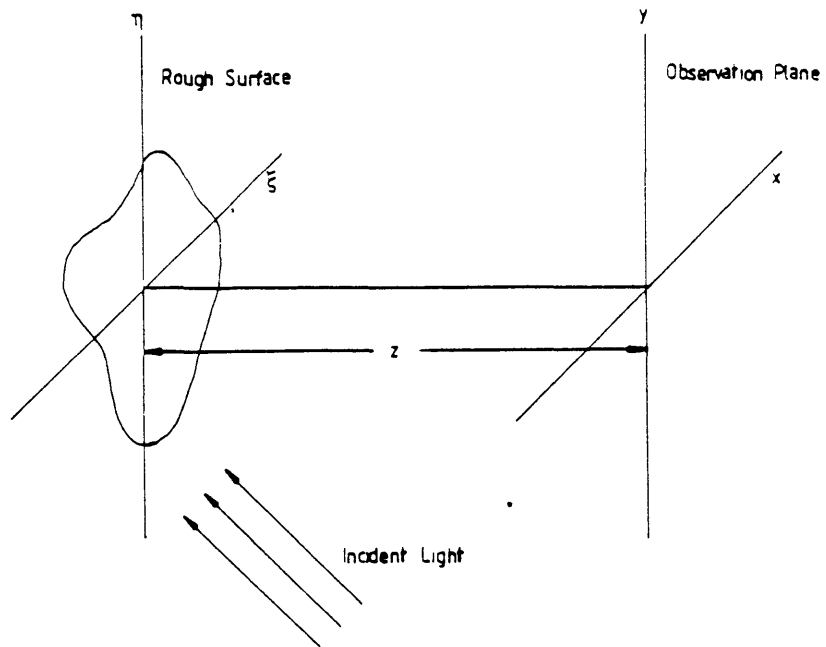


Figure 4.3. Co-ordinate system used for free space geometry.

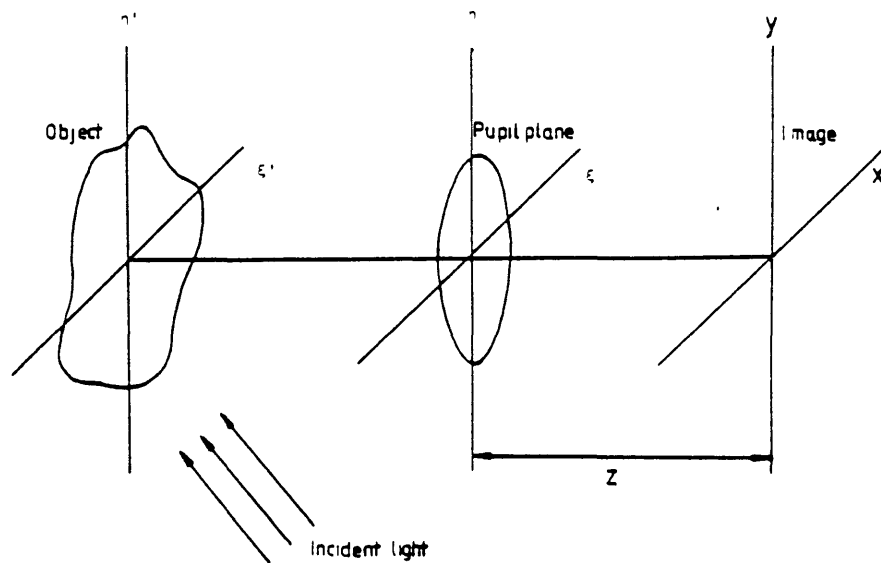


Figure 4.4. Co-ordinate system used for imaging geometry.

$$C(\Delta x, \Delta y) =$$

$$\langle I \rangle^2 \left[1 + \frac{\iint_{-\infty}^{\infty} |P(\xi, \eta)|^2 \exp \left[i \frac{2\pi}{\lambda z} (\xi \Delta x + \eta \Delta y) \right] d\xi d\eta}{\iint_{-\infty}^{\infty} |P(\xi, \eta)|^2 d\xi d\eta} \right]^2 \quad (4.23)$$

where $|P(\xi, \eta)|^2$ is the intensity distribution incident on the scattering area, and λ is the wavelength.

The power spectrum of the intensity distribution in the Fraunhofer plane is given by the Fourier transform of the autocorrelation function (section 1.7 i)

$$G(u, v) = \langle I \rangle^2 \left\{ \delta(u, v) + \frac{\iint_{-\infty}^{\infty} |P(\xi, \eta)|^2 |P(\xi - \lambda z u, \eta - \lambda z v)|^2 d\xi d\eta}{\left[\iint_{-\infty}^{\infty} |P(\xi, \eta)|^2 d\xi d\eta \right]^2} \right\} \quad (4.24)$$

The expression (4.24) can be interpreted as follows. The power spectrum consists of a δ function at zero frequency, together with a second component which is extended over frequency. This second component has the shape of the autocorrelation function of $|P(\xi, \eta)|^2$ i.e. of the intensity distribution incident on the scattering area.

When an imaging system is present the results (4.23) and (4.24) may still be applied, but $|P(\xi, \eta)|^2$ should be interpreted as the intensity transmittance of the lens pupil. In other words, the pupil plane is treated as a uniformly bright rough surface, followed by free space propagation to the image plane. The new coordinate system is shown in figure 4.4. Aberrations in the imaging system affect only the phase of $P(\xi, \eta)$, so they will have no effect on the autocorrelation function, which is dependent only on $|P(\xi, \eta)|^2$.

We can now see that the intensity distribution in the speckle pattern has potential for use as a random function whose statistical properties correspond to those of a required model. Let us consider first the probability density function. For a single pattern, as shown in equation (4.11) above, the density has a negative exponential form. In the following sections two ways will be shown of combining speckle fields to obtain a close approximation to a Gaussian probability density function. In each case the contrast of the resulting pattern will also be derived, as a very low contrast result would be of little practical use.

4.3 Speckle pattern resulting from the incoherent addition of a number of fully developed speckle patterns

4.3 i) Probability density function

The probability density function of total intensity resulting from the addition of N uncorrelated speckle patterns, each of mean intensity λ_k is given by (Goodman 1984 p23)

$$p_I(I) = \sum_{k=1}^N \frac{\lambda_k^{N-2}}{\prod_{\substack{p=1 \\ p \neq k}}^N (\lambda_k - \lambda_p)} \exp \left[-\frac{I}{\lambda_k} \right] \quad (4.25)$$

for $I \geq 0$, zero for negative I .

Clearly this expression depends on the values of N and λ_k . In the special case when λ_k are all equal, the equivalent expression is

$$p_I(I) = \frac{I^{N-1} N^N}{(N-1)!} \exp(-NI) \quad (4.26)$$

for $I \geq 0$, zero for negative I .

Figure 4.5 shows curves for various N , for the special case when λ_k are all equal. Figure 4.6 shows the curve for $N=10$ compared with a Gaussian curve with the same mean and standard deviation. In chapter five, figures 5.6 and 5.7 correspond to these but the λ_k are not all equal.

4.3 ii) Contrast

The contrast of any fully developed speckle pattern, whether it results from a summation or not, is given by

$$C = \sigma / \langle I \rangle \quad (4.21)$$

To apply this expression to a sum of N incoherently added speckle patterns, it is necessary to find an expression for the standard deviation of the result. If the standard deviation of each pattern is σ_k , then (Papoulis 1985 p153)

$$\sigma^2 = \sum_{k=1}^N \sigma_k^2 \quad (4.27)$$

but, in a fully developed speckle pattern (section 4.2 i)

$$\sigma^2 = \langle I \rangle^2 \quad (4.16)$$

therefore

$$\sigma_k^2 = \lambda_k^2 \quad (4.28)$$

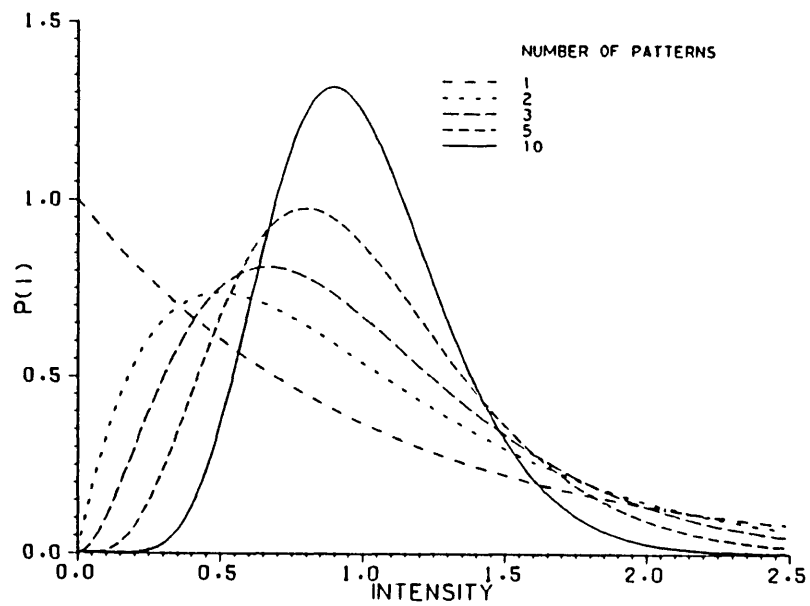


Figure 4.5. Probability density function for the incoherent sum of a number of fully developed speckle patterns, each of the same mean intensity.

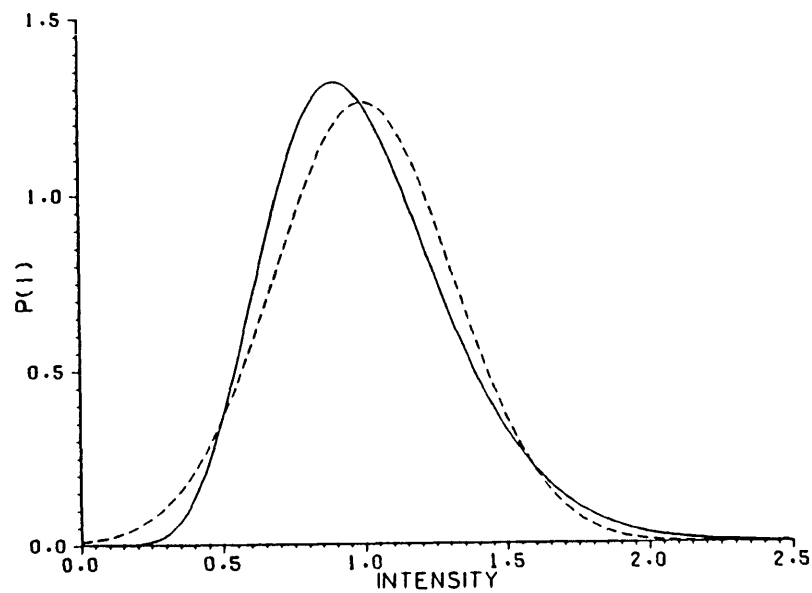


Figure 4.6. Probability density function (solid) for the incoherent sum of 10 fully developed speckle patterns each of the same mean intensity, compared with a Gaussian curve (dashed) of the same mean and standard deviation.

and

$$\sigma^2 = \sum_{k=1}^N \lambda_k^2 \quad (4.29)$$

Thus, the standard deviation of the incoherent sum of N speckle patterns, each of mean intensity λ_k , is

$$\sigma = \left[\sum_{k=1}^N \lambda_k^2 \right]^{\frac{1}{2}} \quad (4.30)$$

Curves showing the change in contrast with different numbers of summed patterns are shown in chapter five, figure 5.8, when the importance of the values taken by λ_k is emphasised.

4.4 Speckle pattern resulting from the addition of a fully developed speckle pattern and a coherent background

The assumptions in this section are

- i) The speckle pattern is perfectly polarized, the complex amplitude obeys circular Gaussian statistics and is independent of the coherent background.
- ii) The background is of constant intensity and has the same polarization as the speckle pattern.

4.4 i) Probability density function

The probability density function of intensity is given by
(Goodman 1984)

$$p_I(I) = \frac{1}{\langle I_s \rangle} \exp\left[\frac{-I+I_B}{\langle I_s \rangle}\right] I_0\left[\frac{2\sqrt{II_B}}{\langle I_s \rangle}\right] \quad (4.31)$$

for $I \geq 0$, zero for negative I .

where $\langle I_s \rangle$ = mean intensity of speckle pattern

I_B = background intensity

$I_0(\)$ = modified Bessel function of first kind, zero order (Watson 1962 chap 3)

This density function depends on the beam ratio

$$r = I_B / \langle I_s \rangle \quad (4.32)$$

and

$$\langle I_s \rangle + I_B = 1 \quad (4.33)$$

So, (4.31) may be rewritten (Dainty 1972a)

$$p_I(I) = (r+1) \exp\left[-r - I(r+1)\right] I_0\left\{2\left[Ir(r+1)\right]^{1/2}\right\} \quad (4.34)$$

Figure 4.7 shows this expression for various values of the beam ratio. For $r=10$ the curve is close to Gaussian (figure 4.8).

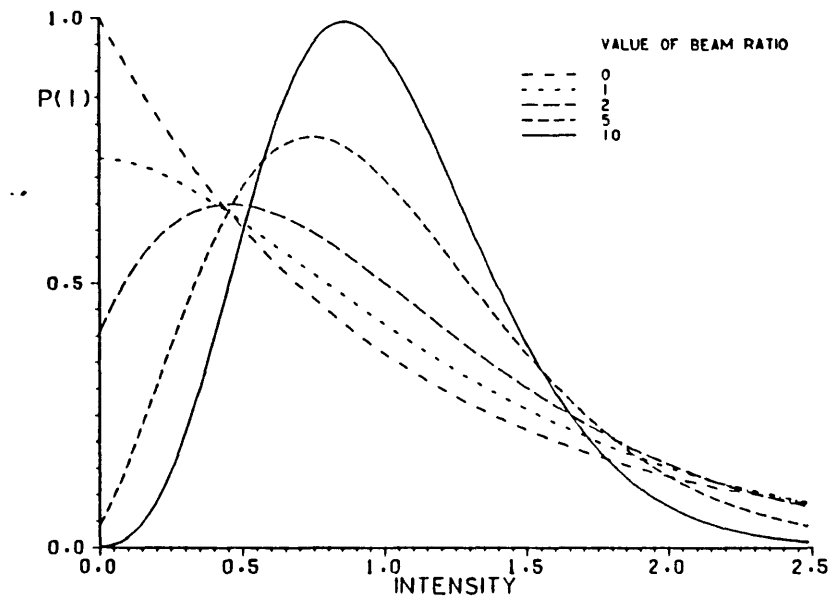


Figure 4.7. Probability density function for the coherent sum of a fully developed speckle pattern and a plane background wave, for several values of the beam ratio r (equation 4.32).

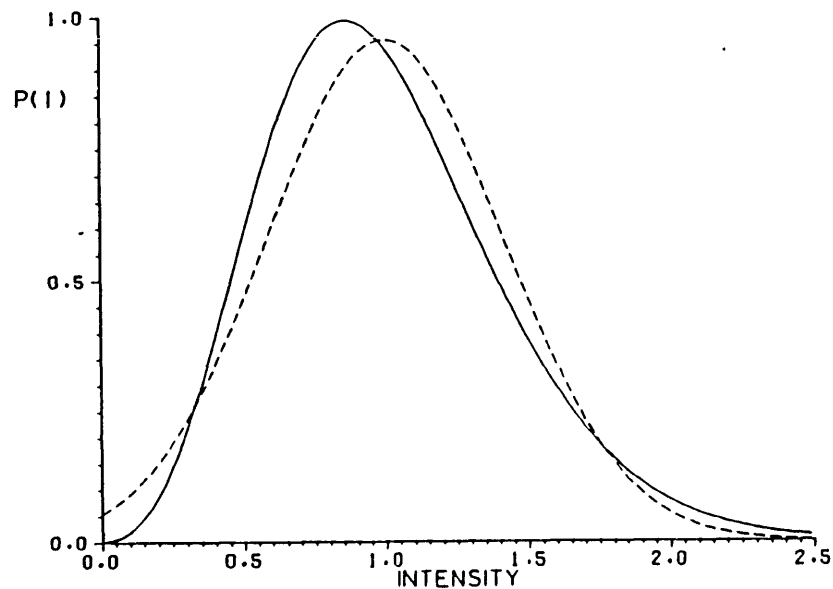


Figure 4.8. Probability density function (solid) for the coherent sum of a fully developed speckle pattern and a plane background wave of beam ratio $r=10$, compared with a Gaussian curve (dashed) of the same mean and standard deviation.

4.4 ii) Contrast

The contrast of the resulting pattern is given by

$$C = (2r+1)^{1/2}/r+1 \quad (4.35)$$

We have now seen that it is possible to obtain a Gaussian probability density function in two ways. Let us now consider the power spectra resulting from such combinations.

4.5 Power spectrum from the incoherent addition of a number of fully developed speckle patterns

The total resulting power spectrum is given simply by the summation of the spectra of the individual patterns.

4.6 Power spectrum from the addition of a fully developed speckle pattern and a coherent background

The situation is less straightforward here. It will be shown that the value of the beam ratio r determines whether the overall spectrum is dominated by that of the speckle pattern alone, or of the combination. It is instructive here to consider the coherent addition from the addition of complex amplitudes onwards. Let the complex amplitude of the speckle at the plane of interest be S , and the complex amplitude of the background field be B . Then the total complex amplitude is given by

$$A = S + B \quad (4.36)$$

The autocorrelation function of the complex amplitudes is given by

$$\begin{aligned}
 \langle A_1 A_2^* \rangle &= \langle (S_1 + B_1) (S_2^* + B_2^*) \rangle \\
 &= \langle S_1 S_2^* \rangle + \langle S_1 \rangle B_2^* + \langle S_2^* \rangle B_1 + B_1 B_2^*
 \end{aligned} \tag{4.37}$$

but $\langle S \rangle = \langle S^* \rangle = 0$, therefore

$$\langle A_1 A_2^* \rangle = \langle S_1 S_2^* \rangle + B_1 B_2^* \tag{4.38}$$

Now, if B has a plane wavefront

$$B_1 B_2^* = I_B \tag{4.39}$$

and when points 1 and 2 coincide

(4.38) may be written

$$\langle I \rangle = \langle I_S \rangle + I_B \tag{4.40}$$

Now, consider the autocorrelation function of intensities

$$\langle I_1 I_2 \rangle = \langle A_1 A_1^* A_2 A_2^* \rangle \tag{4.41}$$

Now, all odd moments of the speckle complex amplitude are zero, equation (4.1). Also if both amplitudes multiplied together in a pair are (or ^{both} are not) complex conjugates then the result is zero, see section 4.2 i) a). Thus, on expansion (4.41) becomes

$$\langle I_1 I_2 \rangle = \langle I_{S1} I_{S2} + I_{S1} I_B + S_1^* S_2 I_B + I_B I_{S2} + S_1 S_2^* I_B + I_B^2 \rangle \tag{4.42}$$

$$= \langle I_{S1} I_{S2} \rangle + I_B (\langle I_{S1} \rangle + \langle S_1^* S_2 \rangle + \langle I_{S2} \rangle + \langle S_1 S_2^* \rangle) + I_B^2 \tag{4.43}$$

i.e.

$$\langle I_1 I_2 \rangle = \langle I_{S1} I_{S2} \rangle + I_B (2\langle I_S \rangle + 2\text{Re}\langle S_1 S_2^* \rangle) + I_B^2 \quad (4.44)$$

But, from equation (4.22)

$$\langle I_{S1} I_{S2} \rangle = \langle I_S \rangle^2 + |\langle S_1 S_2^* \rangle|^2 \quad (4.22)$$

Thus, the autocorrelation function of intensities in the combination is

$$\langle I_1 I_2 \rangle = \langle I_S \rangle^2 + |\langle S_1 S_2^* \rangle|^2 + I_B (2\langle I_S \rangle + 2\text{Re}\langle S_1 S_2^* \rangle) + I_B^2 \quad (4.45)$$

For clarity, instead of using the autocorrelation function to predict the power spectrum, the autocovariance will be used instead. This is the autocorrelation function with the mean square of the intensity subtracted (equation 1.44). The power spectrum is strictly defined as the Fourier transform of the autocorrelation function. The Fourier transform of a constant is a delta function, so if the autocovariance is transformed instead of the autocorrelation the result will omit this feature. The component of the power spectrum which is extended over frequency will, however, be unaffected.

$$\text{cov}(I_1 I_2) = \langle I_1 I_2 \rangle - \langle I \rangle^2 \quad (4.46)$$

and

$$\langle I \rangle = \langle I_S \rangle + I_B \quad (4.40)$$

so

$$\text{cov}(I_1 I_2) = |\langle S_1 S_2^* \rangle|^2 + 2\text{Re}\langle S_1 S_2^* \rangle I_B \quad (4.47)$$

The power spectrum is given by the Fourier transform of this expression, it will have two components

Component 1

$$F \left[|\langle S_1 S_2^* \rangle|^2 \right]$$

Component 2

$$F \left[2I_B \langle S_1 S_2^* \rangle \right]$$

where F denotes Fourier transformation. $\langle S_1 S_2^* \rangle$ is the real part of $\langle S_1 S_2^* \rangle$. The Fourier transform of a real and even function is also real and even. Now, $\langle S_1 S_2^* \rangle$ is the autocorrelation function of the complex amplitudes of the speckle field. Goodman (1984 p38) has shown that this function, which is also known as the mutual intensity of the field, is proportional to the intensity distribution incident on the scattering area (assuming that the microstructure of the scattering surface is too fine to be resolved by a lens in the observation plane the size of the observation region) $|P(\xi, \eta)|^2$.

Component 1 can then be interpreted as having the shape of the autocorrelation function of $|P(\xi, \eta)|^2$, just as in equation (4.24) for a single speckle pattern.

Component 2 has the shape of $|P(\xi, \eta)|^2$ itself.

It was shown in section 4.4 i) that the probability density function would be close to Gaussian for a beam ratio of 10, i.e.

$$I_B = 10 \langle I_S \rangle \quad (4.48)$$

In this case component 2 predominates. The spectrum has the shape of the scattering area.

CHAPTER FIVE

METHODS USED TO FABRICATE MULTISCALE

RANDOM ROUGH SURFACES

5.1 Introduction

This chapter is divided into two parts. In part A it is shown how the theory of speckle expounded in chapter four can be applied to produce a one dimensional multiscale, random rough surface. The power spectrum, probability distribution, and contrast of such a surface are considered in turn. The various practical set-ups considered are described, before a full description of the method finally chosen is given. In part B two alternative methods for producing two dimensional surfaces are proposed. The first uses the same principles as the one dimensional method. The second is based on the alternative way of obtaining a Gaussian probability density outlined in chapter four. Again the power spectrum, probability density and contrast are considered in turn. Practical considerations conclude each discussion.

5A A one dimensional surface

5A.1 Principles

5A.1 i) Power spectrum

In chapter four it was shown that the power spectrum of a speckle pattern, for the non-imaging geometry of figure 4.3, has, in addition to a delta function component at zero frequency, a component which depends on the intensity distribution incident on the scattering area. The power spectrum of intensity in the observation plane is written

$$G(u,v) = \langle I \rangle^2$$

$$\left\{ \delta(u,v) + \frac{\iint_{-\infty}^{\infty} |P(\xi,\eta)|^2 |P(\xi-\lambda zu, \eta-\lambda zv)|^2 d\xi d\eta}{\left[\iint_{-\infty}^{\infty} |P(\xi,\eta)|^2 d\xi d\eta \right]^2} \right\} \quad (4.24)$$

The second component has the shape of the autocorrelation of $|P(\xi,\eta)|^2$, where $|P(\xi,\eta)|^2$ is the intensity distribution incident on the scattering area.

Consider then, the effect of choosing the scattering area shown in figure 5.1. The power spectrum for the u component of spatial frequency will have the shape of the autocorrelation function of the chosen area. This may be thought of as the result of shearing the pair of slits. The shape so produced is shown in figure 5.2.

In order to build up a power spectrum of power law dependence on spatial frequency, patterns of this kind are superimposed for several different slit separations. The individual triangular spectra sum to give the required resultant curve. This is shown in figure 5.3. The triangles are arranged so that the peak of one is at the same abscissa as the feet of the two on either side. Figure 5.3 shows both the required curve and that actually obtained from the summation of spectra. The difference between them is minimal when likely experimental variations are considered.

It may be easily shown that the width of each triangle is $2w/\lambda z$, and they must be spaced at intervals in frequency of $w/\lambda z$. w is the slit width; λ is the wavelength; z is the distance from slit to observation plane. Consider Young's double slits, in figure 5.4. Bright fringes appear at

$$y = n\lambda z/d \quad (5.1)$$

i.e. at frequency $d/\lambda z$.

Figure 5.1.

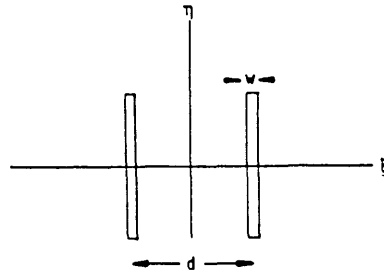


Figure 5.2.

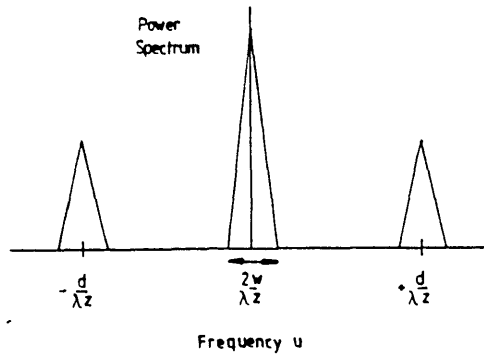


Figure 5.1. Proposed shape of scattering area.

Figure 5.2. Power spectrum corresponding to scattering area in figure 5.1.

Thus, the peak in the power spectrum for slit separation d , occurs at frequency $u = d/\lambda z$.

Due to the shearing effect of autocorrelation the frequency width of the triangular spectrum must be twice the width of each slit, i.e. $2w/\lambda z$. Thus for a spectrum built up as suggested with each peak corresponding with the abscissa of those on either side,

$$d/\lambda z = 2w/2\lambda z \quad (5.2)$$

i.e.

$$d = w \quad (5.3)$$

So, the successive slit separations must differ by the value of the slit width.

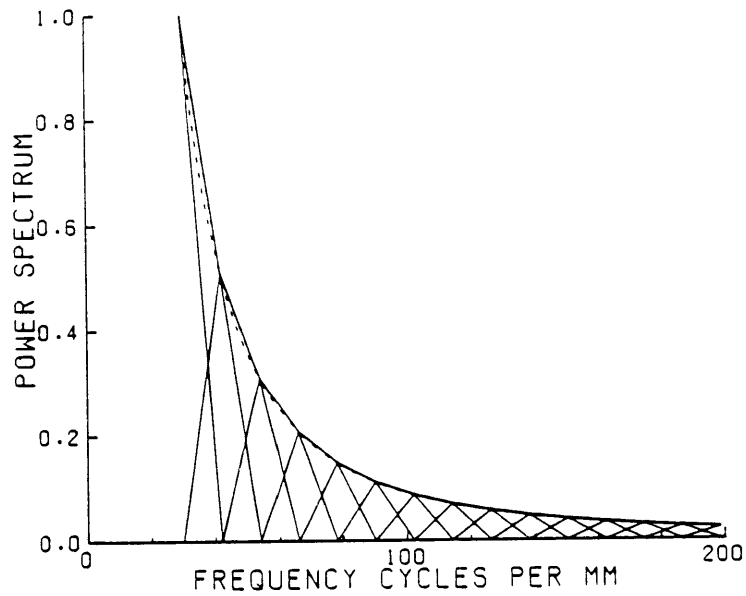


Figure 5.3. Spectrum built up from the summation of individual triangular spectra (solid), compared with an inverse power law curve (dashed).

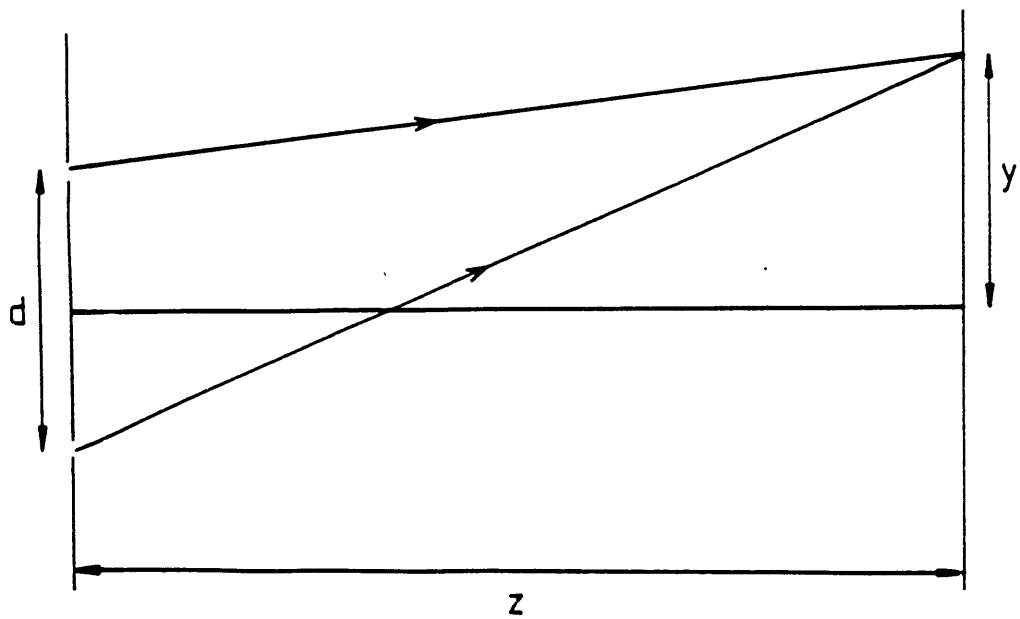


Figure 5.4. Geometry of Young's slits.

From equation (4.24) it is clear that the height of each triangle is proportional to the square of the intensity of the corresponding speckle pattern, due to the factor $\langle I \rangle^2$. When recording in photoresist, assuming that reciprocity failure is absent and the response is linear, the mean intensity term is equivalent to the mean exposure given. A less intense speckle pattern exposed for a longer time would be indistinguishable from a more intense one with a shorter exposure time. Thus, when recording, we say instead that the height of each triangle is proportional to the square of the exposure, $E(u)$, for the corresponding speckle pattern.

$$E(u) = \langle I \rangle \cdot t(u) \quad (5.4)$$

therefore

$$G(u) = k \langle I \rangle^2 t^2(u) \quad (5.5)$$

The form required is

$$G(u) = k' u^{-\alpha} \quad (5.6)$$

$$t^2(u) = K' u^{-\alpha} \quad (5.7)$$

$$t(u) = K u^{-\alpha/2} \quad (5.8)$$

i.e. the exposure times must follow an inverse power law of index $\alpha/2$.

Consider now the effect of this choice of scattering area on the v component of spatial frequency. Consider shearing the pair of slits in the direction as shown in figure 5.5.

The resulting spectrum for the v component of frequency must be a broad triangle of frequency width $2\ell/\lambda z$.

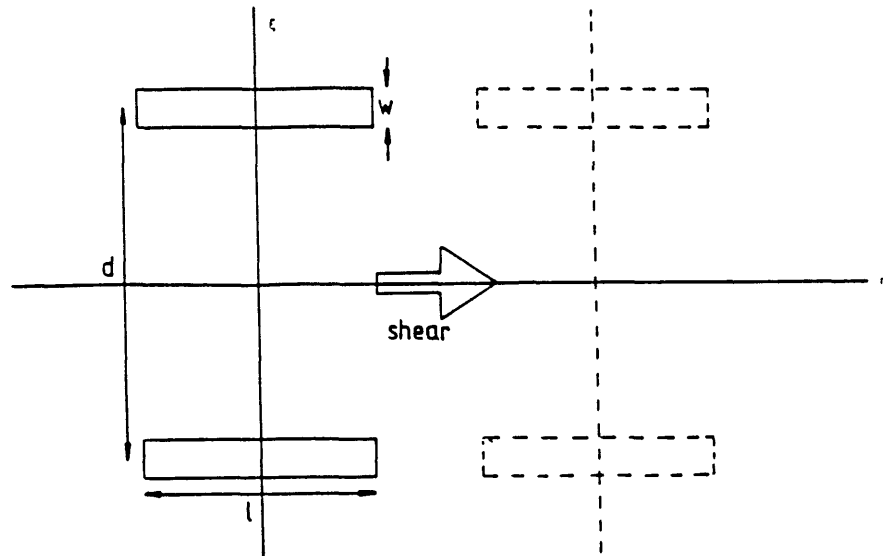


Figure 5.5. To illustrate the shearing of two slits in the η direction.

The same result will occur whatever the slit separation d . So, many frequencies will be present in the y direction, but they have a power law spectrum of a truncated linear form. This is why method A is said to produce a one dimensional multiscale surface.

5A.1 ii) Probability density

If N totally uncorrelated speckle patterns are summed incoherently, the probability density of total intensity is (Goodman 1984; section 4.3 i)

$$p(I) = \sum_{k=1}^N \frac{\lambda_k^{N-2}}{\prod_{\substack{p=1 \\ p \neq k}}^N (\lambda_k - \lambda_p)} \exp\left(\frac{-I}{\lambda_k}\right) \quad (4.25)$$

for $I \geq 0$, and zero for negative I . λ_k are the mean intensities of the individual speckle patterns (assumed all different) each with the usual negative exponential distribution. The shape of this distribution for several values of N is shown in figure 5.6. The λ_k have been chosen to give $\alpha = 2$. For larger N the shape is reasonably close to Gaussian. The curve for $N = 10$ is shown compared with a Gaussian curve in figure 5.7.

5A.1 iii) Contrast

The contrast of any fully coherent speckle pattern, whether it results from a summation or not, is given by (Goodman 1984; section 4.2 i) e))

$$C = \sigma / \langle I \rangle \quad (4.21)$$

The standard deviation of the incoherent sum of N speckle patterns, each of mean intensity λ_k , is

$$\sigma = \left[\sum_{k=1}^N \lambda_k^2 \right]^{\frac{1}{2}} \quad (4.29)$$

Contrasts for incoherently summed patterns are shown in figure 5.8 for a selection of numbers of patterns. The crosses show how if the mean intensity of each pattern is the same then the resultant pattern has a very low contrast once more than about five patterns are summed. The loss of contrast is far less severe for the other set of points. Here the mean intensities are such that the power spectrum of the resultant pattern will follow a power law, with index -2.

Note added after oral examination.

The crosses shown in figure 5.8 are incorrectly plotted. The contrast does not become as low as is shown when patterns of the same mean intensity are summed incoherently.

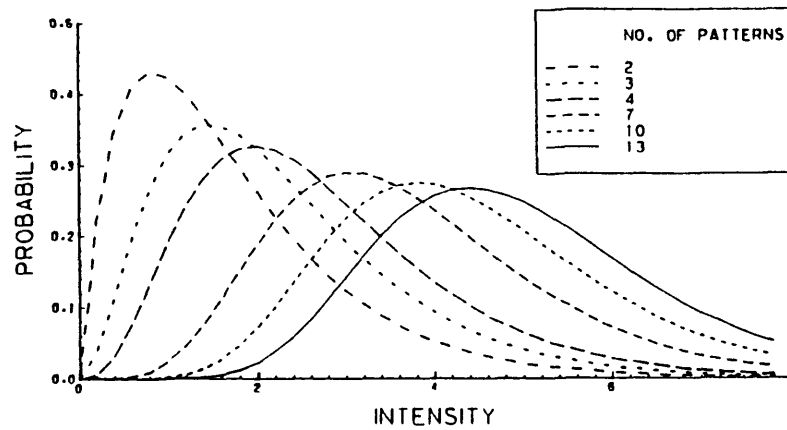


Figure 5.6. Probability density function for the incoherent sum of a number of fully developed speckle patterns, each of mean intensity chosen to give $\alpha = 2$.

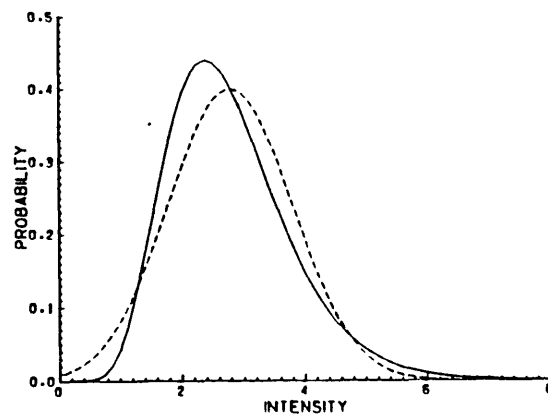


Figure 5.7. Probability density function (solid) for the incoherent sum of 10 fully developed speckle patterns, each of mean intensity chosen to give $\alpha = 2$, compared with a Gaussian curve (dashed) of the same mean and standard deviation.

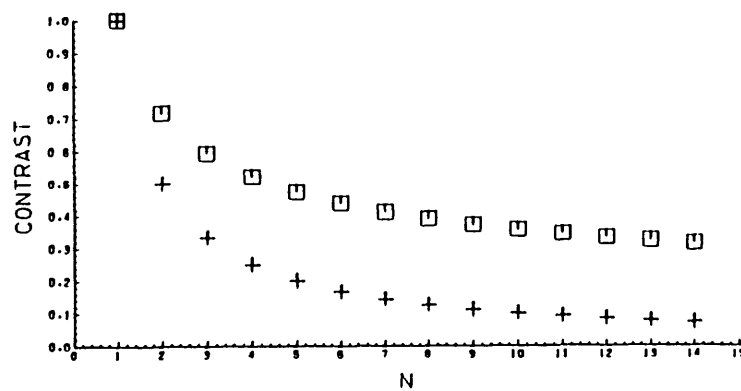


Figure 5.8. Fall off in contrast as the number, N , of incoherently summed speckle patterns is increased. Comparison between case where the individual mean intensities are all equal (+), and where they are chosen to give $\alpha = 2$ ($\boxplus$).

5A.2 Practice

5A.2 i) Possibilities

It can be seen that a system is required which allows the separation between the slits to be adjusted. An early method considered was to use several sets of slits, each of an appropriate separation. The slits were to be drawn on microfilm using the Imperial College Computer. This method was abandoned primarily because there was appreciable leakage of light through the dark background areas of the film.

An optical method seemed more appropriate. The first possibility was to exploit the principles of the Lloyd's mirror arrangement. Figure 5.9 shows how the Lloyd arrangement can mimic a Young's double slit set-up. Light reflected from the mirror interferes with light which has travelled straight from the slit to the screen. For the purposes of my work the effective slit separation could be changed by altering the distance d shown in figure 5.9. A diffuser is placed next to the slit.

Plate 5.1 a) shows a speckle pattern obtained from a single slit diffuser. b) shows the speckle when a Lloyd's mirror is included.

It can be seen from figure 5.10 that the extent and position of the region in which the fringes occur depend on a number of parameters. From figure 5.10

$$\theta_1 = d/2l_1 = b/(m+l_2) \quad (5.9)$$

$$\theta_2 = d/2(l_1+m) = a/l_2 \quad (5.10)$$

therefore

$$a = l_2 d/2(l_1+m) \quad (5.11)$$

$$b = (m+l_2) d/2l_1 \quad (5.12)$$

The extent of the region of overlap = $b-a$

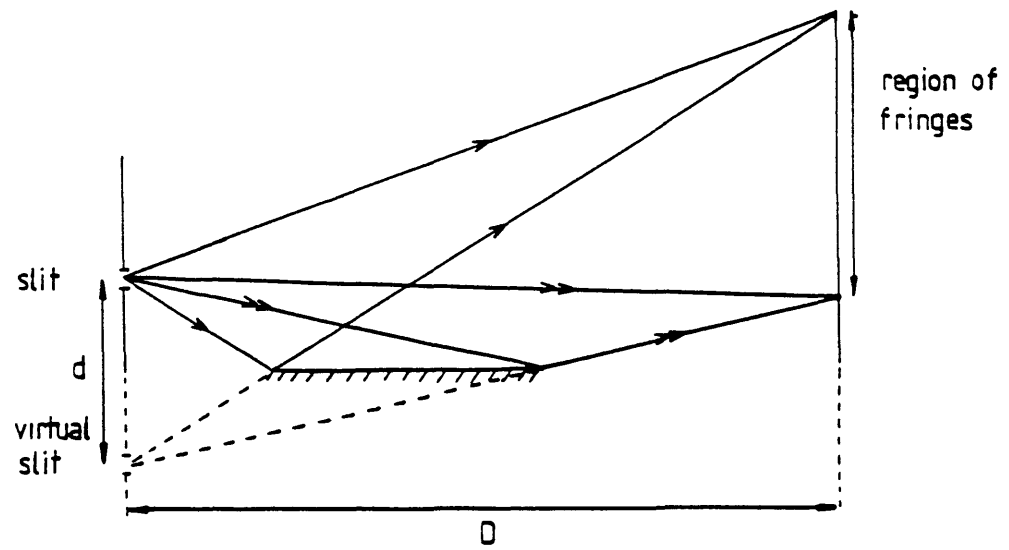


Figure 5.9. Lloyd's mirror arrangement.

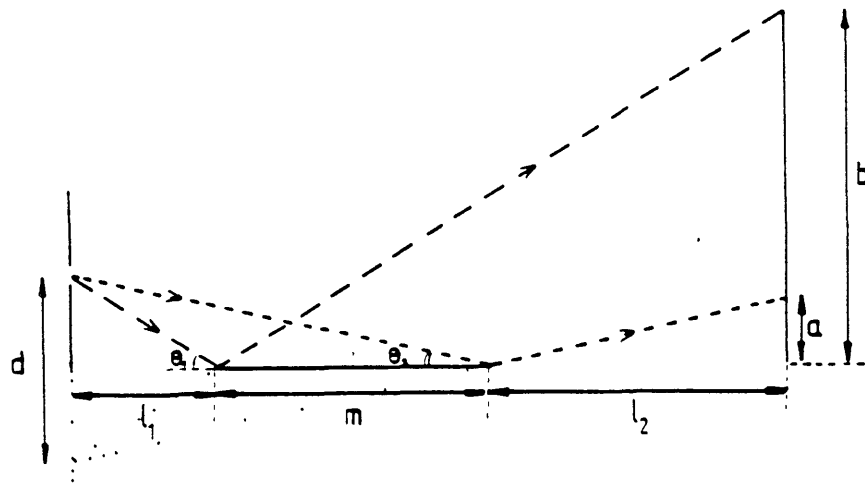


Figure 5.10. Lloyd's mirror to illustrate parameters in equations (5.9) to (5.13).

Plate 5.1. Speckle from Lloyd's mirror arrangement of figure 5.9. a) without mirror, b) with mirror.

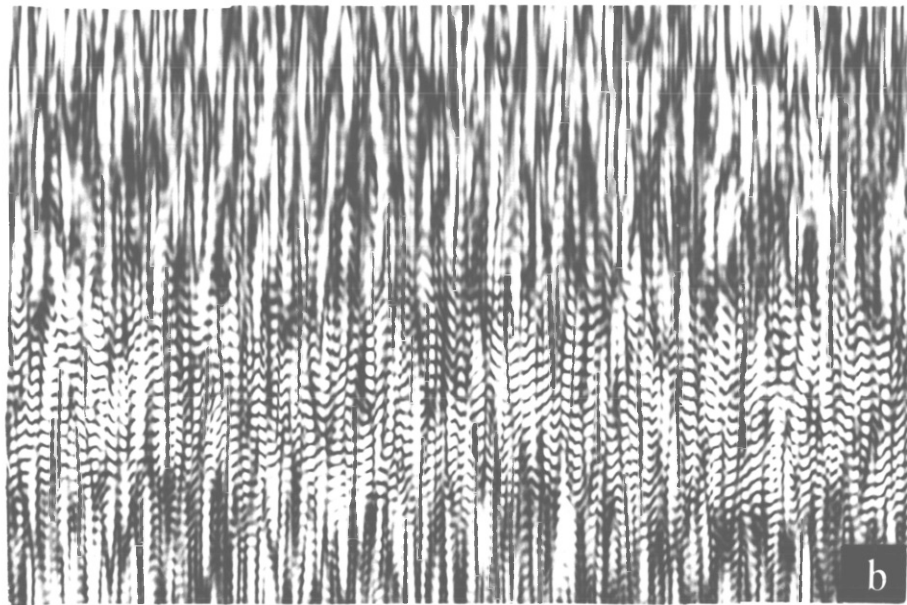
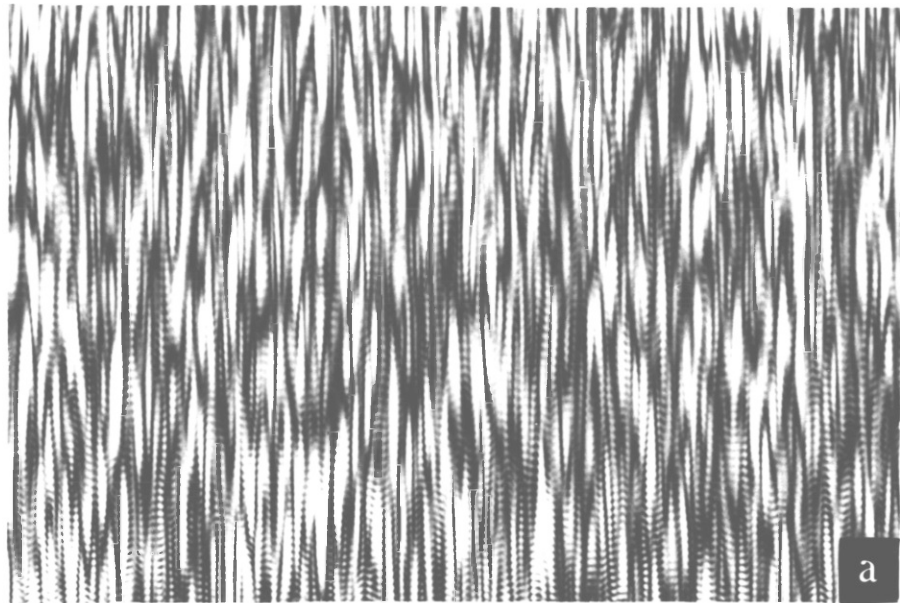


PLATE 5.1

$$= md (l_1 + l_2 + m)/2 l_1 (l_1 + m) \quad (5.13)$$

Typical parameters tried were

$$m = 5 \text{ cm}$$

$$d = 2 \text{ mm}$$

$$l_1 = 30 \text{ cm}$$

$$l_2 = 1.65 \text{ m}$$

these lead to a region around one millimetre in extent, which was considered too limited for practical use. In addition the method was inelegant and difficult to set up and use.

The last method tried proved successful for producing a one dimensional surface, this involved using a cube beam splitter to produce the virtual slit.

5A.2 ii) Method chosen

5A.2 ii) a) Use of the cube beam splitter

This was used as indicated in figure 5.11. The cube beam splitter is set so that the incident beam is split into two emerging parallel beams. The effective separation of the pair of virtual slits is changed by translating the cube as shown in figure 5.12. The cube has rotation and translation movements; the rotation sets the parallelism of the beams, and the translation adjusts the separation.

The separation of the slits, for about ten positions of the translation stage, was measured using a travelling microscope. A calibration line was plotted using these data, so that the slit separations required could be easily set.

Speckle patterns obtained from this arrangement are shown in plate 5.2. The slit to plate distance for these exposures was 2.17m, the slit width 2mm, and the slit separations 5 mm in a) and 28 mm in b). The expected dominant frequencies would be 5000 per metre, and 28000 per metre respectively. These values suggest that the

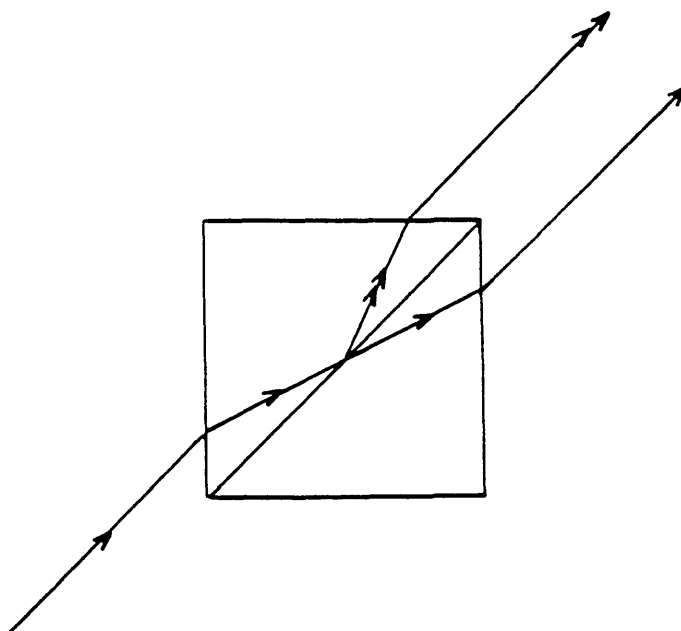


Figure 5.11. Use of cube beamsplitter to give virtual slit.

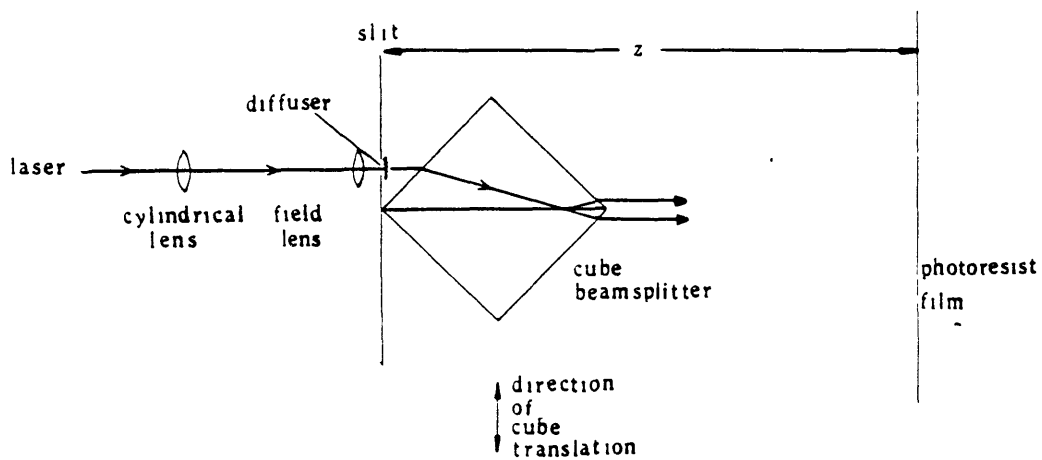


Figure 5.12. Experimental arrangement for a one-dimensional surface.

Plate 5.2. Speckle from cube beamsplitter arrangement of figure 5.12. a) slit separation 5mm, b) slit separation 28mm.

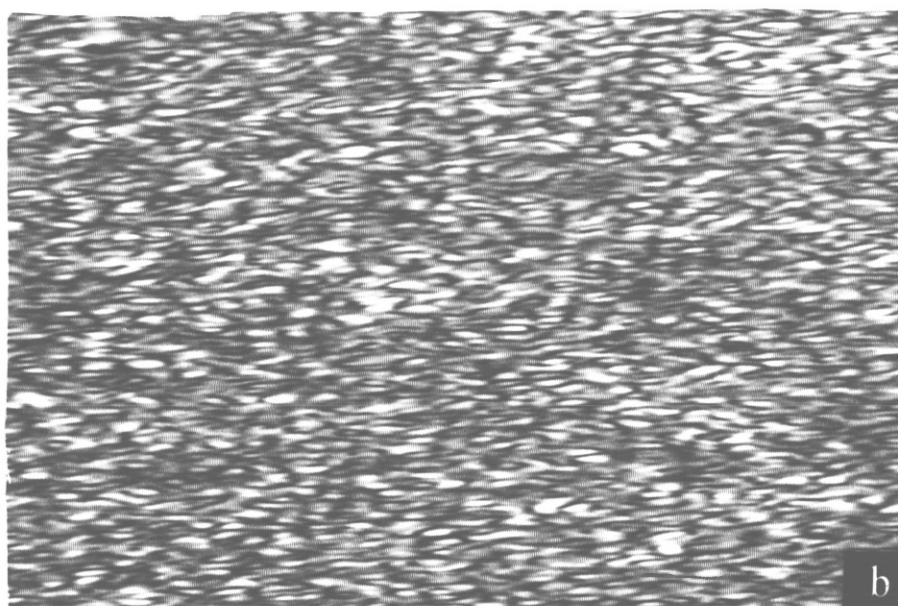
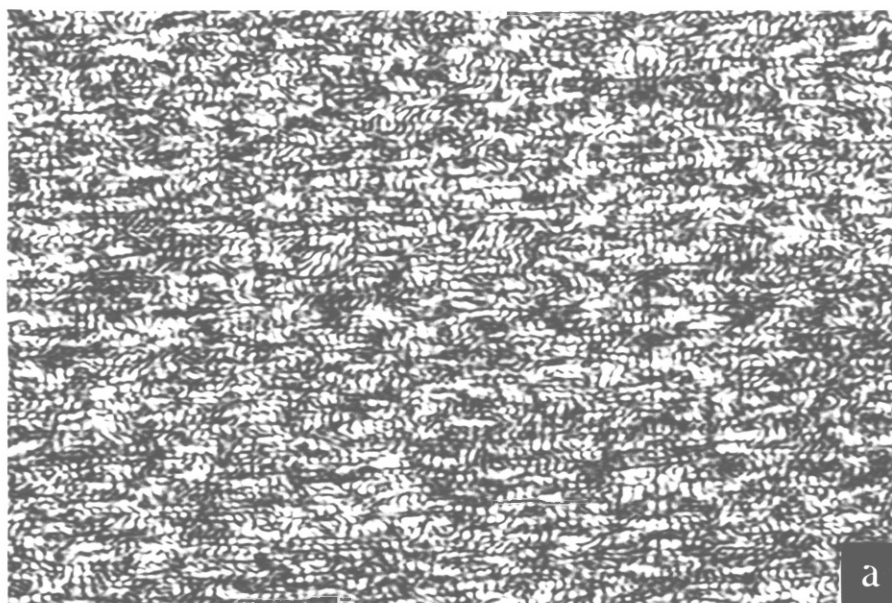


PLATE 5.2

separations of the 'stripes' on the speckle patterns a) and b) should be in the ratio approximately 6:1 given the same degree of enlargement.

5A.2 ii) b) Selection of parameters

The available range of slit separations is limited by two factors. First there is a limit to the number of exposures which can be made on the photoresist. Too many exposures produce a decrease in contrast in the summed speckle pattern. There is also a practical limit to the thickness of resist which can be used (see section 3.3). With excess exposure a thin layer will be simply burned through. Secondly the cube dimensions (40mm by 45mm by 45mm) define the range of slit separations. In figure 5.13 the limiting ray is drawn. If travelling in direction a) it represents the closest separation possible. A ray entering the cube at a distance from the axis greater than $S_{\max}$ will not encounter the boundary, and so not be divided. If travelling in direction b), the ray represents the maximum separation possible.

From figure 5.13, where ℓ is side length of cube

$$\text{In triangle DEF, } \tan \theta_2 = DF/EF = DF/\ell \quad (5.14)$$

$$\text{In triangle EGB, } \cos \theta_1 = GB/EB = S_{\max}/EB \quad (5.15)$$

But, $EB = \ell - DF$

therefore from (5.14)

$$EB = \ell (1 - \tan \theta_2) \quad (5.16)$$

and, $\theta_1 = 45^\circ$

So, into (5.15)

Exposures are made at separations differing successively by the width of the slit, as indicated in figure 5.3 and in section 5A.1 i). The total number of exposures which fit within the possible range of separations will depend on the slit width. Since ten exposures would result in a probability distribution approaching Gaussian (figure 5.7), but not result in excessive exposure, a slit width of 2mm was selected.

It was indicated in section 3.3 iii that an argon ion laser was used to make exposures. The output of the laser was such that z was chosen to be between 30 and 50 cm to keep the exposure times reasonably short.

5A.2 ii) c) Comments on the practical set-up

The slit mount was designed to be positioned as closely as possible to the cube.

It was found to be necessary to coat the backs of the recording plates with black paint to prevent internal reflections producing fringing in the resist. Reflections from the table and surrounding equipment were prevented by a covering of black cloth, or by painting.

5B Two dimensional surfaces

5B a Summation of Gaussians

5B a.1 Principles

5B a.1 i) Power spectrum

Once again the expression (4.24) is invoked. The power spectrum of intensity in the observation plane is written

$$G(u,v) = \langle I \rangle^2 \left\{ \delta(u,v) + \frac{\iint_{-\infty}^{\infty} |P(\xi, \eta)|^2 |P(\xi - \lambda_z u, \eta - \lambda_z v)|^2 d\xi d\eta}{\left[\iint_{-\infty}^{\infty} |P(\xi, \eta)|^2 d\xi d\eta \right]^2} \right\} \quad (4.24)$$

The second component has the shape of the autocorrelation function of the intensity distribution incident on the scattering area. The method proposed to produce a surface with a power law spectrum in both u and v directions of spatial frequency is identical to that described in section 5A. Bearing in mind that the autocorrelation function of a Gaussian function is also Gaussian, speckle patterns produced by Gaussian scattering areas will be superimposed. Clearly the relative widths and heights of the Gaussians must be chosen to result in a power law curve. The Gaussian curves to be summed are all of zero mean, and are described by

$$a \exp(-x^2/\sigma^2) \quad (5.21)$$

where a represents the height of the curve, and σ^2 its variance. x is chosen as the variable for ease of representation. It has no connection with the co-ordinate system. The result of summing N such curves is given by

$$\sum_{i=1}^N a_i \exp(-x^2/\sigma_i^2) \quad (5.22)$$

Physically, a_i must be related to the exposure, and σ_i to the width of the illuminated area. In the following derivation it is assumed that the σ_i are known.

The curve that the summation should approximate is $\gamma x^{-\alpha}$. The approximation is made by a method of least squares. The sums of the squares of the residuals are minimised with respect to a_m , over a range of x values X_1 to X_2 , i.e.

$$\frac{\partial}{\partial a_m} \int_{X_1}^{X_2} \left[\sum_{i=1}^N a_i e^{-x^2/\sigma_i^2} - \gamma x^{-\alpha} \right]^2 dx = 0 \quad (5.23)$$

$$2 \int_{X_1}^{X_2} \left[\sum_{i=1}^N a_i e^{-x^2/\sigma_i^2} - \gamma x^{-\alpha} \right] e^{-x^2/\sigma_m^2} dx = 0 \quad (5.24)$$

$$\underbrace{\int_{X_1}^{X_2} \left[\sum_{i=1}^N a_i e^{-x^2/\sigma_i^2} \right] e^{-x^2/\sigma_m^2} dx}_{\text{Term 1}} - \gamma \underbrace{\int_{X_1}^{X_2} x^{-\alpha} e^{-x^2/\sigma_m^2} dx}_{\text{Term 2}} = 0 \quad (5.25)$$

Term 1

This term may be expressed as

$$\sum_{i=1}^N a_i \left\{ \int_0^{X_2} e^{-x^2 \left(\frac{1}{\sigma_i^2} + \frac{1}{\sigma_m^2} \right)} dx - \int_0^{X_1} e^{-x^2 \left(\frac{1}{\sigma_i^2} + \frac{1}{\sigma_m^2} \right)} dx \right\} \quad (5.26)$$

$$\sum_{i=1}^N \frac{a_i \sqrt{\pi}}{2 \left(\frac{1}{\sigma_i^2} + \frac{1}{\sigma_m^2} \right)^{\frac{1}{2}}} \left[\operatorname{erf} \left\{ x_2 \left[\frac{1}{\sigma_i^2} + \frac{1}{\sigma_m^2} \right]^{\frac{1}{2}} \right\} \right. \\ \left. - \operatorname{erf} \left\{ x_1 \left[\frac{1}{\sigma_i^2} + \frac{1}{\sigma_m^2} \right]^{\frac{1}{2}} \right\} \right] \quad (5.27)$$

Term 2

This term can be expressed in terms of error functions if α is set to 2.

$$\int_{x_1}^{x_2} x^{-2} e^{-x^2/\sigma_m^2} dx \quad (5.28)$$

$$= \gamma \left\{ \int_{x_1}^{\infty} x^{-2} e^{-x^2/\sigma_m^2} dx - \int_{x_2}^{\infty} x^{-2} e^{-x^2/\sigma_m^2} dx \right\} \quad (5.29)$$

from Gradshteyn and Ryzhik (1980, p337)

$$\int_u^{\infty} \frac{e^{-p^2 x^2}}{x^2} dx = -p\sqrt{\pi} \left[1 - \operatorname{erf}(pu) \right] + \frac{e^{-p^2 u^2}}{u} \quad (5.30)$$

Substituting gives

$$\begin{aligned} \text{Term 2} = & \gamma \left\{ \frac{\sqrt{\pi}}{\sigma_m} \left[\operatorname{erf} \left(\frac{x_1}{\sigma_m} \right) - \operatorname{erf} \left(\frac{x_2}{\sigma_m} \right) \right] \right. \\ & \left. + \frac{1}{x_1} \exp \left(\frac{-x_1^2}{\sigma_m^2} \right) - \frac{1}{x_2} \exp \left(\frac{-x_2^2}{\sigma_m^2} \right) \right\} \end{aligned} \quad (5.31)$$

Thus, overall there are N equations, for $m=1, 2, 3, \dots, N$, of the form

$$\begin{aligned} \sum_{i=1}^N \frac{a_i \sqrt{\pi}}{2 \left(\frac{1}{\sigma_i^2} + \frac{1}{\sigma_m^2} \right)^{\frac{1}{2}}} \left[\operatorname{erf} \left\{ x_2 \left[\frac{1}{\sigma_i^2} + \frac{1}{\sigma_m^2} \right]^{\frac{1}{2}} \right\} \right. \\ \left. - \operatorname{erf} \left\{ x_1 \left[\frac{1}{\sigma_i^2} + \frac{1}{\sigma_m^2} \right]^{\frac{1}{2}} \right\} \right] = \gamma \left\{ \frac{\sqrt{\pi}}{\sigma_m} \left[\operatorname{erf} \left(\frac{x_1}{\sigma_m} \right) - \operatorname{erf} \left(\frac{x_2}{\sigma_m} \right) \right] \right. \\ \left. + \frac{1}{x_1} \exp \left(\frac{-x_1^2}{\sigma_m^2} \right) - \frac{1}{x_2} \exp \left(\frac{-x_2^2}{\sigma_m^2} \right) \right\} \end{aligned} \quad (5.32)$$

Let,

$$\begin{aligned} z_{mi} = & \frac{\sqrt{\pi}}{2 \left(\frac{1}{\sigma_i^2} + \frac{1}{\sigma_m^2} \right)^{\frac{1}{2}}} \left[\operatorname{erf} \left\{ x_2 \left[\frac{1}{\sigma_i^2} + \frac{1}{\sigma_m^2} \right]^{\frac{1}{2}} \right\} \right. \\ & \left. - \operatorname{erf} \left\{ x_1 \left[\frac{1}{\sigma_i^2} + \frac{1}{\sigma_m^2} \right]^{\frac{1}{2}} \right\} \right] \end{aligned} \quad (5.33)$$

$$\text{Term 2} = D_m \quad (5.34)$$

then

$$\sum_{i=1}^N z_{mi} a_m = D_m \quad (5.35)$$

i.e.

$$\begin{bmatrix} z_{11} & z_{12} & \dots & z_{1N} \\ \cdot & & & \\ \cdot & & & \\ \cdot & & & \\ z_{N1} & & \dots & z_{NN} \end{bmatrix} \begin{bmatrix} a_1 \\ \cdot \\ \cdot \\ \cdot \\ a_N \end{bmatrix} = \begin{bmatrix} D_1 \\ \cdot \\ \cdot \\ \cdot \\ D_N \end{bmatrix} \quad (5.36)$$

Two methods were used to solve this equation for a_m . The first is a black box routine, provided by the NAG library, for matrix solution. The second is a Gauss-Seidel iterative method. In either case there is a strong possibility that one or more of the values returned for a_m will be negative. This is mathematically valid, but is of little practical use - all exposures are positive. The Gauss-Seidel method of solution gives less precise answers, although they are quite accurate enough to be used experimentally. For the values of σ_i tried the method resulted in fewer negative a_m than the first technique, and was consequently favoured.

Note that for values of α other than 2, D_m , given by

$$\gamma \int_{X_1}^{X_2} x^{-\alpha} e^{-x^2/\sigma_m^2} dx \quad (5.37)$$

can be evaluated numerically, and equation (5.35) solved by the same techniques.

It is now necessary to consider how the values of a_m are related to the exposure times. The situation is slightly different from the summation of triangles described in section 5A.1 i). Once again the heights in the power spectrum

$$\begin{aligned} a_i &\propto (\text{exposure})^2 \\ &\propto t_i^2, \text{ for constant } \langle I \rangle \end{aligned}$$

Thus to get the power spectrum to fall off as $1/\text{freq}^2$, Gaussians with heights t_i^2 should be summed. The exposure times t_i are in the same ratios as the square roots of the a_m calculated from equation (5.35). It will be seen in section 5Ba.2 ii) that possible values for σ_i are dictated by the focal lengths of the lenses in the experimental arrangement. A set of σ_i which was found to result in positive a_i to give a γx^{-2} curve, is shown in table 5.1, with the corresponding a_i .

Note that $\sigma_5/\sigma_1 = 25$. This corresponds to the ratio of maximum to minimum frequency, and is much better than for the one dimensional case where the value was 5. In addition, as the beam profile is Gaussian in any radial direction, the power spectrum will have the required form in any direction.

5Ba.1 ii) Probability density

As for method A, the surface is made by summing N totally uncorrelated speckle patterns, each of mean intensity λ_k . The probability density of total intensity is given by equation (4.25) (Goodman 1984; section 4.3 i).

TABLE 5.1

Solutions to equation (5.35)

σ	a	relative exposure time
1	1.659	1
2	.2866	.4157
4	.09905	.2444
8	.02289	.1175
25	.005230	.05615

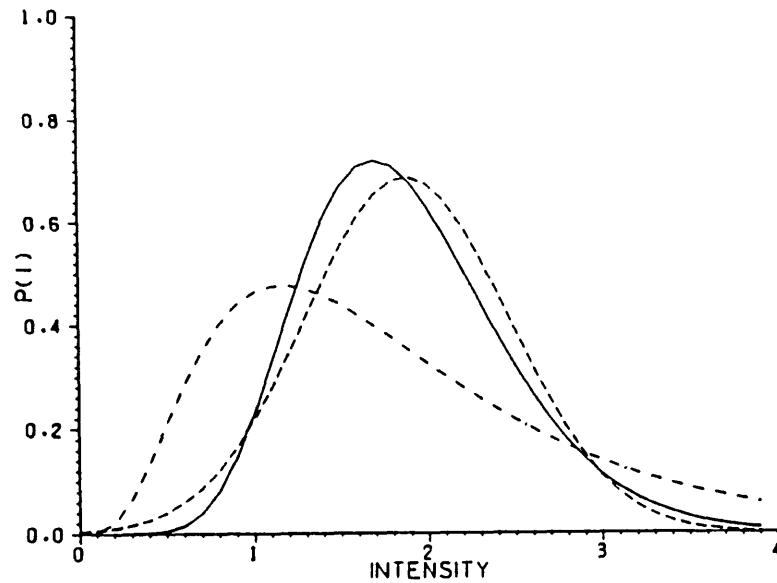


Figure 5.14. Probability density function (dashed and dotted) for incoherent sum of five fully developed speckle patterns whose relative exposure times are tabulated in table 5.1. The solid curve shows the effect of splitting the longer times so that twelve exposures are made. Also shown is a Gaussian curve (dashed) of the same mean and standard deviation.

$$p(I) = \sum_{k=1}^N \frac{\lambda_k^{N-2}}{\prod_{\substack{p=1 \\ p \neq k}}^N (\lambda_k - \lambda_p)} \exp \left(-\frac{I}{\lambda_k} \right) \quad (4.25)$$

for $I \geq 0$, zero for negative I .

The shape of this distribution for five patterns whose exposure times have been calculated to result in a power spectrum with $\alpha=2$, and are tabulated in table 5.1, is shown dotted and dashed in figure 5.14, compared with a dashed Gaussian curve with the same mean and standard deviation.

It is clear that in this case the predicted probability density is not as close to Gaussian as was the corresponding curve for method A (see figure 5.7). To overcome this problem, the longer exposure times are split, so that a larger number of uncorrelated patterns are summed. This is shown by the solid curve in figure 5.14.

5Ba.1 iii) Contrast

The contrast

$$C = \sigma / \langle I \rangle \quad (4.21)$$

where

$$\sigma = \left[\sum_{k=1}^N \lambda_k^2 \right]^{\frac{1}{2}} \quad (4.29)$$

and

$$\langle I \rangle = \sum_{k=1}^N \lambda_k \quad (5.38)$$

If the same exposure times are used as tabulated in table 5.1, then $C = 0.56$. This is slightly greater than the value of 0.36 for method A, and again corresponds to the smaller relative exposure times.

5Ba.2 Practice

5Ba.2 i) Experimental arrangement

The experimental set-up is shown in figure 5.15.

The size of the illuminated area on the ground glass diffuser can be changed by changing the focal length of the collimating lens. Collimation of the beam was checked by inserting a limiting aperture, of diameter 1cm, in the beam and then reflecting it from a mirror over a path of about 5m. If the spot size remained the same, the beam was assumed to be collimated.

5Ba.2 ii) Selection of parameters

The approximate size of the illuminated area may be deduced geometrically. Consider first the objective, in figure 5.16a

$$\tan\theta = d_b/2f_{obj} \quad (5.39)$$

secondly, for the collimating lens, from figure 5.16b

$$\tan\theta = d/2f_c \quad (5.40)$$

Thus,

$$d = f_c d_b / f_{obj} \quad (5.41)$$

Note that d and d_b are taken to be the distances between the $1/e^2$ points of the beam profile.

Also, the focal length of an objective lens is simply related to its magnification M

$$f_{obj}(\text{mm}) = 160/M \quad (5.42)$$

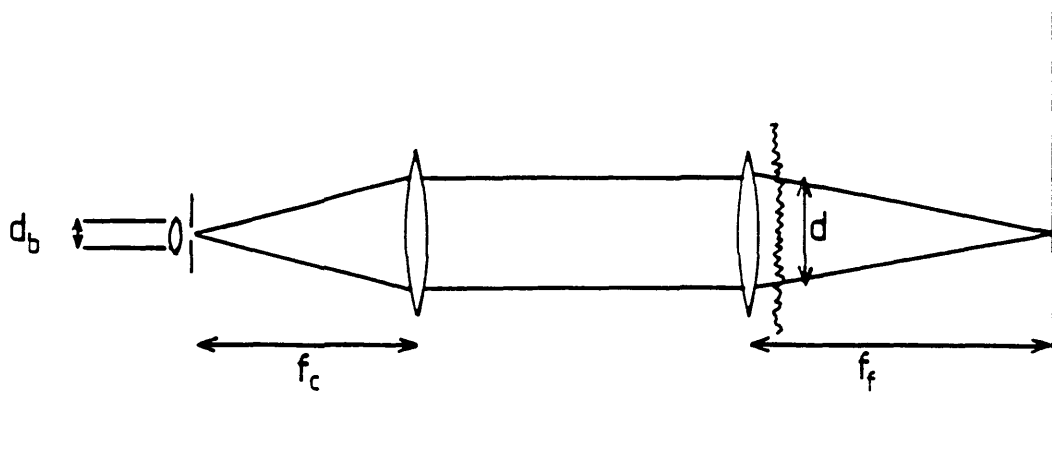


Figure 5.15. Experimental arrangement for a two-dimensional surface.

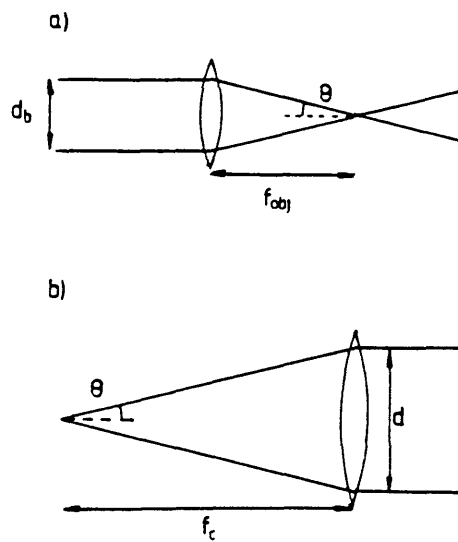


Figure 5.16. a) d_b , f_{obj} and θ for an objective lens.
b) d , f_c and θ for a collimating lens.

Now, from equation (5.41), the diameter of the illuminated area d is proportional to f_c . The values σ_i in the equations of section 5Ba.1 i) are directly proportional to d , since the Gaussian curves in the power spectrum at the recording plane have a shape resulting from shearing Gaussians of width d . Thus $\sigma_i \propto f_{ci}$.

Possible choices for σ_i are limited to the focal lengths of easily available lenses. In addition, as was pointed out in section 5Ba.1 i), the numerical method predicts a best fit without regard to the sign of the curves used. Thus, very often, negative exposure times would be recommended. So a second limitation on σ_i is that the values returned for a_i should be positive. A set of values which satisfy both strictures is tabulated in table 5.1. They correspond to lenses of focal lengths, in millimetres, of 20, 40, 80, 160, and 500.

The diameter of the illuminated area is given by

$$d = f_c d_b / f_{obj} \quad (5.41)$$

now, the speckle size

$$\Sigma = 1.2\lambda z/d \quad (4.18)$$

i.e.

$$\Sigma = 1.2\lambda f_f/d \quad (5.43)$$

Thus, a typical frequency in the pattern at the recording plane

$$= 1/\text{speckle size} \quad (5.44)$$

$$= f_c d_b / 1.2\lambda f_{obj} f_f \quad (5.45)$$

$$If, d_b = 2 \text{ mm}$$

$$\lambda = 457.9 \text{ nm}$$

$$f_{obj} = 16 \text{ mm (x10 objective)}$$

$$f_f = 300 \text{ mm}$$

then the typical frequency = $f_c \times 758 \text{ 300 cycles per metre.}$

For $f_c = 20 \text{ mm}$, frequency = 15 170 per metre, and for $f_c = 500 \text{ mm}$, frequency = 379 200 per metre. As for method A the choice of f_f was determined by the typical exposure times required.

5Bb Speckle plus coherent background

5Bb.1 Principles

5Bb.1 i) Power spectrum

It was shown in section 4.6 that the power spectrum for the sum of a speckle pattern plus a coherent background has two components. The first is the Fourier transform of the square of the real part of the autocorrelation function of the complex amplitudes of the speckle field,

i.e.

$$\text{Component 1} = F \left[\left| \langle S_1 S_2^* \rangle \right|^2 \right]$$

where F denotes Fourier transformation.

The second is given by

$$\text{Component 2} = F \left[2I_B \langle S_1 S_2^* \rangle \right]$$

where I_B is the intensity of the coherent background beam.

If the scattering area is described by $|P(\xi, \eta)|^2$, then component 1 has the shape of the autocorrelation function of $|P(\xi, \eta)|^2$. Component 2 is shaped as $|P(\xi, \eta)|^2$.

In order to attain a probability density function close to Gaussian, the ratio of the background beam intensity to the speckle field intensity should be around 10 (see figure 4.8). In this case, component 2 will be dominant and the power spectrum has the shape of the intensity distribution of the scattering area. The required spectrum has a power law dependence on spatial frequency, thus the intensity distribution across the scattering area should also obey a power law.

5Bb.1 ii) Probability density

From section 4.4 i) the probability density function is given by

$$p_I(I) = (r+1) \exp \left[-r - I(r+1) \right] I_0 \left\{ 2 \left[Ir(r+1) \right]^{1/2} \right\} \quad (4.34)$$

where

$$r = I_B / \langle I_S \rangle \quad (4.32)$$

and $I_0\{\}$ is a modified Bessel function of the first kind, zero order (Watson 1962 chap 3). Curves from equation (4.34) are shown in figure 4.7.

5Bb.1 iii) Contrast

The contrast C of the resulting pattern is given by

$$C = (2r+1)^{1/2} / r+1 \quad (4.35)$$

For $r = 10$, $C = 0.42$. This is of the same order as the contrasts resulting from the incoherent summation of 10 speckle patterns described in section 5A, and shown in fig 5.8.

5Bb.2 Practice

It is clear that the first requirement of this method is an apodizing filter which ensures that the distribution of intensity on the diffuser has a radial distribution, $r^{-\alpha}$. The technique used to produce such filters is outlined below.

The method proposed involves exposing a rotating photographic plate through a sector mask of transmitting and opaque regions. The mask will have four sectors, each of the same shape, as indicated in figure 5.17, where the shaded regions are transparent. The transparent regions are defined by the circumference of the circle, a radius, and a calculated curve which will hereafter be called the mask curve. The use of four sectors instead of a single one ensures that the effect of random irregularities in the mask curve are minimised, and reduces exposure times.

The light incident on the rotating plate at a particular radius r_i must depend on the time the plate takes to transverse the transparent section of the mask at that radius.

$$\tau(r_i) \propto \text{length of arc/velocity} \quad (5.46)$$

$$= kr_i \theta_i / r_i \omega \quad (5.47)$$

$$= k\theta_i / \omega \quad (5.48)$$

For constant ω

$$\tau(r_i) = k\theta \quad (5.49)$$

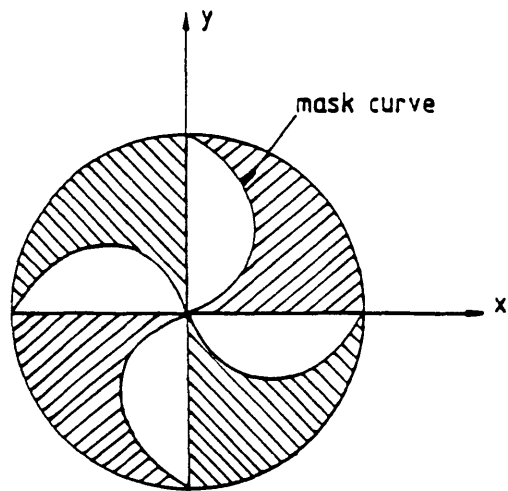


Figure 5.17. General form of a sectored mask.

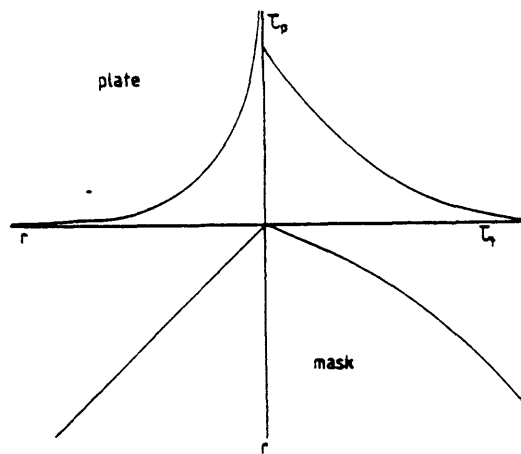


Figure 5.18. The four quadrants used to predict the form of mask required.

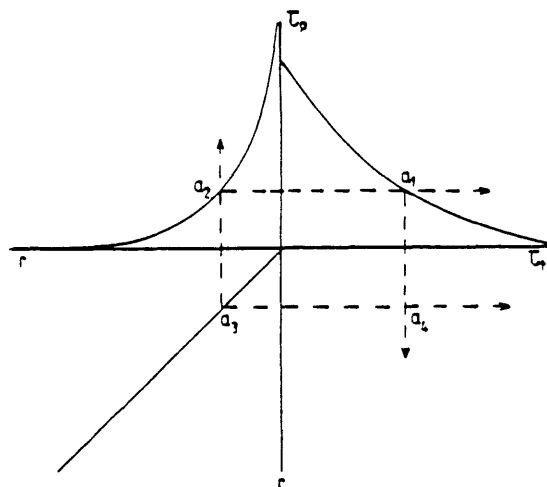


Figure 5.19. Mapping between the quadrants.

when $r_i = r_{\max}$, $\theta_i = \theta_{\max}$

($\theta_{\max} = 90^\circ$ for example shown)

$$k = \tau(r_{\max}) / \theta_{\max} \quad (5.50)$$

so,

$$\theta_i = \tau(r_i) \theta_{\max} / \tau(r_{\max}) \quad (5.51)$$

Thus the form of the mask curve can be predicted, using cartesian co-ordinates as follows

$$\begin{aligned} x_i &= r_i \cos \theta_i \\ y_i &= r_i \sin \theta_i \end{aligned} \quad (5.52)$$

together with expression (5.51) for θ_i . Only the form of $\tau(r_i)$, the mask transmission, remains to be found. Knowledge of the photographic response of the plate must be incorporated, showing the relation between mask transmission and the resulting transmission of the plate. A calibration curve is obtained by exposing a plate through a step tablet, which has 21 steps, each of different density, ranging from 0.05 to 3.05. The density variation across both step tablet and exposed plate is measured with a densitometer. A plot of tablet transmission vs plate transmission can be made, using the relation

$$\tau = 10^{-D} \quad (5.53)$$

where τ is the transmission and D the density.

A graphical prediction of the required mask transmission with radius can be obtained. The method is based on the objective tone reproduction curve described by James and Higgins (1960). A set of axes is drawn to define four quadrants as shown in figure 5.18. In the first quadrant is a plot of the tablet transmission vs plate transmission. The third quadrant contains a line of unit slope, representing the radial co-ordinate r . A curve can be drawn in the

second quadrant showing the required relation between the plate transmission and radius. A prediction can then be made of the required mask transmission with radius. The mask transmission relation is obtained by mapping a point on the line in the third quadrant directly into the fourth quadrant, and indirectly via the curves in the second and first quadrants to the fourth quadrant. The point of intersection of these in the fourth quadrant gives a point on the mask transmission curve. An example is shown in figure 5.19. A smooth curve is fitted by computer to the points obtained from the graph. This curve is the $\tau(r_i)$ curve of equation (5.51).

So, the mask curve is described by

$$\begin{aligned}x_i &= r_i \cos \theta_i \\y_i &= r_i \sin \theta_i\end{aligned}\tag{5.52}$$

where

$$\theta_i = \tau(r_i) \theta_{\max} / \tau(r_{\max})\tag{5.51}$$

and

$$\theta_{\max} = 90^\circ$$

$\tau(r_i)$ has been obtained graphically, and also yields $\tau(r_{\max})$

Masks produced by this technique are shown in figure 5.20. Note that in a) the mask curve does not begin until a radius of $0.24 r_{\max}$. This is due to the form chosen for the required plate transmission with radius curve in the second quadrant. In case a) τ_p was very large for small r , and exceeded the maximum measured τ_p of quadrant one. Alternative masks were produced for which the mask curve could be calculated for the whole range of r , and 5.20 b) is an example.

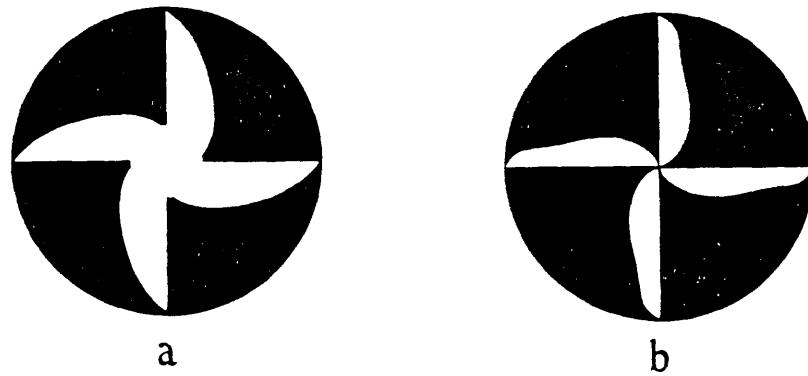


Figure 5.20. Masks used. a) Non-transmitting centre.
b) Transmitting centre.



Figure 5.21. Arrangement for exposure of photographic plates through mask.

The arrangement for exposing the plates is shown in figure 5.21. The plates used were Agfa-Gevaert 8E75 holographic plates, which are designed to respond best to helium-neon laser radiation. Laser light, however, was not used for the exposures because of problems of unwanted fringe production. A white light source, with red filter was used instead. The plate was mounted on a small plate on the spindle of a 6V dc motor. Exposure times depended on the brightness setting of the lamp, and its distance from the plate, but were generally in the region of 10 to 30 seconds.

The plates were developed in Kodak D19 developer for 2 minutes at 20°C , with continuous gentle agitation; rinsed in water; fixed for 2 minutes; and again rinsed in water for about 10 minutes. A final rinse in isopropanol is necessary to remove the blue dye present in the plates, which otherwise has a tendency to cause run marks.

Plates produced however, did not have the required transmission profile. Reasons for the discrepancies are not clear.

1. Whether mask type a or b was used the resulting plate was unexposed to a radius of about $0.2 r_{\text{max}}$.
2. The maximum density should occur at r_{max} . A narrow region of smaller densities might be expected at radii greater than r_{max} . Plates produced always had their maximum density at about $0.8 r_{\text{max}}$. The power law relation was not obeyed at radii $> 0.5 r_{\text{max}}$.
3. Other common faults included the occurrence of dense concentric rings and a mottled appearance.

In the best cases then, the filters produced obeyed the $r^{-\alpha}$ relation over a tiny range of radii, for example a ratio of only 2.5 to 1. A fault of the one dimensional method described in section 5A was that it resulted in a value of only 5 for the ratio maximum to minimum frequency. It was felt that this method was not worth pursuing further, in spite of its apparent simplicity, as the frequency ratio would be inadequate.

CHAPTER SIX

THE PRODUCTION OF SURFACES; MECHANICAL MEASUREMENT OF SURFACES

In this chapter, the parameters used in producing both one and two dimensional multiscale surfaces, and single scale surfaces, are described.

Measurements made on the surfaces with the Talysurf instrument are presented. Comments are included on the accuracy and significance of the results quoted. In particular the experimental determination of the autocorrelation function is considered.

6.1 Parameters used in surface production

6.1 i Parameters for the production of surface 141

The best one dimensional surface produced by the method suggested in chapter 5A is numbered 141. The exposure times, with their corresponding slit separations are shown tabulated in table 6.1. The experimental layout is shown in figure 5.12. The laser beam was expanded and cleaned using a x3 objective with a 50 μ m pinhole. Collimation was by a lens of focal length 10cm. The cylindrical lens used to direct the light onto the slit, was a piece of glass rod of diameter 5mm. The field lens was of 2.75 dioptries, that is of approximately 36cm focal length.

6.1 ii Parameters for the production of surface 200

The best two dimensional surface produced by the method described in chapter 5B, where different sized illuminated areas are used, each with a Gaussian intensity profile, is numbered 200. The exposure times are shown tabulated in table 6.2.

The experimental layout is shown in figure 5.15. The beam was expanded using a x10 objective, with a 25 μ m pinhole. The focal

TABLE 6.1

The Recording of Plate 141

Slit separation cm	Exposure time	
	min	sec
.5	26	56
.7	19	14
.9	14	58
1.1	12	14
1.3	10	22
1.5	8	59
1.7	7	55
1.9	7	06
2.1	6	25
2.3	5	52

TABLE 6.2

The recording of plate 200

Focal length collimating lens mm	Exposure time	
	min	sec
20	4	55
	4	40
	5	09
	4	25
	5	24
40	3	11
	2	56
	4	04
80	2	27
	3	32
160	2	53
500	1	23

lengths of the collimating lenses are listed in table 6.2. The field lens was of focal length 300 mm.

6.1 iii Parameters for the production of surface 201

This surface was produced by the method of Gray (1978), to provide a single scale surface to compare with the multiscale ones. The arrangement was the same as that for plate 200 (section 6.1 ii) and shown in figure 5.15. The focal length of the collimating lens was 20mm. The field lens was of focal length 300mm. Ten exposures were made, each of 4.5 minutes, a different area of the diffuser being used each time.

6.1 iv Parameters for the production of surface 198

This surface was also produced by the method of Gray (1978), to provide a single scale surface to compare with the multiscale ones. The arrangement and method was the same as that for plate 201 (section 6.1 iii) and shown in figure 5.15, but the focal length of the collimating lens was 500mm.

6.2 Talysurf measurements

Talysurf measurements were made using the instrument in the Department of Mechanical Engineering, with the help of Martin West and Martin Webster.

The Talysurf is a stylus instrument, and is the standard device used in engineering work to assess surface texture. A height profile is obtained by tracing a diamond stylus across the rough surface. The vertical movement of the stylus is converted to an electrical signal which is amplified and sent to a computer for storage. The signal can also be sent to a chart recorder.

The stylus used in the measurements presented here has a tip diameter of $1\mu\text{m}$, which puts a limit on the resolution of the measurements, as the stylus will be unable to reproduce faithfully, deep, narrow valleys.

1024 readings were taken at separations of $2\mu\text{m}$. This leads to a total trace length of 2.048mm . The Nyquist frequency is the highest frequency about which meaningful information can be obtained from a set of data,

$$f_N = 1/2\Delta x \quad (6.1)$$

where Δx is the sampling separation. For $\Delta x = 2\mu\text{m}$, $f_N = 250000$ cycles per metre.

The lowest frequency about which meaningful information can be obtained is

$$f_L = 1/L \quad (6.2)$$

where L is the total sample length. For the trace length above, $f_L = 488$ cycles per metre.

Values for the surface height at each sampling point were recorded on floppy discs. These floppy discs can be read by an LSI RT11 computer. Programs were written for the LSI to compute the probability density function, autocorrelation function, structure function and variance from these data. The power spectrum was calculated from the autocorrelation data, on the ICCC CDC computer using a NAG library routine to accomplish Fourier transformation.

For surface 141 ten traces of 1024 points were made across the surface in the x direction, as defined in figure 4.3. For the other surfaces, ten traces were made in each of the x and y directions.

The recordings on 141 had frequencies between approximately 30000 and 138000 m^{-1} , which fall comfortably within the range offered by the sampling separation and number of points selected. For plate 200 frequencies between 15170 and 379200 m^{-1} were expected.

The Nyquist frequency is, in this case, too low if the sampling separation is kept at $2\mu\text{m}$. To try and overcome this problem, an additional ten traces each of 2048 points separated by $1\mu\text{m}$ were made. These would increase the Nyquist limit to 500000m^{-1} . However, when the power spectrum was calculated, it was apparent that no extra information was obtained by this procedure. The spread of points at high frequency was very large, as the finite size of the stylus took effect.

Because of this limit of resolution of the Talysurf, measurements from surface 198 are not presented in this chapter, as a spatial frequency larger than $250\,000\text{m}^{-1}$ is expected. Scattering measurements from this surface do appear in chapter seven.

Talysurf measurements are presented for each function in turn:

- i) Probability density function
- ii) Autocorrelation function
- iii) Variance
- iv) Structure function
- v) Power spectrum

6.2 i) Probability density function

Surface depths are featured, rather than heights to allow direct comparison with the curve predicted by equation (4.25). The expression (4.25) is concerned with the intensity of the light in a speckle pattern. As areas of high intensity in the exposing speckle pattern correspond to deep regions in the developed resist, the intensity curve can be compared directly with the depth measurements. This distinction between depths and heights is important, as the predicted curve is asymmetrical.

The points shown on the plots indicate the statistical uncertainty in the height of each block of the graph. Assuming a Poisson distribution of each bin height, this is taken to be $\pm (N)^{1/2}$, where N is the number of observations in a particular bin.

The block graphs are shown compared with two curves. All three are normalised to have unit area and standard deviation. The solid curve is that predicted by equation (4.25) for relative intensities equivalent to the exposure times used in producing the plate in question. The dashed curve is a Gaussian shown for comparison.

6.2 i) Plate 141

A block graph of the probability density function of the surface depths of plate 141 is shown in figure 6.1.

It can be seen from the figure that the measured histogram is not exactly of the form predicted by equation (4.25). It is lacking in depths around the peak of the predicted curve, that is those which are slightly shallower than the mean. There is however, as predicted, a slight skewness in the measured distribution. The requirement of the model (section 2.8) was a Gaussian probability density. The measured histogram is closer in form to Gaussian than to the predicted curve. This is seen as an advantage rather than as a failing.

6.2 i) Plate 200

Block graphs showing the probability density function of the surface depths for the x and y directions are shown in figures 6.2 i and ii. These plots were made using the traces of 2048 points.

The x direction results agree well with the theoretical curve. The y direction block graph is, like that for plate 141, lacking in depths around the peak of the predicted curve, that is those which are slightly shallower than the mean. The curve is, however acceptably close to Gaussian.

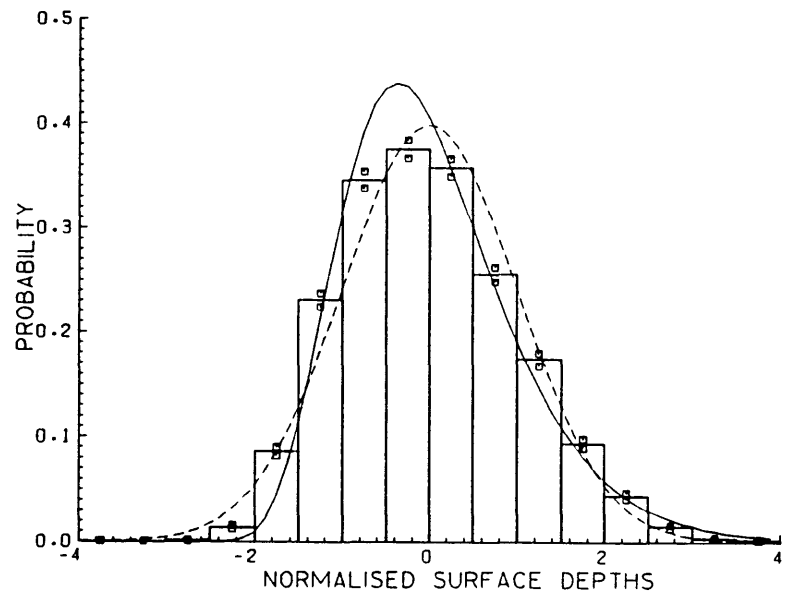


Figure 6.1. Probability density function for plate 141 in x direction.

6.2 i) Plate 201

Block graphs showing the probability density function of the surface depths for the x and y directions are shown in figures 6.3 i and ii. These plots were made using traces of 1024 points.

The y direction results agree quite closely with the Gaussian curve shown for comparison. In the x direction the block graph is somewhat flattened around the mean depth. Similar features were apparent on graphs drawn from traces of 1024 points on plate 200. Graphs drawn from the traces of 2048 points were smoother. It seems likely, therefore, that using more points would result in closer agreement between measured and theoretical curves.

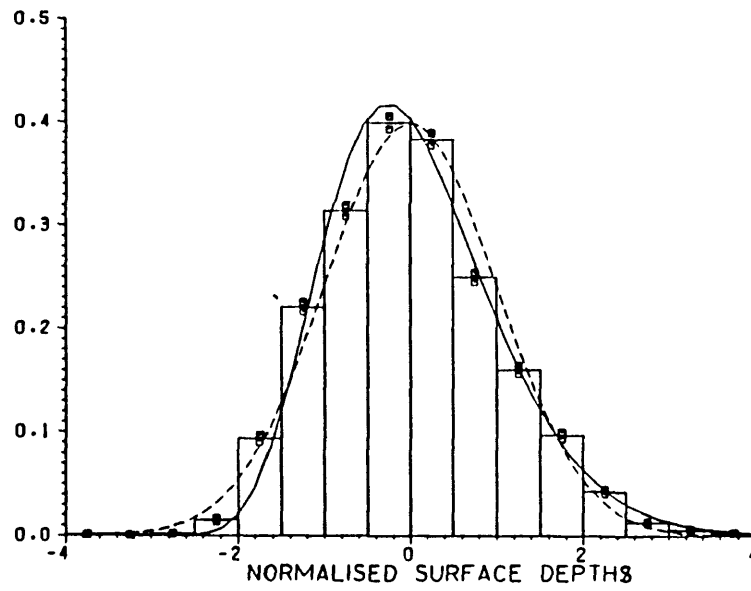


Figure 6.2 i). Probability density function for plate 200 in x direction.

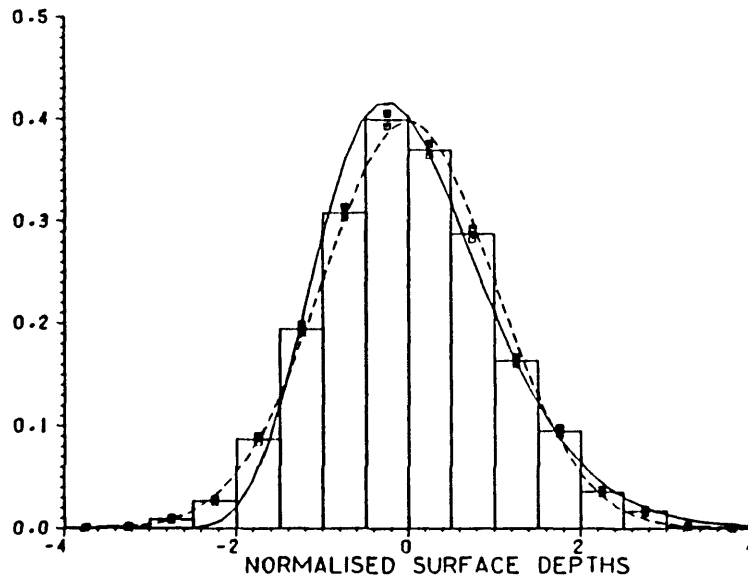


Figure 6.2 ii). Probability density function for plate 200 in y direction.

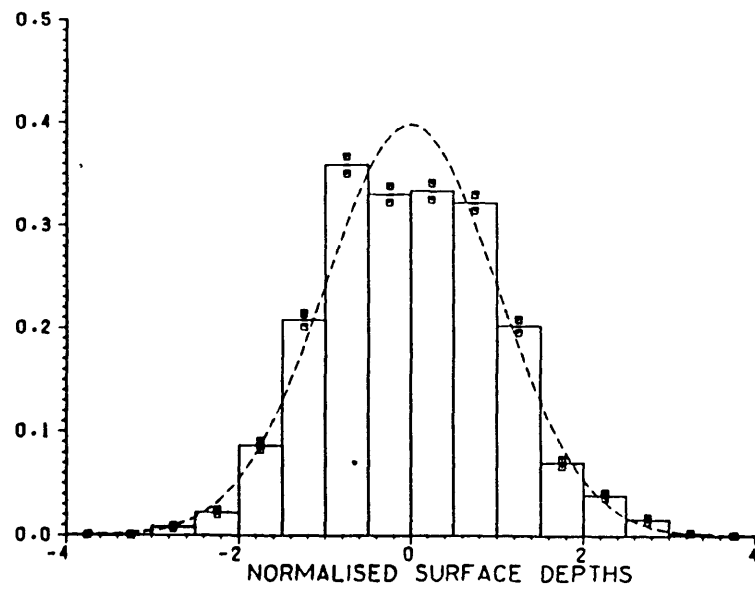


Figure 6.3 i). Probability density function for plate 201 in x direction.

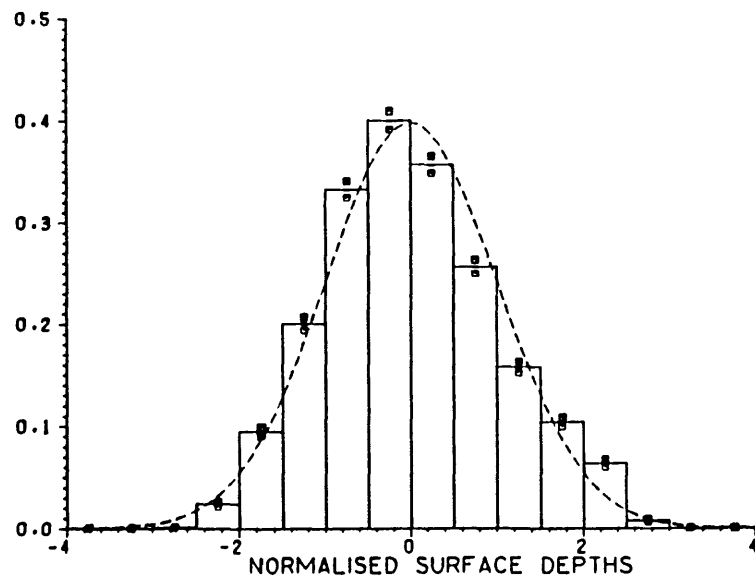


Figure 6.3 ii). Probability density function for plate 201 in y direction.

6.2 ii) Autocorrelation function

The measured, normalised autocorrelation functions are shown. The curves shown are the mean of ten curves, each calculated from a single Talysurf trace of 1024 values. Before each curve was calculated (using equation 1.21) the mean level of the trace was found by a least squares analysis, and each height reading adjusted so that the set of values was a zero mean process.

The error bars in figures 6.4 to 6.6 were found using

$$\sigma_{C'(X)}(X) = \left[\frac{1}{N} \int_{-\infty}^{\infty} \left\{ \left[C'(S) \right]^2 + C'(S+X) C'(S-X) \right\} ds \right]^{\frac{1}{2}} \quad (6.3)$$

where $C'(X)$ is the measured autocorrelation function, and N the total number of lags used in the estimation. This expression has been derived from one given by Jenkins and Watts (1968, p178). The values obtained were then adjusted to allow for the normalisation.

Measured autocorrelations are subject to a degree of biasing which arises because the sample mean (taken to be zero), is not a true population mean, having been found from a limited number of data points. This problem will be discussed further in section 6.3.

6.2 ii) Plate 141

The autocorrelation function in the x direction of plate 141 is shown in figure 6.4.

The curve is unusual in that there is a definite kink which effectively divides the curve into two parts. The autocorrelation function does not therefore obviously fit the expression (2.4)

$$C(X) = \begin{cases} 1 - |X|/\Gamma & |X| < \Gamma \\ 0 & \text{otherwise} \end{cases} \quad (2.4)$$

6.2 ii) Plate 200

The measured, normalised autocorrelation functions for the x and y directions are shown in figure 6.5. That for the x direction is solid, the dashed curve is for the y direction.

These curves are of more satisfactory form than that for plate 141. The curves can be taken as being of truncated linear form, as given by equation (2.4)

$$C(X) = \begin{cases} 1 - |X|/\Gamma & |X| < \Gamma \\ 0 & \text{otherwise} \end{cases} \quad (2.4)$$

The width of the autocorrelation function is given by the outer scale size Γ . Clearly the outer scale size is larger in the y direction, which corresponds to a smaller illuminated region of diffuser. That the scale sizes differ in the x and y directions suggests that the illuminated area on the diffuser was not circular. This point is discussed further in section 6.2 vi).

6.2 ii) Plate 201

The measured, normalised autocorrelation functions for the x and y directions are shown in figure 6.6. That for the x direction is solid the dashed curve is for the y direction.

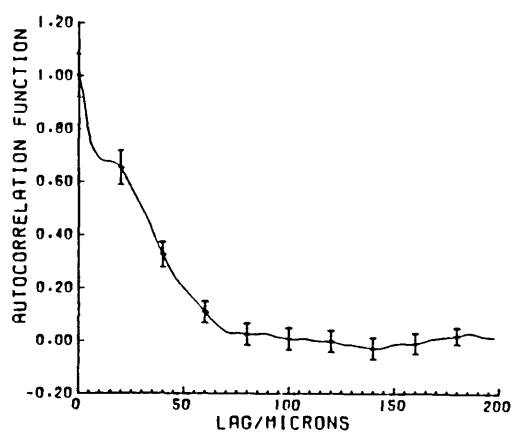


Figure 6.4. Autocorrelation function for plate 141 in x direction.

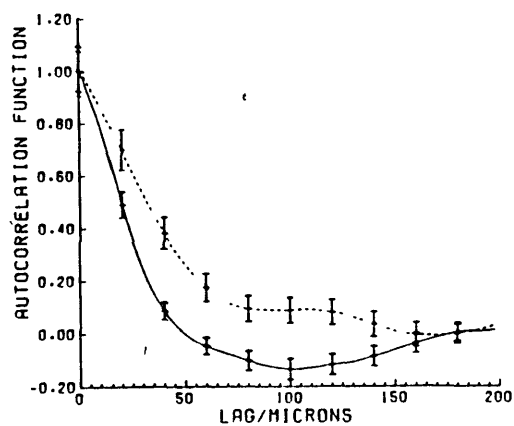


Figure 6.5. Autocorrelation functions for plate 200, x direction solid, y direction dashed.

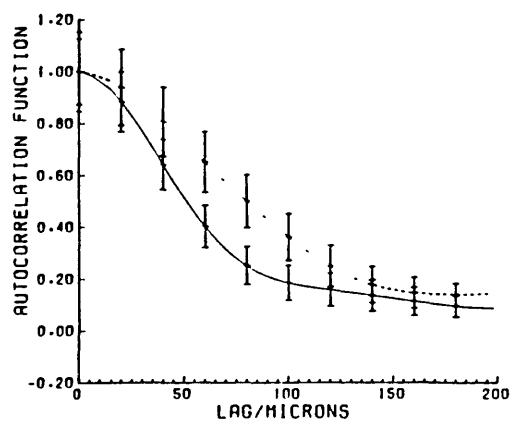


Figure 6.6 i). Autocorrelation functions for plate 201, x direction solid, y direction dashed.

Once again the width of the autocorrelation function is greater in the y than the x direction, corresponding to an illuminated area larger in the ξ than the η direction (figure 4.3). As expected the curves are close to Gaussian in form, and of quite different appearance from those for plate 200. Figure 6.6 ii) shows the x direction curve (solid) compared with a Gaussian curve of correlation length $66 \mu\text{m}$. In figure 6.6 iii) the y direction curve is shown solid, compared with a dashed Gaussian curve of correlation length $95 \mu\text{m}$.

6.2 iii) Variance

As the autocorrelation function is shown normalised it is not possible to read a value for the variance from the graphs presented. The accuracy for the variances shown here has been obtained using equation (6.3).

6.2 iii) Plate 141

The mean value from ten results calculated using equation (1.23) for each trace is $0.17 \pm 0.01 \mu\text{m}^2$.

6.2 iii) Plate 200

In the x direction, variance = $0.29 \pm 0.02 \mu\text{m}^2$

In the y direction, variance = $0.28 \pm 0.02 \mu\text{m}^2$

6.2 iii) Plate 201

In the x direction, variance = $0.34 \pm 0.03 \mu\text{m}^2$

In the y direction, variance = $0.35 \pm 0.04 \mu\text{m}^2$

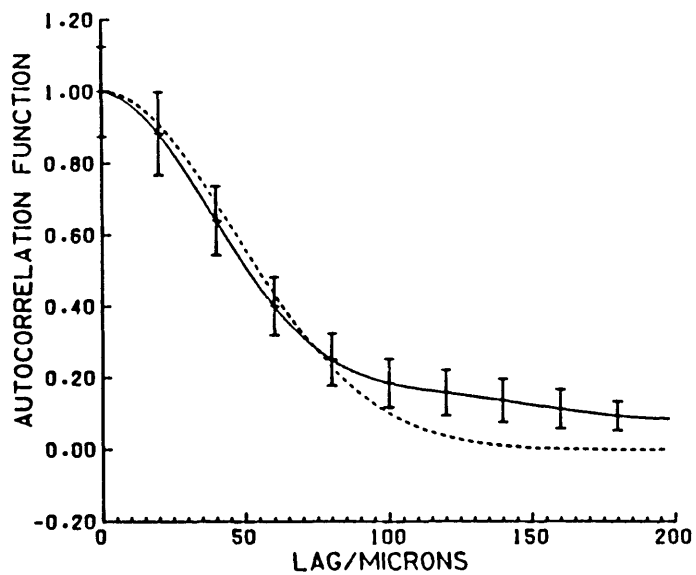


Figure 6.6 ii). Autocorrelation function for plate 201, x direction (solid), compared with a Gaussian curve of correlation length 66 μm .

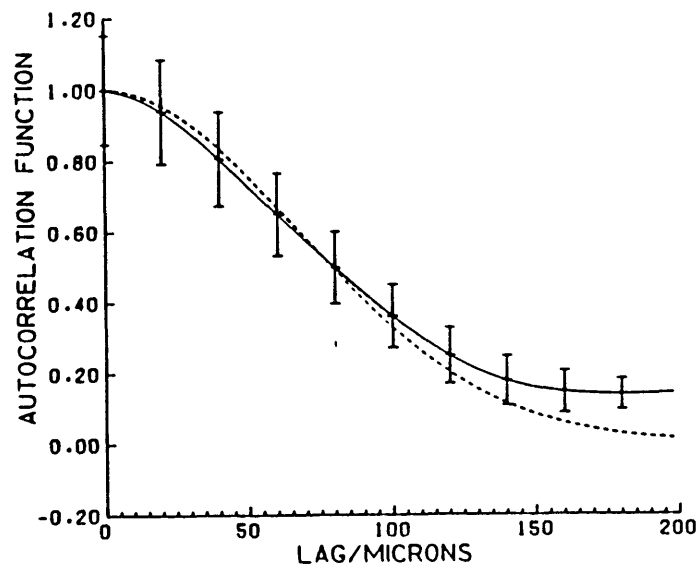


Figure 6.6 iii). Autocorrelation function for plate 201, y direction (solid), compared with a Gaussian curve of correlation length 95 μm .

6.2 iv) Structure function

6.2 iv) Plate 141

The structure function for the x direction is shown in figure 6.7. The slope of the graph in the region of zero lag, before the kink, is $0.015 \mu\text{m}$. Using equation (2.21), and taking $\alpha = 2$,

$$D(X) = T |X| \quad (6.4)$$

so, $T = 0.015 \mu\text{m}$.

From equation (7.23) $T = 2\sigma^2/\Gamma$,

so the corresponding outer scale size is $22 \mu\text{m}$.

A more detailed discussion of the possible reasons for the kink at the limit of this linear region is given in section 6.2 vi).

6.2 iv) Plate 200

The structure functions for the x and y directions are shown in figure 6.8. The solid curve is for the x direction, the dashed one is for the y direction.

For the linear regions near zero lag, the slopes and, from equation (2.21), the topothesis are

$$T_x = 0.0152 \pm 0.0002 \mu\text{m}$$

$$T_y = 0.00886 \pm 0.00006 \mu\text{m}$$

Note that the slopes and the errors in the slopes have been calculated using a least squares analysis. The error evaluation took into account only the deviation of each point from the fitted line; it did not allow for the uncertainties in the position of each point.

The corresponding outer scale sizes, given by equation (7.23) $\Gamma = 2\sigma^2/T$ are

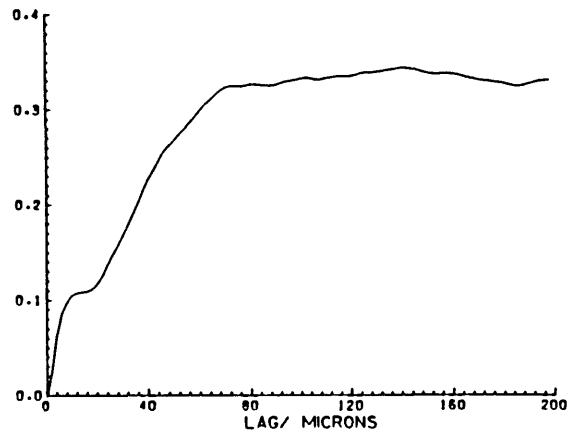


Figure 6.7. Structure function for plate 141, in x direction.

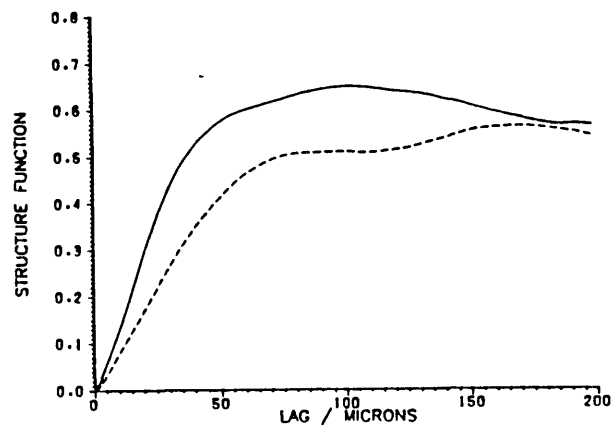


Figure 6.8. Structure functions for plate 200, x direction solid, y direction dashed.

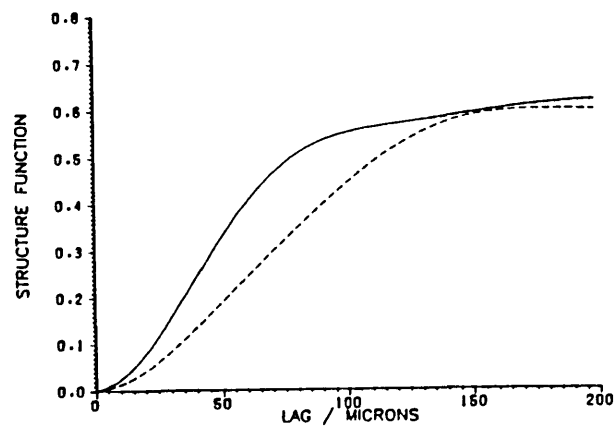


Figure 6.9. Structure functions for plate 201, x direction solid, y direction dashed.

$$\Gamma_x = 38 \mu\text{m}$$

$$\Gamma_y = 63 \mu\text{m}$$

6.2 iv) Plate 201

The structure functions for the x and y directions are shown in figure 6.9. The solid curve is for the x direction, the dashed curve is for the y direction.

These curves differ from those for the multiscale surfaces, as they do not exhibit a linear portion near to zero lag.

6.2 v) Power spectrum

6.2 v) Plate 141

The power spectrum for surface 141 is shown plotted in log log form in figure 6.10.

The slope of the graph in figure 6.10, over the range of frequencies for which the power law should be obeyed, is -2.0 ± 0.1 , where the slope and error were both calculated using a least squares analysis.

Two observations can be made about the spectrum

- 1) The linear region is very well defined between the expected limits of spatial frequency.
- 2) There are, however, components in the spectrum with both lower and higher frequencies than those predicted simply by the summation of the triangular spectra of each exposure.
 - a) The maximum frequency expected due to the central peak shown in figure 5.2, is about 12000 cycles per metre. This component then, accounts for the spectrum at values of log frequency < 4 .

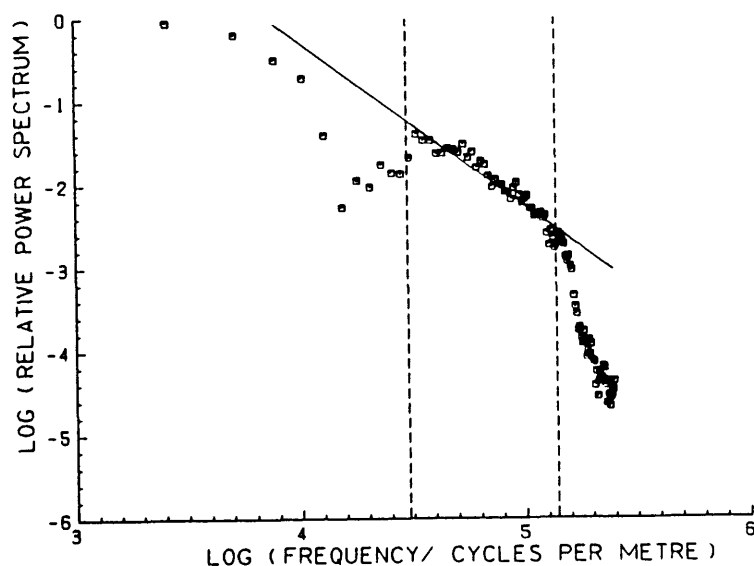


Figure 6.10. Log log plot of power spectrum for plate 141, in x direction.

- b) The lowest frequency triangle, occurring at a frequency $d/\lambda z$ where $d = 5\text{mm}$ will extend in low frequency to $(d-w)/\lambda z$, i.e. 18000 per metre. This half triangle can thus account for the spectrum between approximate values of log frequency 4.2 and 4.5.
- c) The high frequency tail appears to extend over a similar frequency range as region b), but this is a distortion of the logarithmic scale. It in fact extends to the Nyquist frequency of 250000 cycles per metre. The values of the spectrum are very low however, and can be thought of as arising from the numerical calculation.

A Talysurf trace was made of a smooth unexposed surface. The spectrum of such a surface could reasonably be expected to contain only low frequencies, any high frequencies occurring would indicate an instrumental response or analysis effect. In fact an apparent spectrum occurred at all frequencies. As only one trace was made it is probable that the effects would be modified were several traces made and a mean spectrum found. Even so, it was found that the background spectrum was small enough to have no effect on the measured slope of the spectrum.

6.2 v) Plate 200

The power spectra for plate 200 in the x and y directions are shown plotted in log log form in figure 6.11 i) and ii). Slopes of these plots were worked out over the frequency range from 15170 to 200000 m^{-1} rather than the expected range of multiscale behaviour from 15170 to 379200 m^{-1} , because of the limitations set by the size of the Talysurf stylus as discussed at the start of section 6.2.

Slope in x direction = -2.30 ± 0.04

Slope in y direction = -2.50 ± 0.06

6.2 v) Plate 201

The power spectra for plate 201 in the x and y directions are shown in figure 6.12 i) and ii), again plotted in log log form. A linear region is not expected in these plots as the power spectra should be Gaussian. From figure 6.12 it is clear that linear behaviour is seen at frequencies greater than 16000 m^{-1} . However, the strength of the variations contributing to the spectrum at these frequencies is small, the predominant spectrum is Gaussian.

6.2 vi) Overall Comments

6.2 vi) Plate 141

The surface produced has an almost Gaussian probability density function of surface height, and a variance of $0.17 \pm 0.01 \mu\text{m}^2$. The power spectrum has a power law region between spatial frequencies 30000 and 138000 m^{-1} , with an index of -2.0 ± 0.1 . In these respects the surface fits the requirements of the model.

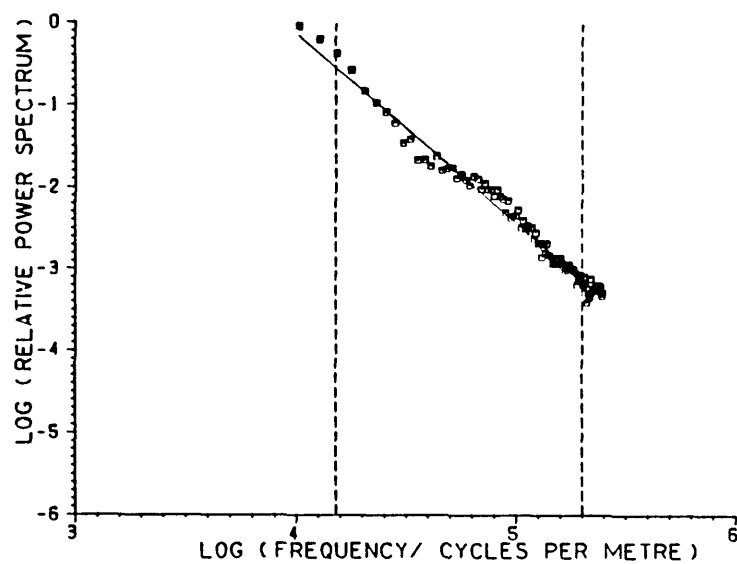


Figure 6.11 i). Log log plot of power spectrum for plate 200 in x direction.

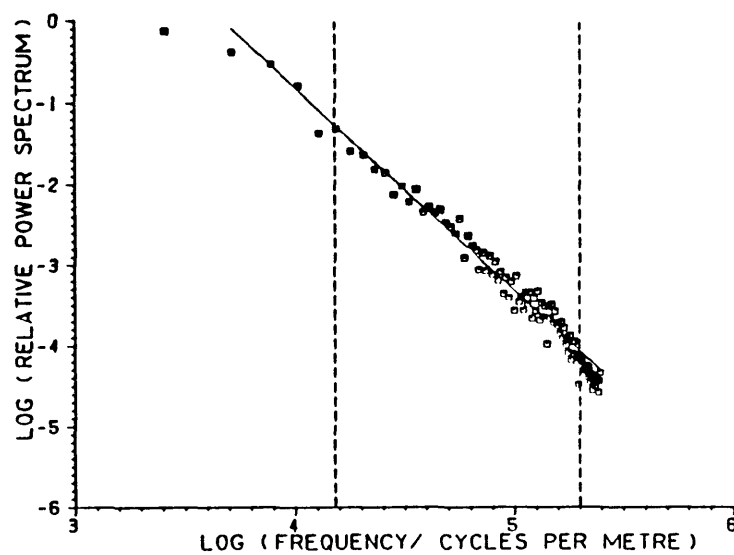


Figure 6.11 ii). Log log plot of power spectrum for plate 200 in y direction.

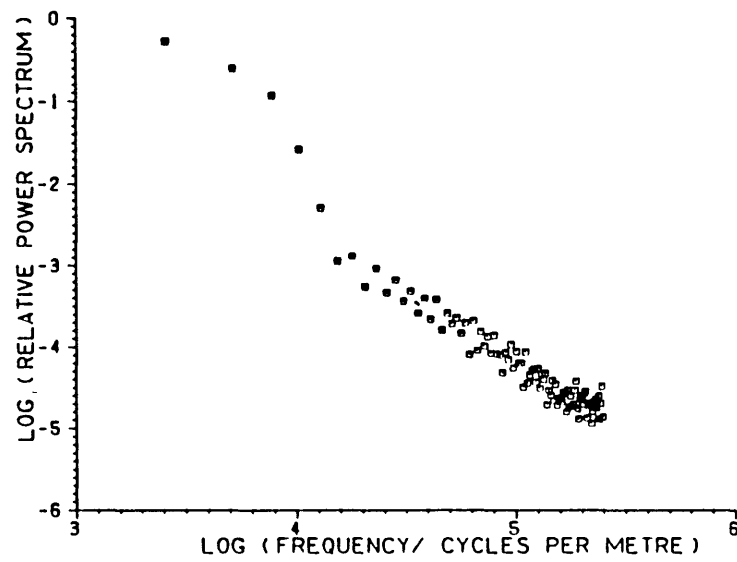


Figure 6.12 i). Log log plot of power spectrum for plate 201 in x direction.

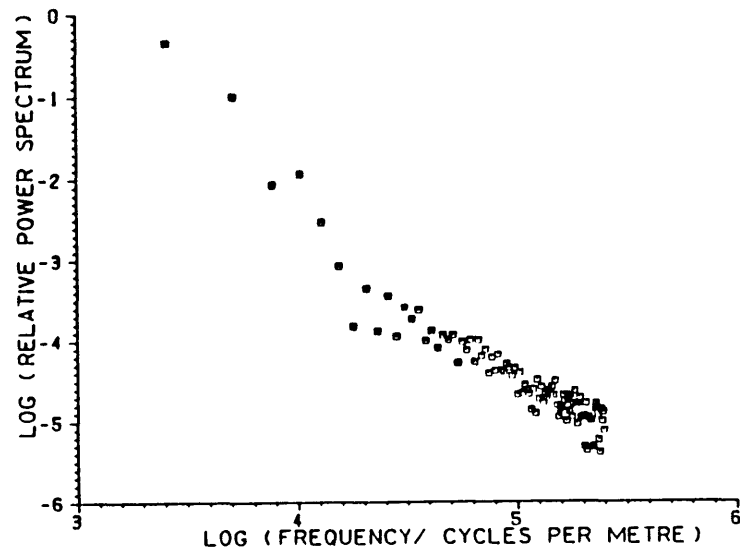


Figure 6.12 ii). Log log plot of power spectrum for plate 201 in y direction.

The forms of both structure and autocorrelation function are unusual. The kink in the curves at small lag means that the parts of the curve described by the condition $|X| \ll \Gamma$ are discontinuous with the main portion of the function. It is concluded that the range of frequencies over which the power law is valid is too small to allow the use of the small lag approximations required by equations (2.4) and (2.6). It seems likely that light scattering from the surface may be dominated by effects other than multiscale, as were the structure and autocorrelation functions. The power spectrum falls by only a factor of 10 between the inner and outer scale size limits. Thus the small scale variations are comparable with the large scale ones and possibly with the wavelength. Jakeman (1982) points out that strong fluctuation at the inner scale size can cause smooth, single scale surface models to be more appropriate in the scattering context. Scattering measurements are presented in section 7.4 iii).

6.2 vi) Plate 200

This surface has a probability density function close to Gaussian in form. Its overall variance is $0.29 \pm 0.01 \mu\text{m}^2$. The power spectrum has a power law region over a range of spatial frequencies between about 15000 and 250000 m^{-1} at least. The resolution of the Talysurf with a $1\mu\text{m}$ stylus is such that information on higher frequencies cannot be obtained (section 6.2). The index of the power law in the x direction differs by 15% from that planned (i.e. -2), in the y direction the difference is 25%. The deviation due to the spread of points from the least squares fitted line is, at most, 2.4%.

The structure function and autocorrelation function are of a much more satisfactory form than those for plate 141. Both have linear regions for small lag. Again the slopes differ in the x and y directions. As suggested in section 6.2 ii), for plate 200, this is likely to be due to a non-circular illuminated area of diffuser. Measurements of the laser beam profiles in orthogonal directions show that the ratio of horizontal to vertical width (measured at the $1/e$ points) is 1.25 ± 0.04 . The corresponding ratio from the autocorrelation function and structure function is 1.72.

The power spectra fall by a factor of about 1000 between the inner and outer scale sizes. This suggests that scattering may be dominated by multiscale effects.

6.2 vi) Plate 201

This surface also has a probability density function close to Gaussian. The overall variance of the surface is $0.34 \pm 0.02 \mu\text{m}^2$. The power spectra for the x and y directions are expected to be Gaussian also, but have a power law region for frequencies greater than 16000 m^{-1} . However, as the value of the spectrum is low relative to the value at zero frequency, this feature is less important than the low frequency part of the curve.

The structure and autocorrelation functions display the same anisotropy as was discussed in the previous section. The autocorrelation functions are good approximations to Gaussian functions.

6.3 Notes on the experimental determination of correlation functions

This topic has been covered in some detail by O'Neill and Walther (1977; Freniere et al 1977). They show that if the autocorrelation function is calculated by using sample heights each with the sample mean subtracted, then

$$\int_0^{X_{\max}} C'(X) dX = 0 \quad (6.5)$$

where $C'(X)$ is the calculated autocorrelation function, and $X_{\max}$ is the maximum lag it is possible to use for the data set.

From equation (6.5), the function $C'(X)$ must change sign at least once between 0 and $X_{\max}$, in order that the total area beneath the curve is zero. As $X_{\max}$ is made larger, the condition must still be met, but as the negative part of the curve must stretch over a

larger number of lags smaller negative values will be reached.

The difference between the calculated autocorrelation $C'(X)$ and the true function depends on the relation of $X_{\max}$ to the correlation width of the function. Armstrong and Gough (1978) quote an expression simpler than that in O'Neill and Walther (1977) relating the true and biased functions

$$C'(X) = (1 - X/X_{\max})C(X) + \int_0^{X_{\max}} K'(\sigma, X)\phi(\sigma)d\sigma \quad (6.6)$$

where

$$K'(\sigma, X) = -2(1 - \sigma/X_{\max})(1 - X/X_{\max})/X_{\max}$$

and $C(X)$ is the true function.

Figure (6.13) shows an example calculated using equation (6.6). The correlation width is 8 lag units. The true curve is shown compared with a curve calculated using a maximum lag 40 times the correlation width, and one where the maximum lag is only 10 times the correlation width. Clearly the bias for the longer sample is much reduced.

The experimental autocorrelation functions presented in section 6.2 ii were calculated from ten Talysurf traces each of 1024 values. In figure 6.4 it can be seen that the correlation width of the function is around 50 microns, which corresponds to 25 lag units. So, the maximum lag in this case is approximately 40 times the correlation width, and the bias introduced by the limited sample length is small. The curve in figure 6.4 is the mean of ten correlation functions worked out for each individual trace, so the bias is further reduced.

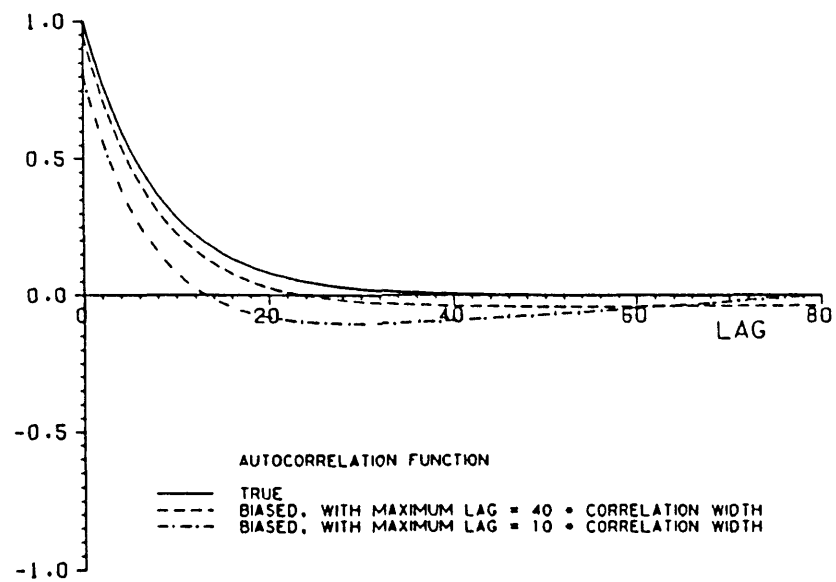


Figure 6.13. Comparison of autocorrelation functions for different maximum lags. True function (solid); biased with maximum lag 40 x correlation length (dashed); biased with maximum lag 10 x correlation length (dashed and dotted).

CHAPTER SEVEN

ESTIMATION OF SURFACE ROUGHNESS STATISTICS BY LIGHT SCATTERING

This chapter begins with a review of methods which can be used to estimate surface roughness statistics by light scattering. The standard deviation of surface height can usually be deduced, sometimes the probability density of surface heights, and the resemblance of the autocorrelation function to particular forms confirmed or otherwise.

Next, four similar approaches which can predict the angular distribution of mean intensity in the far field are discussed. The simplifying assumptions are listed, and equations derived for the mean intensity distribution from two forms of surface height autocorrelation function.

The experimental technique used is described, results presented, and their comparison with theoretical predictions discussed.

7.1 Scattering theory

Light scattering is a subject which, if treated rigorously is extremely complicated. In many practical cases the problem can be considerably simplified by the application of assumptions and approximations.

Many treatments of elastic light scattering in optics begin with diffraction theory. The complex amplitude of an electromagnetic wave is made up of vector electric and magnetic fields. The coupling between these fields is described by Maxwell's equations. A vector scattering theory must take account of the behaviour of these fields when they are near to charges or current. Clearly this would include occasions when the wave is close to the diffracting screen.

Considerable simplification can be obtained by considering the wave only at times when the complex amplitude can be considered as the Cartesian component of either field vector- that is, as a scalar. The theory is then known as scalar diffraction theory. Scalar theory is accurate for an opaque screen containing a clear aperture when

- a) The diffracting aperture is large compared with the wavelength.
- b) The diffracted fields are observed far from the aperture.

In practice, these conditions are obeyed approximately in most situations with visible light.

7.2 Estimating statistics from light scattering

A number of methods have been proposed for estimating the surface statistics from scattered radiation. Reviews have been published by Welford (1977) and Chandley (1979). The methods can be broadly subdivided into those which use coherent illumination, and those using partially coherent illumination. In the former case further subdivision is possible between cases where the measurements are made in the far field and those in the image plane.

As Chandley (1979) points out, the method which involves measuring the ensemble averaged far-field scattered coherent light is relatively simple. The theory of this technique was developed first by Davies (1954). The work of Beckmann (and Spizzichino 1963; 1967) was found to be more accurate for strong scattering (Houchen and Hering 1967). Beckmann's theory was later reformulated, and extended to apply to transmissive scattering by Chandley and Welford (1975). The Beckmann and Chandley approaches are outlined in greater detail in sections 7.3 and 7.4.

The first experimental estimate of surface roughness was made by Bennett and Porteus (1961) who used the theory of Davies (1954). Since then Beckmann's theory has been used by numerous authors to estimate the standard deviation of surface height, and correlation length (by assuming a Gaussian autocorrelation function of surface heights). See, for example, Hensler (1972). Chandley's work (1979) has shown that although the theory applies strictly only to those situations where the Kirchhoff theory is valid, measurements outside this region still agree with the theory.

It was pointed out in chapter four that a fully developed far field speckle pattern carries no information about the scattering surface height statistics. Hence the Beckmann theory is used in cases of weak scatter. However, if the size of the illuminated area is reduced until only a small number of scatterers is illuminated, then the speckle becomes non-Gaussian and carries information about the surface even if the scattering is strong (Jakeman and Pusey 1975). In order to obtain values for the mean square phase deviation and correlation length measurements must be made of the ratio of the second moment to the square of the mean far field intensity. Jakeman and McWhirter (1977) have predicted the distribution of mean intensity with angle for different surface height autocorrelation functions. This approach will be discussed in section 7.5. Non-Gaussian speckle is today the subject of much interest (Jakeman 1980), but it is not yet well enough understood to be used as a measuring tool.

If image plane speckle is used, an estimate of the standard deviation of surface heights can be obtained, for weak scatterers, by measuring the speckle contrast. The speckle contrast is defined by equation (4.21), and is the ratio of the standard deviation of the intensity to the mean intensity. Chandley (1979) emphasises that although good results have been obtained by such methods (Ohtsubo and Asakura 1975), the experimental technique is slightly more complicated than that based on Beckmann theory, and the speckle statistics need to be measured to find the standard deviation of surface heights.

The mean intensity at only three positions in the far field pattern need be measured to give an estimate of this parameter using Chandley's approach.

Spatially partially coherent illumination is derived from an extended monochromatic source. Temporally partially coherent, or polychromatic, illumination arises when two or more slightly different wavelengths are used together.

Spatially partially coherent light can be used to produce image plane speckle, and the standard deviation of surface height found from the speckle contrast as already suggested for coherent light. However, as the speckle contrast is reduced, there seems little advantage in not using coherent light when in the laboratory.

Polychromatic light can also be used. The intensity at any point in a far field speckle pattern depends on the vertical scale of the scatterer, and on the wavelength of the light. Thus, if the two wavelengths illuminating a surface are very far apart, completely different far field patterns result. If they are very close then the two patterns may well be identical. For a pair of wavelengths between these categories the far field patterns will be correlated to a certain degree (Parry 1984). The contrast of the overall pattern, for a particular probability density, must depend on the standard deviation of surface height.

A number of techniques based on this principle have been described (Parry 1974, George and Jain 1974). Because a far field polychromatic speckle pattern is spatially non-stationary except at the centre of the far field, measurements must be restricted to this region. These methods yield the standard deviation of surface height and can give the probability density. Parry (1975) reported measurements agreeing fairly well with the theory. More recent experiments utilizing polychromatic speckle have been reported (Stansberg 1979, Giglio et al 1979, Nakagawa and Asakura 1978). An advantage of these methods is that standard deviations of more than one wavelength can be measured.

There are clearly several ways of estimating the standard deviation of surface height using scattered light. The methods may now be grouped according to the range of surface roughness for which they are most appropriate. For weak scatter the method of Chandley (1979), or coherent image plane speckle (Ohtsubo and Asakura 1975). For strong scatter, either non-Gaussian speckle (Jakeman and Pusey 1975), or polychromatic speckle (Parry 1975). However, successful experimental

techniques for the latter two have yet to be fully developed.

The purpose of the measurements made in this work is primarily to check that the surfaces have the required shape of autocorrelation function of surface height, although an estimate of the standard deviation is also of interest to compare with mechanical measurements. A method based on the Beckmann and Chandley theory seems most applicable, as measurements of the mean intensity of the far field speckle envelope intensity can be made fairly simply. Different forms for the autocorrelation function can be incorporated to predict the form of the envelope.

7.3 Prediction of angular distribution of mean intensity

Four approaches to the prediction of the angular distribution of mean intensity in the far field are considered. It transpires that although the methods initially appear to differ, in fact they are equivalent within the limits of scalar diffraction, and lead to the same result at near normal incidence. The four approaches have been described by Beckmann (and Spizzichino 1963; 1967), Chandley (1979; Chandley & Welford 1975), Goodman (1984) and Jakeman (Jakeman & Pusey 1973; Jakeman & McWhirter 1977).

7.3 i) The thin phase screen model

A thin phase screen is one which occupies a very narrow region, and introduces a random distribution of phase changes in the incident wave. To predict the far field scattering then, the Kirchhoff diffraction integral is applied across a plane containing the wave scattered from the screen. This is the approach used by Chandley (1979), Goodman (1984), and Jakeman (Jakeman & Pusey 1973; Jakeman & McWhirter 1977). The last two do so implicitly. Use of the thin phase screen means that the Kirchhoff diffraction integral is evaluated over a plane which includes the random phase changes introduced by the thin phase screen. If the wavefront has gentle slopes, then the angles of incidence and scattering will not vary appreciably over the plane, and so no inclination factor need be

included in the integral.

Beckmann does not use the thin phase screen model. Instead he applies the Kirchhoff diffraction integral directly to the random rough surface. In this case the direction of the normal to the surface of integration changes, depending on position. Thus the angle of incidence of light with the surface also changes. In order to be able to take the inclination factor outside the integral, Beckmann assumes perfect conductivity of the rough surface, so that the reflectivity is unity whatever the angle of incidence. In practice, perfect conductivity is unlikely, but constant reflectivity over the surface can be obtained by dividing results by measurements from an identical, but plane, surface - so long as the surface slopes are small. In effect then, it doesn't matter whether the thin phase screen is used or not; providing the surface slopes are gentle the inclination factor may be ignored.

7.3 ii) Assumptions

In each case the basic assumptions used are very similar.

Assumption one

Scalar diffraction theory is used.

Assumption two

The Kirchhoff diffraction integral is applied, so the predictions are strictly valid only for small angles of scatter. If the surface has low slopes then there will be no obliquity factor within the integral, and the angles of scatter will also be small. A quantitative expression defining low slopes is given by Nieto-Vesperinas & Garcia (1981) as

$$\sigma/L < 0.05 \quad (7.1)$$

where σ is the standard deviation of the surface heights, and L the correlation length.

Assumption three

No multiple scattering or shadowing effects.

Assumption four

No depolarization of scattered light.

Assumption five

No variations of reflectivity or transmissivity with angle. Similarly, the phase change on reflection changes negligibly with change in incident angle.

Assumption six

Random rough surface is isotropic, and has stationary statistics.

Further assumptions are introduced to facilitate the solution of the integral.

Assumption seven

Chandley, Beckmann, and Goodman assume that the incident wave is plane and coherent. Jakeman takes it to have a Gaussian amplitude profile.

Assumption eight

In all cases the first and second order probability densities of surface height, and of the phase of the scattered wave, are Gaussian.

Assumption nine

The illuminated area is generally taken to be greater than the correlation area of the surface. Jakeman does not make this assumption when considering non-Gaussian scattering.

Assumption ten

A particular form is chosen for the normalised autocorrelation function of the surface heights, or phase fluctuations.

The equations derived by Chandley (1979) will now be considered.

7.3 iii) The Chandley equations

The general expression given by Chandley (1979), for the normalised ensemble averaged far field intensity, where ξ, η are the far field co-ordinates, is

$$\begin{aligned} \langle I(\xi, \eta) \rangle = F^2 \left[\frac{|\langle V \rangle|^2}{A^2} K(\xi, \eta) + \right. \\ \left. \frac{2\pi}{SA^2} \int_0^\infty \left[\langle V(x_1) V^*(x_1 - x) \rangle - |\langle V(x_1) \rangle|^2 \right] \right. \\ \left. J_0 \left\{ \frac{kx}{f} (\xi^2 \cos^2 \theta_1 + \eta^2)^{\frac{1}{2}} \right\} x dx \right] \end{aligned} \quad (7.2)$$

where F is the Kirchhoff inclination factor. Chandley (1979) gives two alternative expressions for F , depending on the position of the limiting aperture. For small angles of incidence and scatter, these reduce to $F=1$, corresponding to the Beckmann (and Spizzichino 1963; 1967) result for normal incidence.

$$k=2\pi/\lambda$$

V is the near field complex amplitude

S is the area of illumination

θ_1 is the angle of incidence with the mean plane of the surface

θ_2 is the scattering angle

$K(\xi, \eta)$ is the normalised intensity point spread function of the limiting aperture and field lens

f is the focal length of the field lens

If assumption eight is applied, i.e. that the rough surface has Gaussian first and second order probability density functions, then

$$\langle I(\xi, \eta) \rangle = F^2 \exp \{-g_c\} \left[K(\xi, \eta) + \frac{2\pi}{S} \int_0^\infty \left[\exp \{g_c C_{\text{NORM}}(X)\} - 1 \right] \right] \quad (7.3)$$

$$J_0 \left\{ \frac{kX}{F} (\xi^2 \cos^2 \theta_1 + \eta^2)^{\frac{1}{2}} \right\} X dX$$

where

$$k = 2\pi/\lambda$$

$$g_c = (ku\sigma_h)^2 \quad (7.4)$$

and

$$u = 2\cos\theta_1 \quad \text{for reflective scattering} \quad (7.5)$$

$$u = (n_p \cos\theta_1' - n_m \cos\theta_1) \quad \text{for transmissive scattering} \quad (7.6)$$

n_p , n_m are the refractive indices of photoresist and the surrounding medium respectively.

θ_1' is the angle of refraction of light within photoresist.

Let the incident light be normal to the surface

$$\theta_1 = \theta_1' = 0,$$

the angle of scatter $\theta_2 = \theta$

and, $F = 1$.

If the normalised autocorrelation function of surface height is chosen to be Gaussian, given by

$$C_{\text{NORM}}(X) = \exp(-X^2/L^2) \quad (7.7)$$

For strong scatter, $g_c \gg 1$ and the first term in equation (7.3) may be neglected. In addition, only the behaviour of $C_{\text{NORM}}(X)$ near the origin is important, so equation (7.7) may be written

$$C_{\text{NORM}}(X) = 1 - (X/L)^2 \quad (7.8)$$

and, as $X \ll L$,

$$\exp(-g_c) \ll \exp -g(X/L)^2$$

So equation (7.3) reduces to

$$\langle I(\xi, \eta) \rangle = \frac{2\pi}{S} \int_0^\infty \exp \left\{ -g \left[\frac{X}{L} \right]^2 \right\} J_0 \left\{ \frac{kX}{f} (\xi^2 \cos^2 \theta_1 + \eta^2)^{\frac{1}{2}} \right\} X dX \quad (7.9)$$

From Gradshteyn and Ryzhik (p717, 1980)

$$\int_0^\infty e^{-\alpha x^2} J_0(\beta x) x dx = \frac{1}{2\alpha} \exp \left[\frac{-\beta^2}{4\alpha} \right] \quad (7.10)$$

thus

$$\langle I(\xi, \eta) \rangle =$$

$$\frac{\pi^2 L^2}{S g_c} \exp \left\{ -k^2 L^2 (\xi^2 + \eta^2) / 4 g_c f^2 \right\} \quad (7.11)$$

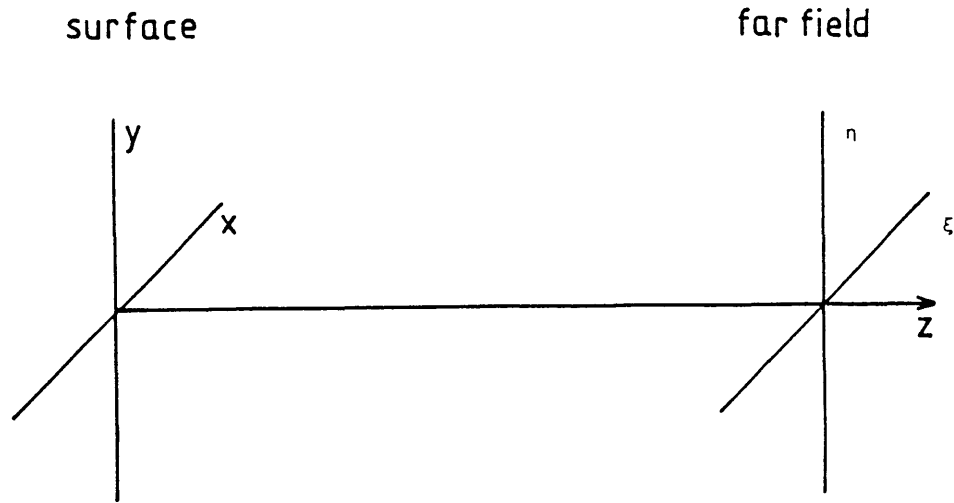


Figure 7.1. Co-ordinate system for scattering measurements.

If measurements are made in the ξz plane, then $\eta = 0$, and $\sin\theta = \xi/f$ (the co-ordinate system is shown in figure 7.1).

$$\langle I(\theta) \rangle = \frac{\pi^2 L^2}{Sg_c} \exp \left\{ -\frac{k^2 L^2 \sin^2 \theta}{4g_c} \right\} \quad (7.12)$$

and similarly for measurements in the ηz plane.

7.3 iv) Comparison with other approaches

It is emphasised that this expression (7.12) is identical to that given by Beckmann (and Spizzichino 1963; 1967), for the small angles where both are valid.

In addition, note that Goodman (1984), implicitly using a thin phase screen model, derives a relation between the autocorrelation of the complex fields scattered by a surface, and the normalised autocorrelation function of surface heights.

$$C_{FNORM}(\Delta x, \Delta y) = \exp \left\{ - \sigma_{\phi}^2 \left[1 - C_{HNORM}(\Delta x, \Delta y) \right] \right\} \quad (7.13)$$

where σ_{ϕ}^2 is the variance of the phase

$C_{FNORM}(\Delta x, \Delta y)$ is the normalised autocorrelation function of the complex fields.

$C_{HNORM}(\Delta x, \Delta y)$ is the normalised autocorrelation function of surface heights.

Now, in the far field, the amplitude of the field is given by the Fourier transform of the reflected fields,

$$F \left[\alpha(x, y) \right] \quad (7.14)$$

where F denotes Fourier transformation

Thus, the far field mean intensity is given by

$$\langle I(\xi, \eta) \rangle = \langle |F \left[\alpha(x, y) \right]|^2 \rangle \quad (7.15)$$

But expression (7.15) is the power spectrum of the field $\alpha(x, y)$, which can also be expressed as

$$F \left[C_{FNORM}(\Delta x, \Delta y) \right]$$

therefore

$$\langle I(\xi, \eta) \rangle = F \left[C_{FNORM}(\Delta x, \Delta y) \right] \quad (7.16)$$

where $C_{FNORM}(\Delta x, \Delta y)$ is given by equation (7.13).

If it is again assumed that only the behaviour of C_{HNORM} near the the origin is important, then, for a Gaussian function only the first terms of the expansion are required, and equation (7.13) becomes

$$C_{\text{FNORM}}(\Delta x, \Delta y) = \exp \{-\sigma_{\phi}^2 (X/L)^2\} \quad (7.17)$$

Fourier transformation of this expression, assuming normal incidence, and measurement in the ξz plane, (see figure 7.1) yields once again the same angular dependence of mean intensity as Beckmann and Chandley. Goodman's approach, however avoids the inclusion of factors relating to the absolute value of scattered intensity.

Jakeman's approach is more general, although the screen is again assumed to be deep, so that scatter is strong, as he retains terms relating to the size of the scattering area in order to allow predictions about non-Gaussian scattering - speckle which does not belong to the fully developed class described in chapter four. Instead, because only a few scattering centres are illuminated, or if the detector is so close to the screen that contributions to the backscattered field come from only a few scatterers, the central limit theorem cannot be applied. Non-Gaussian scattering is, however, beyond the scope of this thesis. When the same simplifying assumptions are applied as before, the results reduce to give, once again, the same predictions.

7.3 v) Applications in this work

The equation (7.12) should apply to the single scale Gaussian scatterers made by the method of Gray, and to the one dimensional multiscale surface in the y direction.

Clearly, in order to predict the scattering from the multiscale surfaces produced a different form of normalised autocorrelation function of surface height (or phase) must be assumed. A suitable form has been suggested by Jakeman (1980; Jordan, Hollins and Jakeman 1983).

7.3 vi) Incorporation of the multiscale model into Chandley's equations

Jakeman has pointed out (1980; Jordan, Hollins and Jakeman 1983) that the use of a truncated linear phase autocorrelation function, i.e.

$$C_{\phi}(X) = \begin{cases} 1 - \frac{|X|}{\Gamma} & |X| < \Gamma \\ 0 & \text{otherwise} \end{cases} \quad (7.18)$$

is equivalent to the use of a power law structure function model. It was shown in the appendix that for a power law spectrum, where

$$P(u) = A / |u|^{\alpha} \quad (2.1)$$

the structure function $D(X)$, is given by

$$D(X) = T^{3-\alpha} |X|^{\alpha-1} \quad (2.2)$$

i.e. for $\alpha = 2$

$$D(X) = T |X| \quad (7.19)$$

for a power law spectrum where T is the topothesy (Berry 1979), and α is related to the Hausdorff-Besicovitch dimension, D , (section 2.5) by

$$\alpha = 5 - 2D \quad (2.19)$$

If the normalised autocorrelation function is given by

$$C_{\text{NORM}}(X) = \begin{cases} 1 - \frac{|X|}{\Gamma} & |X| < \Gamma \\ 0 & \text{otherwise} \end{cases} \quad (7.20)$$

then, for $X < \Gamma$, the structure function can be written, from equation (1.63)

$$D(X) = 2\sigma_h^2 (1 - C_{\text{NORM}}(X)) \quad (1.63)$$

where $D(X)$, $C_{\text{NORM}}(X)$ are the structure function and normalised autocorrelation function respectively of surface heights, rather than phase.

So,

$$D(X) = 2\sigma_h^2 (1 - (1 - |X|/\Gamma)) \quad (7.21)$$

$$D(X) = 2\sigma_h^2 |X|/\Gamma \quad (7.22)$$

$$\text{If } T = 2\sigma_h^2/\Gamma, \text{ then} \quad (7.23)$$

$$D(X) = T|X| \quad (7.19)$$

and the structure function has the form of that for a power law spectrum, of index $\alpha=2$.

Let us substitute expression (7.20) into equation (7.3)

$$\begin{aligned} \langle I(\xi, \eta) \rangle = F^2 \exp(-g_c) & \left[K(\xi, \eta) + \frac{2\pi}{S} \int_0^\infty \left[\exp\{-g_c C_{\text{NORM}}(X)\} - 1 \right] \right. \\ & \left. J_0 \left\{ \frac{kX}{f} (\xi^2 \cos^2 \theta_1 + \eta^2)^{\frac{1}{2}} \right\} X dX \right] \end{aligned} \quad (7.3)$$

$$= F^2 \left[K(\xi, \eta) \exp(-g_c) + \frac{2\pi}{S} \int_0^\infty \left[\exp \left\{ -g_c \frac{|X|}{\Gamma} \right\} - \exp(-g_c) \right] J_0 \left\{ \frac{kX}{f} (\xi^2 \cos^2 \theta_1 + \eta^2)^{\frac{1}{2}} \right\} X dX \right] \quad (7.24)$$

As in section 7.3 iii) for the Gaussian autocorrelation function for strong scatter, $g_c \gg 1$, so ^{we} may neglect the first term.

$$\langle I(\xi, \eta) \rangle = F^2 \frac{2\pi}{S} \int_0^\infty \left[\exp(-g_c \frac{|X|}{\Gamma}) - \exp(-g_c) \right] J_0 \left\{ \frac{kX}{f} (\xi^2 \cos^2 \theta_1 + \eta^2)^{\frac{1}{2}} \right\} X dX \quad (7.25)$$

Also, $\Gamma \gg X$, so

$$\exp(-g_c) \ll \exp(-g_c \frac{|X|}{\Gamma})$$

thus

$$\langle I(\xi, \eta) \rangle = F^2 \frac{2\pi}{S} \int_0^\infty \left[\exp(-g_c \frac{|X|}{\Gamma}) \right] J_0 \left\{ \frac{kX}{f} (\xi^2 \cos^2 \theta_1 + \eta^2)^{\frac{1}{2}} \right\} X dX \quad (7.26)$$

Now, from Erdelyi et al (1954, p182)

$$\int_0^{\infty} e^{-pt} J_0(\alpha t) t dt = \frac{p}{(p^2 + \alpha^2)^{3/2}} \quad (7.27)$$

Thus,

$$\langle I(\xi, \eta) \rangle = \frac{F^2 2\pi g_c}{S\Gamma} \left[\frac{g_c^2}{\Gamma^2} + \frac{k^2}{f^2} (\xi^2 \cos^2 \theta_1 + \eta^2) \right]^{-3/2} \quad (7.28)$$

For normal incidence $\theta_1 = 0$, and let $F = 1$

$$\langle I(\xi, \eta) \rangle = \frac{2\pi g_c}{2\Gamma} \left[\frac{g_c^2}{\Gamma^2} + \frac{k^2}{f^2} (\xi^2 + \eta^2) \right]^{-3/2}$$

Now, in practice measurements are made in the ξz plane, see figure 7.1, i.e. $\eta = 0$, $\sin \theta = \xi/f$

$$\langle I(\theta) \rangle = \frac{2\pi g_c}{S\Gamma} \left[\frac{g_c^2}{\Gamma^2} + k^2 \sin^2 \theta \right]^{-3/2} \quad (7.29)$$

and similarly for measurements in the ηz plane.
where

$$g_c = \left[2k\sigma \right]^2 \quad \text{for reflection} \quad (7.5)$$

$$g_c = \left[k(n_p - n_m)\sigma \right]^2 \quad \text{for transmission} \quad (7.6)$$

This result may be compared with the prediction of Jordan, Hollins and Jakeman (equation 8, 1983), by substituting for g_c , noting that the topography is given by

$$T = 2\sigma_h^2/\Gamma \quad (7.23)$$

Thus, in transmission

$$g_c = k^2(n_p - n_m)^2 \Gamma T/2 \quad (7.30)$$

$$\langle I(\theta) \rangle = \frac{\pi T}{Sk} (n_p - n_m)^2 \left[\frac{T^2 k^2 (n_p - n_m)^4}{4} + \sin^2 \theta \right]^{-3/2} \quad (7.31)$$

i.e.

$$\langle I(\theta) \rangle \propto \left[\frac{k^2 T^2 (n_p - n_m)^4}{4} + \sin^2 \theta \right]^{-3/2} \quad (7.32)$$

which corresponds exactly to equation 8 in Jordan et al (Jordan, Hollins and Jakeman 1983).

7.3 vii) Applications in this work

This expression will be compared with measurements made from both the one and the two dimensional multiscale surfaces produced. It should be noted that the equation has been derived assuming small angle diffraction theory, though in the expression (7.29) multiscale behaviour is not expected to be obvious until larger angles of scatter.

In addition, assumption two (section 7.3 ii) cannot be applied to a multiscale surface. A fractal set was defined in section 2.5 as being non differentiable. This prevents the use of the assumption of low slopes, since the slope is not defined. Instead, it has been suggested (Jordan, Hollins & Jakeman 1983) that the criterion will be based on the topothesy, and that Kirchhoff theory may be applied if

$$T \ll \lambda \quad (7.33)$$

The topothesy is a measure of the length over which the rms slope of a chord between two points on the surface is unity.

$$T = \sqrt{\langle [h(x) - h(x+T)]^2 \rangle} \quad (7.34)$$

It thus shows how quickly changes in surface height are taking place. T will be smaller for more rapid changes in surface height, and smaller still if these are of low magnitude. Thus, if $T \ll \lambda$ the surface height fluctuations of those elements smaller than size T will be of low magnitude, and the criterion is equivalent to assuming low slopes.

7.4 Experimental far field, coherent light scattering measurements

7.4 i) Experimental arrangements

Two alternative experimental arrangements were used. These are shown schematically in figures 7.2 a and b.

In each case, the aperture was used to define the area of illumination on the diffuser. This was generally about two millimetres in diameter, ensuring that the illuminated area was larger than the correlation area of the sample, but small enough to allow ten different areas of the diffuser to be used for measurements.

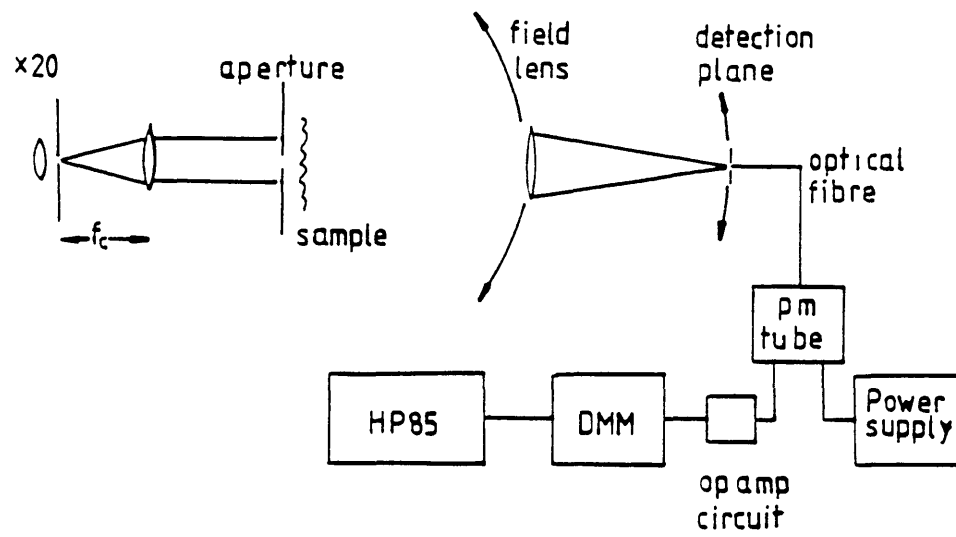


Figure 7.2 a). Schematic of initial experimental arrangement.

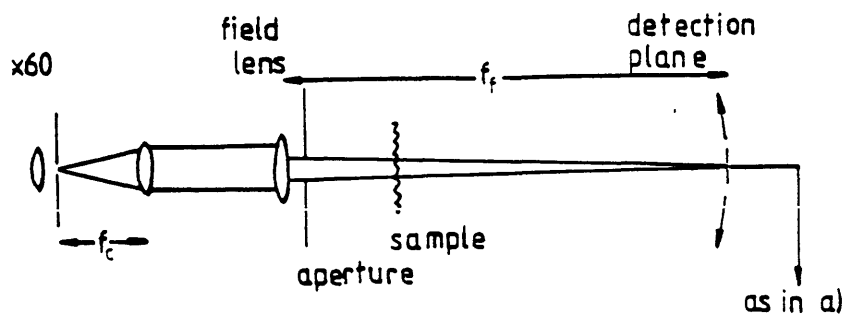


Figure 7.2 b). Schematic of experimental arrangement.

It is not vital in these measurements that the light incident on the diffuser in arrangement a) should be a plane wave, so long as both the radius of curvature and Gaussian beam width are large compared with the correlation length of the surface (Jakeman and Pusey 1975).

The choice of the expansion and collimation lenses, and the diameter of the limiting aperture in arrangement a) was such that the beam was plane to within $\pm 1\%$. From chapter 5B, the Gaussian beam diameter emerging from the collimating lens is given by

$$d = f_c d_b / f_{obj} \quad (5.14)$$

where d_b is the Gaussian beam diameter before expansion, f_c the focal length of the collimating lens, and f_{obj} that of the objective lens.

Now the beam profile may be described by

$$\exp \left[-(r/r_0)^2 \right] \quad (7.35)$$

where r is the radial co-ordinate, and r_0 the Gaussian beam half-width.

Across the limiting aperture, the beam should be uniform to within $\pm 1\%$, i.e.

$$\exp \left[-(r/r_0)^2 \right] \approx 0.99 \quad (7.36)$$

$$\exp \left[-(2rf_{obj}/f_c d_b)^2 \right] \approx 0.99 \quad (7.37)$$

if $f_{obj} = 8 \text{ mm}$

$f_c = 10 \text{ cm}$

$d_b = 1 \text{ mm}$

then

$r = 0.6 \text{ mm}$

Thus the limiting aperture should have a diameter of approximately 1.2 mm.

The rough surface was mounted on the table of a spectrometer, and could be moved using the table controls. In arrangement 7.2 a) the field lens was that of the telescope of the spectrometer, detection taking place in its focal plane. This arrangement was found to be suitable only for small scattering angles. At angles greater than about 5° reflections from the walls of the telescope rendered this method unsuitable. In arrangement 7.2 b) the field lens and aperture are placed before the diffuser. Changes in the angle of measurement were made simply by moving the telescope, and reading the angular position from the Vernier scale.

Averaging can be done in two ways. Firstly measurements at each angle can be made for several (typically ten) different positions of the scatterer. If readings are made at angular separations of ϕ , and if the distance from the plate is R , then the distance between reading points in the detection region is given by $R\phi$. Typically $R = 25$ cm, $\phi = 0.25^{\circ}$, leading to a separation of 1.09 mm. Thus for maximum coverage the detector aperture should be about 1 mm in diameter. Ideally the detector aperture should cover several speckles in the pattern. The size of each speckle depends on the area illuminated

$$\Sigma = 1.2\lambda R/d \quad (7.38)$$

where d is the diameter of the illuminated area.

Thus, for $R = 25$ cm, $\lambda = 633$ nm, and $d = 2$ mm, $\Sigma = 95$ μm . So, a detector aperture of 1mm gives adequate coverage.

Fibre Optic cable supplied by RS Components has a diameter of 1 mm. It was thus used both to define the area of detection, and to lead the light to the photomultiplier tube, which was permanently positioned and did not move with the spectrometer telescope.

Alternatively, a larger area may be illuminated so that a large number of speckles form over the detection area. The relative statistical uncertainty in one estimate of the mean intensity measured this way is given by (Dainty 1972b).

$$\sigma_I / \langle I \rangle = .61 \lambda R / dr \quad (7.39)$$

where d is diameter of illuminated area

R is the detector-diffuser distance

r is the radius of the circular detector

To obtain then an accuracy of 3%, and using the same 1mm detector area as before, $d = 6\text{mm}$.

The second method of averaging was chosen as it was very much quicker and easier.

The current generated in the photomultiplier tube was converted to voltage by a simple op-amp circuit, which has a linear response of .5 V per μA . Values of the voltage were displayed on a Keighley digital multimeter, and fed via an IEEE interface to the HP85 computer. A program was written for the HP85 which issues prompts requesting that the scatterer be changed, moved, or removed, or the scattering angle altered. The angles selected were input from the keyboard, and recorded on tape. A background count was made before each set of data was taken.

A second program was written to process the data once it had all been collected. It outputs the sine of the scattering angle, the mean voltage reading for this angle, and the standard deviation of this mean value. If the standard deviation of the mean voltage reading gives an error of less than that expected from equation (7.39), then equation (7.39) is applied.

7.4 ii) Equations to be applied

For both strong and weak scatter, equation (7.3) for the normalised speckle envelope intensity may be applied.

We have seen (section 7.3 iii) that if $C_{\text{NORM}}(X)$ is Gaussian, then for strong scatter

$$\langle I(\theta) \rangle \propto \exp \left\{ -\frac{k^2 L^2 \sin^2 \theta}{4g_c} \right\} \quad (7.12)$$

If $C_{\text{NORM}}(X)$ is truncated linear (section 7.3 w), then

$$\langle I(\theta) \rangle \propto \left[\frac{g_c^2}{\Gamma^2} + k^2 \sin^2 \theta \right]^{-3/2} \quad (7.29)$$

or, in terms of the topothesy

$$\langle I(\theta) \rangle \propto \left[\frac{k^2 T^2 (n_p - n_m)^4}{4} + \sin^2 \theta \right]^{-3/2} \quad (7.32)$$

From equation (7.12), a linear relation exists between $\log \langle I \rangle$ and $\sin^2 \theta$, with slope given by

$$L^2 \log_{10} e / 4 (n_p - n_m)^2 \sigma_n^2$$

Similarly, from equation (7.32), if

$$\sin^2 \theta \gg k^2 T^2 (n_p - n_m)^4 / 4$$

then

$$\langle I(\theta) \rangle \propto \left[\sin \theta \right]^{-3} \quad (7.40)$$

and a plot of $\log \langle I \rangle$ versus $\log (\sin \theta)$ has a slope of -3.

7.4 iii) Measurements from one-dimensional multiscale surfaces

Measurements of the mean intensity scattered with angle from plate 141 are presented in figures 7.3 and 7.4. Significant measurements above background could be obtained up to scattering angles of about 15° .

ξz plane

The curve shown in figure (7.3) is calculated from equation (7.32), using the value $T = 0.015\mu\text{m}$ which was obtained from the structure function in figure 6.7, and quoted in section 6.2 iv. It is clear from figure 7.3 that the scattering data does not agree with the multiscale prediction, with $T = 0.015\mu\text{m}$. Even if a different value for the topothesy were chosen (perhaps corresponding to the slope of the structure function at higher lags beyond the kink in figure 6.7) the curve would still have the same slope of -3 at high angles.

It was pointed out in section 6.2 vi) that since the small scale size fluctuations were only one tenth smaller than those at the outer scale, the scattering might not be multiscale in character. These results appear to bear this out.

ηz plane

No Talysurf measurements were made on the surface in the y direction, so no results corresponding to these scattering measurements appear in chapter 6.

As was mentioned in section 5A.1 i, the power spectrum of the surface heights in this direction should be a triangle of frequency width $2\ell/\lambda z$, where ℓ is the length of the slit. The autocorrelation function is not Gaussian.

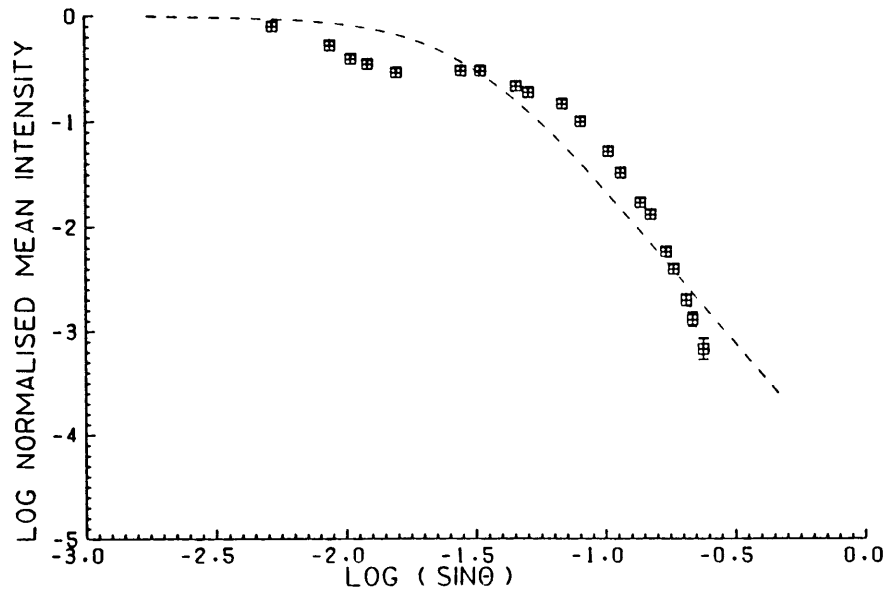


Figure 7.3. Measurements from plate 141 in the ξz plane compared with the multiscale prediction of equation (7.32) for $T = 0.015\mu\text{m}$.

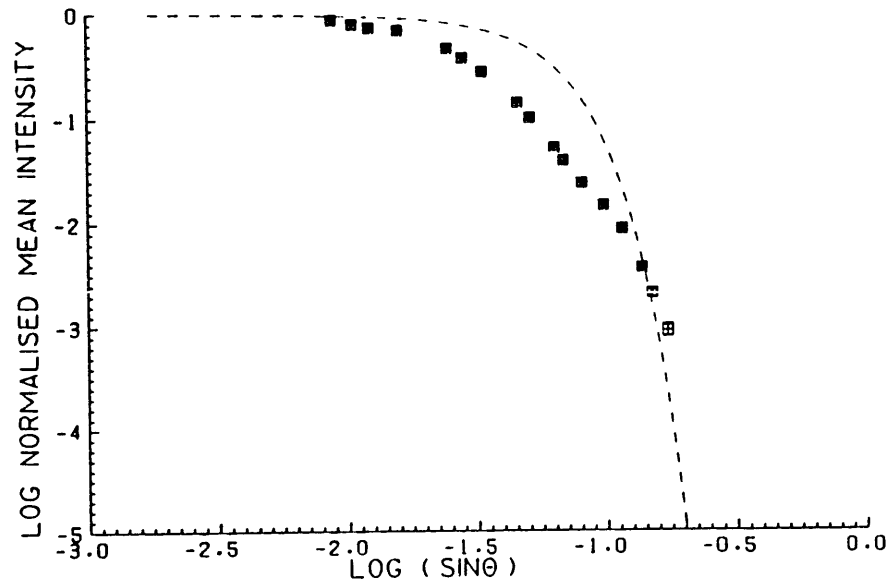


Figure 7.4. Measurements from plate 141 in the ηz plane compared with the single scale prediction of equation (7.12) for $L = 11\mu\text{m}$.

In figure 7.4 the measured values are compared with a curve predicted by equation (7.12) for a single scale surface with $L = 11\mu\text{m}$ which corresponds to a slit length of 1.5 cm. There is little agreement, which is not surprising as the autocorrelation function is not Gaussian. The plot is shown as a log log plot of the mean intensity versus $\sin\theta$, where a linear region is expected if the behaviour is multiscale.

7.4 iv) Measurements from two-dimensional multiscale surfaces

Measurements of the mean intensity scattered with angle from plate 200 are presented in figures 7.5 and 7.6. Significant readings above background could be obtained up to scattering angles of about 18° .

ξz plane

Three different sets of data are presented in figure 7.5, the curve is calculated from equation (7.32) with $T = 0.015\ \mu\text{m}$ which was obtained from the structure function in figure 6.8 and quoted in section 6.2 iv.

The agreement between the measured points and the curve is good, especially for scattering angles up to 12° . At larger angles less light is scattered than predicted by the theory.

ηz plane

Only two sets of data are presented in figure 7.6, as there was very much less variation between data sets in this case. The curve is calculated from equation (7.32), with $T = 0.009\ \mu\text{m}$, which was obtained from the structure function in figure 6.8 and quoted in section 6.2 iv.

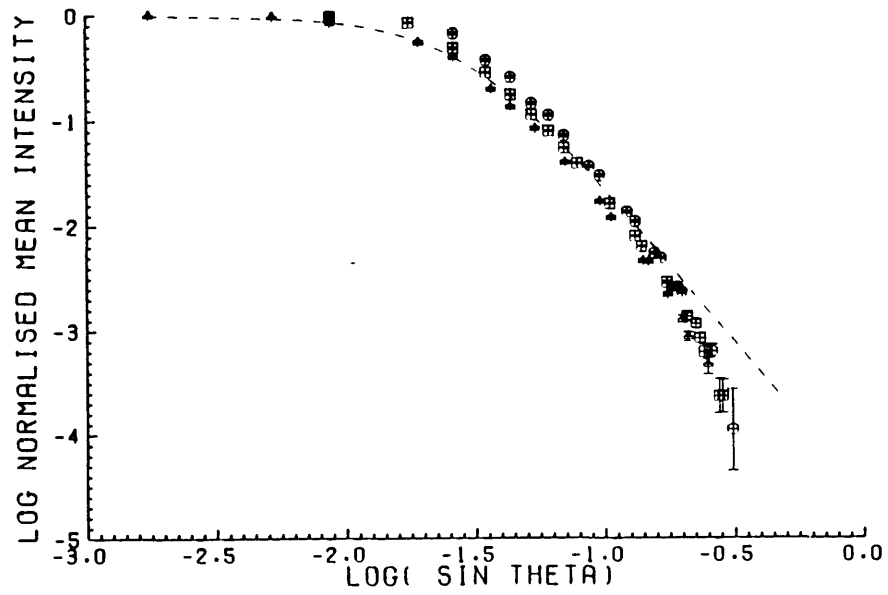


Figure 7.5. Measurements from plate 200 in the ξz plane compared with the multiscale prediction of equation (7.32) for $T = 0.015 \mu\text{m}$.

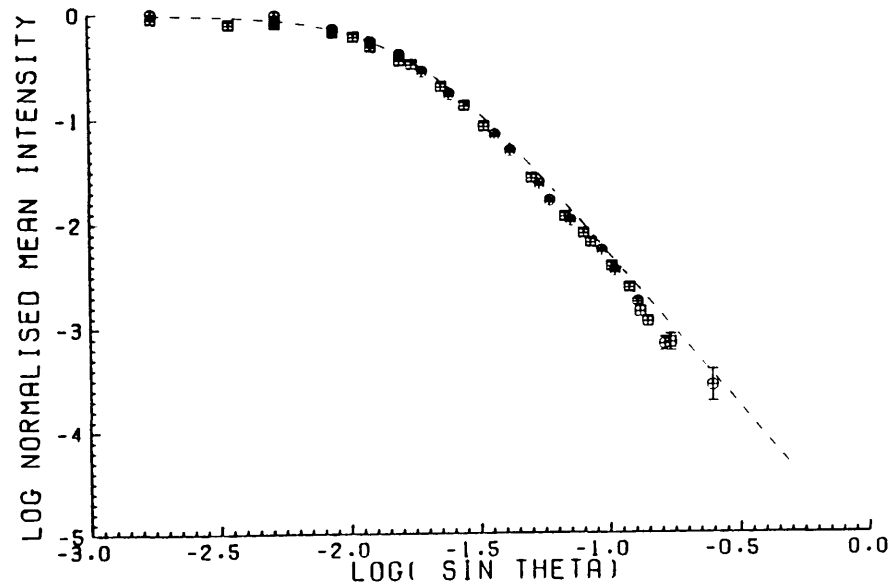


Figure 7.6. Measurements from plate 200 in the $n z$ plane compared with the multiscale prediction of equation (7.32) for $T = 0.0009 \mu\text{m}$.

The agreement between the measured points and the curve is good, even at the larger scattering angles where the agreement broke down for measurements in the ξz plane.

It seems likely that the broadening of the beam in the x direction which is responsible for the anisotropy of the surfaces could also account for the departures of the measurements in the ξz plane from the predicted form. Non-Gaussian intensity profiles cause the results derived by the method described in section 5Ba 1 i) to be inappropriate.

7.4 v) Measurements from single scale surfaces

Measurements of the mean intensity scattered with angle are presented in figures 7.7 to 7.10. Significant measurements above background could be obtained up to scattering angles of 2.2° for plate 201, and 36° for plate 198, reflecting the different correlation lengths of the surfaces.

Surface 201

ξz plane

Four sets of data are presented in figures 7.1 i) and ii), as once again there was more variation between sets in this plane. The curve has been calculated from equation (7.12) with $L = 66 \mu\text{m}$. The results are shown in two graphs, 7.7 i) has values for $\log(\sin\theta)$ as abscissa; in ii) the abscissa are $\sin^2 \theta$. The latter plot emphasises the single scale nature of the surface. In addition, from 7.7 i) it would appear that at scattering angles greater than 2° there is a region displaying multiscale behaviour; there is a distinctly linear part which diverges from the predicted form. When the data are plotted as a function of $\sin^2 \theta$ however, where linear it would imply single scale behaviour, there is again a linear part. The surface could be considered to have two predominant scales.

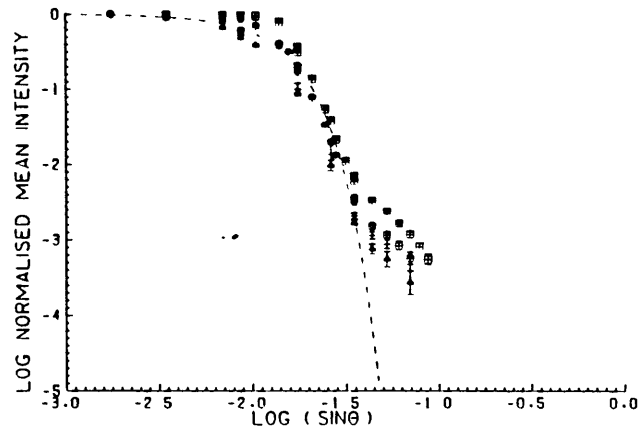


Figure 7.7 i). Measurements from plate 201 in the ξz plane compared with the single scale prediction of equation (7.12) for $L = 66 \mu\text{m}$.

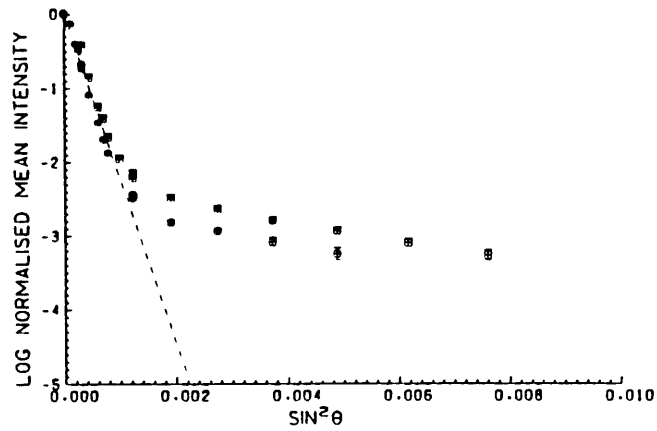


Figure 7.7 ii). As for figure 7.7 i) but plotted to emphasise the single scale nature of the surface.

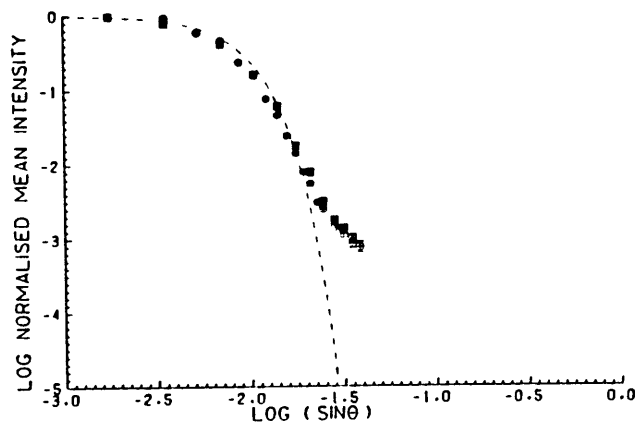


Figure 7.8. Measurements from plate 201 in the nz plane compared with the single scale prediction of equation (7.12) for $L = 110 \mu\text{m}$.

ηz plane

As for plate 200, there was less spread when measurements were made in this plane. The curve shown in figure 7.8 was calculated from equation (7.12) with $L = 110 \text{ } \mu\text{m}$. There is again a linear tail, this time for angles greater than 1.3° .

It should be noted that this region of enhanced scattering is not due to a defect in the scattering measurements. It is also predicted by the Talysurf data. The power spectrum for plate 201 shown in figure (6.12) displays two different regions of behaviour. In addition, when the mean scattered intensity with angle is predicted from knowledge of the autocorrelation function (as described in section 7.5) a similar discontinuity is apparent.

Surface 198

ξz plane

In figure 7.9 the curve shown compared with the scattering data was calculated from equation (7.12) with $L = 4.5 \text{ } \mu\text{m}$. The standard deviation was assumed to be the same as for plate 201, as they received the same exposure times at the same laser output. This surface was not measured on the Talysurf because it was expected to have a major frequency of 379200 m^{-1} which is beyond the resolution limit of the Talysurf (section 6.2).

In figure 7.9 it is clear that there is no discernable high angle tail such as appeared for plate 201. There are two possible reasons. Firstly, it may exist but at light levels too low to be detected using the current arrangement. Secondly, if the feature on plate 201 is due to an interference effect during exposure, then it is of a particular frequency. That frequency is lower than the predominant one for plate 198, and the scattering effects are consequently swamped by the stronger fluctuations.

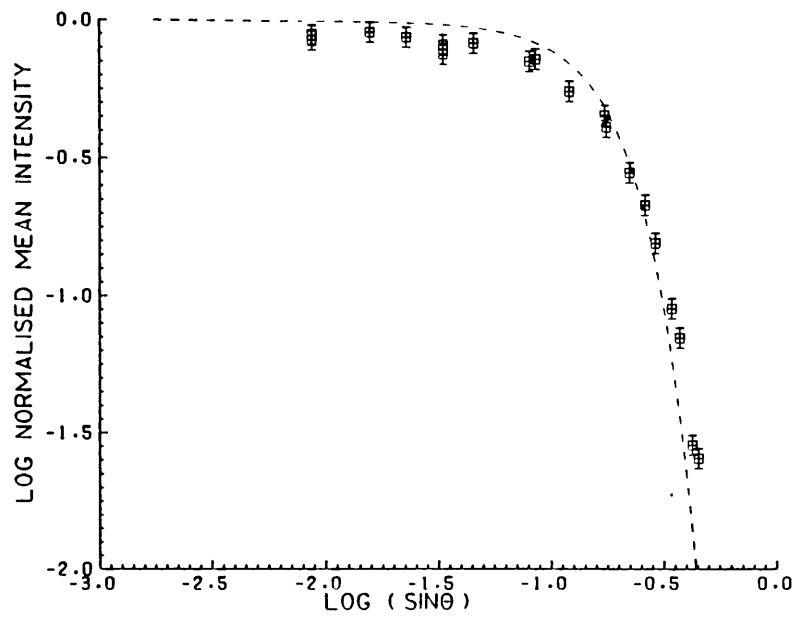


Figure 7.9. Measurements from plate 198 in the ξz plane compared with the single scale prediction of equation (7.12) for $L = 4.5 \mu\text{m}$.

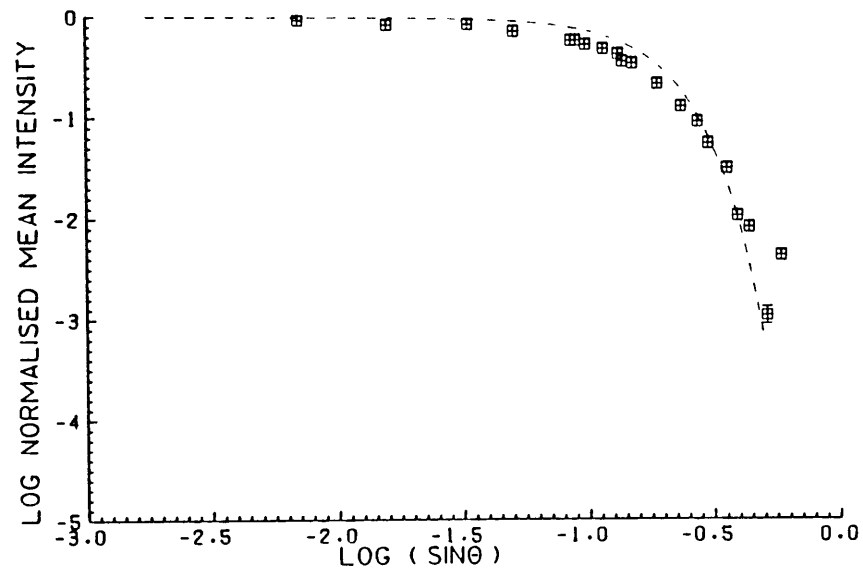


Figure 7.10. Measurements from plate 198 in the ηz plane compared with the single scale prediction of equation (7.12) for $L = 5 \mu\text{m}$.

ηz plane

The comparison curve in figure 7.10 was calculated from equation (7.12) with $L = 5 \mu m$. The same assumption about the standard deviation was made as for the ξz plane. Again the behaviour is characteristically single scale.

7.5 Prediction of mean scattered intensity from knowledge of the autocorrelation function of surface heights

It was shown in section 7.3 iv) that

$$\langle I(\xi, \eta) \rangle = F \left[C_{FNORM}(\Delta x, \Delta y) \right] \quad (7.16)$$

where F denote Fourier transformation.

$C_{FNORM}(\Delta x, \Delta y)$ is the normalised autocorrelation function of the complex fields immediately next to the scatterer, and is given by

$$C_{FNORM}(\Delta x, \Delta y) = \exp \{ -\sigma_{\phi}^2 \left[1 - C_{HNORM}(\Delta x, \Delta y) \right] \} \quad (7.13)$$

$C_{HNORM}(\Delta x, \Delta y)$ is the normalised autocorrelation function of surface heights.

In this work the autocorrelation function is known as a function of only one variable. Similarly scattering measurements are made in either the ξz or ηz plane (see figure 7.1), thus

$$\langle I(\frac{\sin \theta}{\lambda}) \rangle = F \left[\exp \{ -\sigma_{\phi}^2 \left(1 - C(X) \right) \} \right] \quad (7.41)$$

Note that the positions of points in frequency space resulting from Fourier transformation will be given by $\xi/f\lambda$ (figure 7.1 for coordinate system)

$$= \frac{\sin \theta}{\lambda}$$

The frequency separations depend on the lag spacings in the autocorrelation function. If the total number of points defining the function to be transformed is N , then the frequency separation is given by

$$1/N\Delta x \quad (6.2)$$

where Δx is the lag separation.

Similarly the maximum (Nyquist) frequency is given by

$$f_N = 1/2\Delta x \quad (6.1)$$

Thus, if the values for the autocorrelation function calculated from the Talysurf traces are used, where $\Delta x = 2 \mu\text{m}$, $f_N = 250000\text{m}^{-1}$. For a wavelength of 633 nm, and using (7.41) $\theta_{\text{max}} = 9^\circ$.

We have seen that appreciable scatter occurs from plate 200 at angles up to 15° , so this value of lag separation is too large for satisfactory predictions to be made. Lag separations of $0.5\mu\text{m}$ would result in $\theta_{\text{max}} = 39^\circ$.

Attempts were made to draw a smooth curve by hand between the points at $2 \mu\text{m}$ separations and so obtain values for the autocorrelation function at $0.5 \mu\text{m}$ separations. Sets of data obtained in this way were found to be unsatisfactory after Fourier transformation. Small rounding errors in the estimated autocorrelation function resulted in large oscillations in the predicted scattering at large angles.

Instead a polynomial curve fitting routine from the NAG library was used to fit a curve to the measured points; again separations of $0.5 \mu\text{m}$ were used. Three curves fit to the Talysurfed autocorrelation function for plate 200 in the x direction are shown in figure 7.11. The degree of polynomial chosen in each case is 6, 9, and 11 respectively. The three curves are virtually indistinguishable except near zero lag. The box shows this section of the curves magnified.

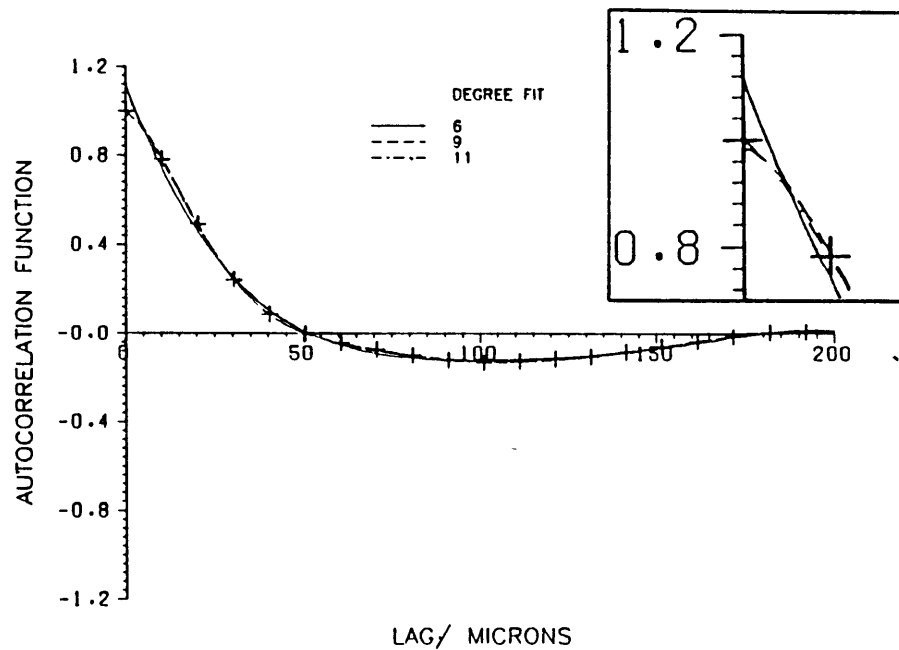


Figure 7.11. Polynomial fits of degree 6, 9 and 11, to the autocorrelation function for plate 200 in the x direction, every fifth point of the curve in figure 6.5 is shown. A magnified portion is shown in the box.

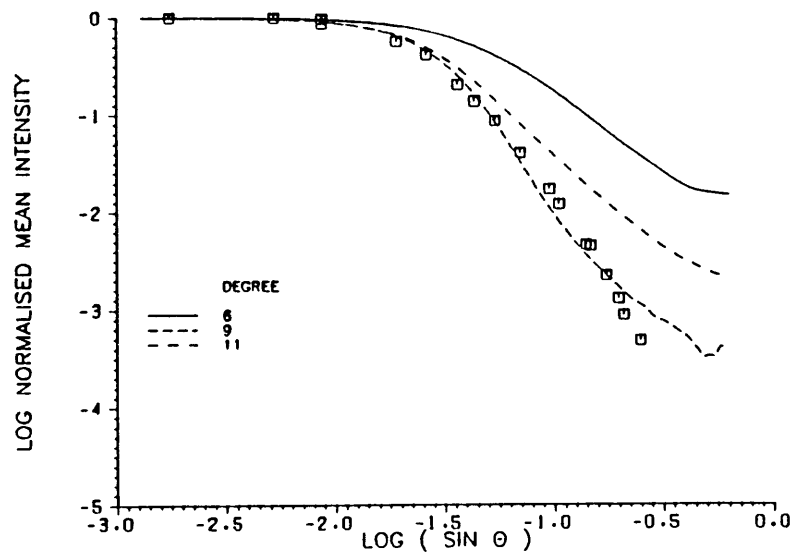


Figure 7.12. The predicted angular scattered mean intensity using equation (7.41) and the fits shown in figure 7.11, compared with scattering data for plate 200.

On the plot only every fifth point of the measured autocorrelation function is shown, to avoid clutter. The result of this small difference between the curve fit to the autocorrelation function is dramatic when the mean scattered intensity is calculated using equation (7.41). The predicted curves are shown in figure 7.12, compared with a set of scattering data from plate 200. The curve predicted from a degree 9 fit to the autocorrelation function agrees quite well with the measured points. However, its validity is doubtful as the curve from a fit of degree 11 is quite different.

Curves calculated in the same way for plate 201 were also very dependent on the degree of fit to the autocorrelation function chosen. In each case though, the predicted scattering curves displayed a linear region for θ greater than 2° , just as in the scattering measurements (section 7.4 v). This feature is not an instrumental fault of the scattering set up.

Overall it is concluded that this method of predicting the mean angular scattered intensity is not satisfactory. There is too much dependence on the degree of fit chosen when interpolating values in the autocorrelation function. The method does allow the existence of features in the scattered envelope to be confirmed, but their actual form and position is not predicted.

CONCLUSION

Two methods, which have been implemented in practice, have been described for producing multiscale surfaces. The first of these is the one-dimensional method which was used to produce plate 141. When Talysurfed, this surface has the required inverse power law spectrum, with index -2.0 to within 5% over the expected range of spatial frequencies. This region is however small, and the power spectrum falls off only by a factor of ten across it. It has been suggested by Jakeman (1982), that if the fluctuations at the inner scale introduce path differences comparable with the wavelength then single scale models are more appropriate to the scattering. When scattering measurements were made from surface 141 the results did not agree with the multiscale predictions, and it seems likely that the behaviour described by Jakeman (1982) has occurred. In fact, the scattering from the orthogonal direction was more multiscale in character. The equations derived in chapter seven cannot be applied in this case, as the index ν is not equal to 1. The one-dimensional method can be considered successful as it resulted in the expected surface. However, the surface would not be of use for testing scattering theories as the multiscale range is small, and the small scale fluctuations too large for the behaviour to be uniquely multiscale. The experimental arrangement is such that a sufficient increase in frequency range cannot be obtained.

The method used to produce the surface multiscale in two directions resulted in a surface with a power law spectrum covering more than a decade of spatial frequencies. More importantly, the power spectrum is smaller at the small scale end by a factor of about 1000 compared with the large scale end. This means that, for the measured standard deviation of $0.5 \mu\text{m}$, the fluctuations at the inner scale are less likely to be comparable with the wavelength. The light scattering measurements from this surface were particularly satisfactory; unambiguous multiscale type scattering was observed.

Talysurf and scattering measurements both emphasised that the method used in producing plate 200 requires modification before truly isotropic multiscale surfaces are produced. The asymmetry of the laser beam can be corrected by using a pair of cylindrical lenses to

alter the width of the beam in one direction, leaving the perpendicular width unchanged. The focal lengths of the collimating lenses should be ascertained to a greater degree of precision (they were found generally to differ from the quoted value by the manufacturers stated tolerance of 5%), and these measured values used in the equations of section 5Ba.1. There is scope for extending the frequency range by the selection of suitable values for the focal length of the objective, collimating and field lenses. Small scale sizes of $1\mu\text{m}$ are easily attainable.

Future work on scattering from rough surfaces is likely to use incident radiation of wavelength $10.6\mu\text{m}$. This means that some features which can be resolved by a Talysurf, even using a $1\mu\text{m}$ stylus, will be of a scale smaller than the wavelength. The effect of such features is of particular interest.

APPENDIX

Derivation of the structure function corresponding to an unmodified power law spectrum

Consider the unmodified power law spectrum

$$G(u) = A |u|^{-(\nu + 1)} \quad (2.1)$$

Now, the autocorrelation function is given by the inverse Fourier transform of the spectrum

$$C(X) = \int_{-\infty}^{\infty} G(u) \exp(iuX) du \quad (1.67)$$

and the structure function by equation (1.60)

$$D(X) = 2 (C(0) - C(X)) \quad (1.60)$$

We have seen however, (section 2.6), that $C(0)$ for an unmodified power spectrum is infinite, so the expression for $C(X)$ will be left in integral form, thus

$$D(X) = 2 \left[\int_{-\infty}^{\infty} G(u) du - \int_{-\infty}^{\infty} G(u) \exp(iuX) du \right] \quad (A.1)$$

$$= 2 \int_{-\infty}^{\infty} (1 - \exp(iuX)) G(u) du \quad (A.2)$$

$$= 2A \int_{-\infty}^{\infty} \frac{(1 - \exp(iuX))}{|u|^{(\nu + 1)}} du$$

The function $G(u)$ is even, so (A.3) may be written

$$D(X) = 4A \int_0^{\infty} \frac{1 - \cos(Xu)}{u^{(\nu + 1)}} du \quad (A.4)$$

$$= 8A \int_0^{\infty} \frac{\sin^2(Xu/2)}{u^{(\nu+1)}} du \quad (A.5)$$

From Gradshteyn and Ryzhik (1980, expression 3.823)

$$\int_0^{\infty} x^{\mu-1} \sin^2(ax) dx = \frac{-\Gamma(\mu) \cos(\mu\pi/2)}{2^{(\mu+1)} a^{\mu}}$$

$$a>0, -2 < \text{Re}\mu < 0 \quad (A.6)$$

So, (A.5) may be written

$$D(X) = 8A \left[\frac{-\Gamma(-\nu) \cos(-\nu\pi/2) X^{\nu}}{2^{1-\nu} 2^{-\nu}} \right] \quad (A.7)$$

$$2 < \nu < 0$$

$$= -4A\Gamma(-\nu) \cos(-\nu\pi/2) X^{\nu} \quad (A.8)$$

$$2 < \nu < 0$$

From Abramowitz and Stegun (1965, expression 6.1.17)

$$\Gamma(z)\Gamma(1-z) =$$

$$-z\Gamma(-z)\Gamma(z) = \frac{\pi}{\sin(\pi z)} \quad (A.9)$$

thus

$$\Gamma(-z) = - \frac{\pi}{z\Gamma(z) \sin(\pi z)} \quad (A.10)$$

Substituting in (A.8)

$$D(X) = 4A \left[\frac{\pi \cos(-\nu\pi/2) X^\nu}{\nu\Gamma(\nu) \sin(\pi\nu)} \right] \quad (\text{A.11})$$

$$\begin{aligned} \text{Expanding } \sin(\pi\nu) &= 2\sin(\pi\nu/2)\cos(\pi\nu/2) \\ &= 2\sin(\pi\nu/2)\cos(-\pi\nu/2) \end{aligned}$$

$$\begin{aligned} D(X) &= \frac{2A\pi X^\nu}{\nu\Gamma(\nu)\sin(\pi\nu/2)} \\ 2 &< \nu < 0 \end{aligned} \quad (\text{A.12})$$

$$\begin{aligned} D(X) &= \frac{2A\pi X^\nu}{\Gamma(\nu+1)\sin(\pi\nu/2)} \\ 2 &< \nu < 0 \end{aligned} \quad (\text{A.13})$$

ACKNOWLEDGEMENTS

I am pleased to thank my two supervisors, Professor W T Welford and Professor J C Dainty, who, during the course of this work, have given much advice, support and encouragement.

Thanks also to my colleagues in the Optics Section whose help in a variety of fields is much appreciated. In particular my thanks to Eugenio Mendez and Neil Emerton.

The financial support of the Science and Engineering Research Council and the Industrial Associates Bursary is gratefully acknowledged.

Finally, I am grateful to Thaira Hakim for typing this manuscript.

REFERENCES

- ABRAMOWITZ M & STEGUN I A (1965)
Handbook of Mathematical Functions
Dover Publications. New York.
- ARMSTRONG S A & GOUGH P T (1978)
Investigations of errors in sample autocovariance
functions and their corresponding power spectra
J Opt Soc Am 68 (5) 568-572
- BECKMANN P & SPIZZICHINO A (1963)
The scattering of electromagnetic waves from
rough surfaces
Pergamon. London.
- BECKMANN P (1967)
Scattering of light by rough surfaces
Progress in Optics VI 53-69
- BENNETT H E (1963)
Specular reflectance of aluminized ground glass
and the height distribution of surface
irregularities
J Opt Soc Am 53 (12) 1389-1394
- BENNETT H E & PORTEUS J O (1961)
Relation between surface roughness and specular
reflectance at normal incidence
J Opt Soc Am 51 (2) 123-129
- BERRY E (1985a)
Making a multiscale random rough surface
Optica Acta 32 (2) 147-153
- BERRY E (1985b)
A technique for producing characterised multiscale
random rough surfaces for use in scattering
experiments
Acta Polytechnica Scandinavica, Appl Phys Series
150 245-248
- BERRY M V (1979)
Diffractals
J Phys A: Math Gen 12 (6) 781-797
- von BESTENREINER R F & WEIERSHAUSEN W (1971)
Die photographische Herstellung streuender
schichten mit beliebiger Indikatrix
Optik 32 (5) 446-455

- BUCKLEY R (1971)
Diffraction by a random phase screen with very
large RMS phase deviation I One dimensional
screen
Aust J Phys 24 351-371
- CHANDLER J S & DEPALMA J J (1968)
High brightness projection screens with high ambient
light rejection.
J Soc Motion Pict Telev Eng 77 1012-1024
- CHANDLEY P J (1976)
Surface roughness measurements from coherent
light scattering.
Optical & Quantum Electronics 8 323-327
- CHANDLEY P J (1979)
Measurement of surface roughness by the scattering
of coherent light.
Thesis. University of London
- CHANDLEY P J & WELFORD W T (1975)
A re-formulation of some results of P Beckmann for
scattering from rough surfaces
Optical and Quantum Electronics 7 393-397
- CLIFFORD S F, BOURICIUS G M B, OCHS G R &
ACKLEY M H (1971)
Phase variations in atmospheric optical propagation
J Opt Soc Am 61 (10) 1279-1284
- DAINTY J C (1972a)
Coherent addition of a uniform beam to a speckle
pattern
J Opt Soc Am 62 595-596
- DAINTY J C (1972b)
Speckle statistics and the detection of images in
hologram reconstructions of bubble chamber tracks
and other objects
Thesis. University of London
- DAINTY J C (1976)
The statistics of speckle patterns
Progress in Optics XIV 1-46
- DAVIES H (1954)
The reflection of electromagnetic waves from rough
surfaces
Proc IEE 101 pt IV 209-214
- ERDELYI A, MAGNUS W, OBERHETTINGER F,
& TRICOMI F G (1954)
Tables of integral transformations
McGraw Hill. London

- FRENIERE E R, O'NEILL E L & WALTHER A (1979)
Problem in the determination of correlation
functions II
J Opt Soc Am 69 (4) 634-635
- GALE M T & KNOP K (1980)
Surface-relief images for color reproduction
Focal Press. (London)
- GEORGE N & JAIN A (1974)
Space and wavelength dependence of speckle intensity
Appl Phys 4 201-212
- GIGLIO M, MUSAZZI S & PERINI U (1979)
Surface roughness measurements by means of speckle
wavelength decorrelation
Opt Commun 28 (2) 166-170
- GOODMAN J W (1984)
Statistical properties of laser speckle patterns in Laser
speckle and related phenomena (Second edition) Topics
in Applied Physics Vol 9
Ed J C Dainty. Springer Verlag (Berlin)
- GRADSHTEYN I S & RYZHIK I M (1980)
Table of integrals, series and products
Academic Press
- GRAY P F (1978)
A method of forming optical diffusers of simple known
statistical properties
Optica Acta 25 (8) 765-775
- HENSLER D H (1972)
Light scattering from fused polycrystalline
aluminium oxide surfaces
Appl Optics 11 (11) 2522-2528
- HEWISH A, SCOTT P F & WILLS D (1964)
Interplanetary scintillation of small diameter radio
sources
Nature 203 1214-1217
- HOLLINS R C & JORDAN D L (1983)
Measurements of 10.6 micron radiation scattered by a
pseudo random surface of rectangular grooves
Optica Acta 30 (12) 1725-1734
- HOUCHEMS A F & HERING R G (1967)
Bidirectional reflectance of rough metal surfaces
Prog Astronaut Aeronaut 20 65-89
- HUREWICZ W & WALLMAN H (1948)
Dimension theory
Oxford University Press

- INTRILIGATOR D S & WOLFE J H (1970)
Preliminary power spectra of the interplanetary plasma
The Astrophysical Journal 162 L187-L190
- IWAMOTO A (1980)
Artificial diffuser for Fourier transform hologram
recording
Appl Optics 19 (2) 215-221
- JAKEMAN E (1980)
Speckle statistics with a small number of scatterers
SPIE 243 Applications of speckle phenomena
- JAKEMAN E (1981)
A simple multiscale scattering model
Optica Acta 28 (4) 435-441
- JAKEMAN E (1982)
Fresnel scattering by a corrugated random surface
with fractal slope
J Opt Soc Am 72 (8) 1034-1041
- JAKEMAN E & HOENDERS B J (1982)
Scattering by a random surface of rectangular grooves
Optica Acta 29 (12) 1587-1598
- JAKEMAN E & McWHIRTER J G (1977)
Correlation function dependence of the scintillation
behind a deep random phase screen
J Phys A 10 (9) 1599-1643
- JAKEMAN E & PUSEY P N (1973)
The statistics of light scattered by a random phase
screen
J Phys A 6 L88-L92
- JAKEMAN E & PUSEY P N (1975)
Non-Gaussian fluctuations in electromagnetic radiation
scattered by a random phase screen I. Theory
J Phys A 8 (3) 369-391
- JAMES T H & HIGGINS G C (1960)
Fundamentals of photographic theory
Morgan & Morgan Inc (New York)
- JENKINS G M & WATTS D G (1968)
Spectral analysis and its applications
Holden Day (San Francisco)
- JORDAN D L, HOLLINS R C & JAKEMAN E (1983)
Experimental measurements of non-Gaussian scattering
by a fractal diffuser
Appl Phys B 31 (3) 179-186
- KERR J R (1972)
Experiments on turbulence characteristics and
multiwavelength scintillation phenomena
J Opt Soc Am 62 (9) 1040-1049

- KOLMOGOROV A N (1941a)
The local structure of turbulence in incompressible
viscous fluid for very large Reynolds numbers
Compte Rendus (Doklady) de l'Academie des Sciences
de l'URSS 30 301-305
- (Reprinted in (1961)
FRIEDLANDER S K & TOPPER L
Turbulence: Classic papers on statistical theory.
Interscience. New York)
- KOLOMOGOROV A N (1941b)
Dissipation of energy in the locally isotropic
turbulence
Compte Rendus (Doklady) de l'Academie des Sciences
de l'URSS 32 16-18
- (Reprinted in (1961)
FRIEDLANDER S K & TOPPER L
Turbulence: Classic papers on statistical theory.
Interscience. New York)
- KURTZ C N (1972)
Transmittance characteristics of surface diffusers
and the design of nearly band-limited binary
diffusers
J Opt Soc Am 62 (8) 982-989
- KURTZ C N, HOADLEY H O & DEPALMA J J (1973)
Design and synthesis of random phase diffusers
J Opt Soc Am 63 (9) 1080-1092
- LAHART M J & MARATHAY A S (1975)
Image speckle patterns of weak diffusers
J Opt Soc Am 65 (7) 769-778
- LEE T C (1974)
Holographic recording on thermoplastic films
Appl Optics 13 (4) 888-895
- LEVINE B M & DAINITY J C (1983)
Non-Gaussian image plane speckle: measurements from
diffusers of known statistics
Opt Commun 45 (4) 252-257
- MacANDREW J A (1980)
Characteristics and applications of positive
photoresist with particular reference to holographic
recording
Thesis. University of London
- MANDELBROT B B (1975)
Les Objets fractals: forme, hasard et dimension
Flammarion, Paris

- MANDELBROT B B (1982)
The fractal geometry of nature
Freeman and Co
- MERCIER R P (1962)
Diffraction by a screen causing large random phase
fluctuations
Proc Camb Phil Soc A 58 382-400
- NAKAGAWA K & ASAKURA T (1978)
Contrast dependence of white light image speckles
on surface roughness
Opt Commun 27 (2) 207-213
- NIETO-VESPERINAS M & GARCIA N (1981)
A detailed study of the scattering of scalar waves
from random rough surfaces
Optica Acta 28 (12) 1651-1672
- NORMAN S L & SINGH M P (1975)
Spectral sensitivity and linearity of Shipley
AZ-1350J photoresist
Appl Optics 14 (4) 818-820
- OBUKHOV A (1941)
On the energy distribution in the spectrum of a
turbulent flow
Comptes Rendus (Doklady) de l'Academie des
Science de l'URSS 32 (1) 19-21
- (Reprinted in (1961)
FRIEDLANDER S K & TOPPER L
Turbulence: Classic papers on statistical theory
Interscience. New York)
- OHTSUBO J & ASAKURA T (1975)
Statistical properties of speckle intensity variations
in the diffraction field under illumination of
coherent light
Opt Commun 14 (1) 30-34
- O'NEILL E L & WALTHER A (1977)
A problem in the determination of correlation functions
J Opt Soc Am 69 (8) 1125-1126
- PAPOULIS A (1985)
Probability, Random variables and stochastic processes.
Second edition.
McGraw Hill (London)
- PARRY G (1974)
Some effects of temporal coherence on the first order
statistics of speckle
Optica Acta 21 (10) 763-772

- PARRY G (1975)
The scattering of polychromatic light from rough
surfaces: first order statistics
Optical and Quantum Electronics 7 311-318
- PARRY G (1984)
Speckle patterns in partially coherent light in
Laser speckle and related phenomena (second edition)
Topics in Applied Physics Vol 9
Ed. J C Dainty. Springer Verlag (Berlin)
- PARRY G, PUSEY P N, JAKEMAN E & McWHIRTER J G (1977)
Focussing by a random phase screen
Opt Commun 22 (2) 195-201
- PORTEUS J O (1963)
Relation between the height distribution of a rough
surface and the reflectance at normal incidence
J Opt Soc Am 53 (12) 1394-1402
- PUSEY P N & JAKEMAN E (1975)
Non-Gaussian fluctuations in electromagnetic radiation
scattered by a random phase screen
II Application to dynamic scattering in a liquid
crystal
J Phys A: Math Gen 8 (3) 392-410
- RATCLIFFE J A (1956)
Some aspects of diffraction theory and their
application to the ionosphere
Rep Prog Phys 17 188-207
- RICHARDSON L F (1922)
Weather prediction by numerical process
Cambridge University Press
- RICKETT B J (1969)
Frequency structure of pulsar intensity variations
Nature 221 158-159
- RUMSEY V H (1975)
Scintillations due to a concentrated layer with a
power-law turbulence spectrum
Radio Science 10 (1) 107-114
- SAYLES R S & THOMAS T R (1978)
Topography of random surfaces
Nature 271 431-434; 273 573
- SHIPLEY CHEMICALS LTD.
Microposit 1300 series photo resist. Data sheet

- STANSBERG C T (1979)
Surface roughness measurements by means of polychromatic
speckle patterns
Appl Optics 18 (23) 4051-4060
- TATARSKI V I (1961)
Wave propagation in a turbulent medium
McGraw Hill
- USCINSKI B J, BOOKER H G & MARIANS M (1981)
Intensity fluctuations due to deep phase screen with
power law spectrum
Proc Roy Soc Lond A 374 503-530
- WAAG R C & KNOX K T (1972)
Power spectrum analysis of exponential diffusers
J Opt Soc Am 62 (7) 877-881
- WALKER J G & JAKEMAN E (1982)
Non-Gaussian light scattering by a ruffled water
surface
Optica Acta 29 (3) 313-324
- WATSON G N (1962)
A treatise on the theory of Bessel functions. Second
edition
Cambridge University Press
- WELFORD W T (1977)
Review. Optical estimation of statistics of surface
roughness from light scattering measurements
Optical and Quantum Electronics 9 269-287

PUBLISHED PAPERS

Reprints of the following publications are submitted as supplementary material

- 1) Optica Acta 32 (2) (1985) 147-153 "Making a multiscale random rough surface".
- 2) Acta Polytechnica Scandinavica, Appl Phys Series 150 (1985) 245-248 "A technique for producing characterised multiscale random rough surfaces for use in scattering experiments".

Making a multiscale random rough surface

ELIZABETH BERRY

Blackett Laboratory, Imperial College, London SW7 2BZ, England

(Received 14 September 1984)

Abstract. A technique is described for making multiscale random rough surfaces (e.g. fractals) of prescribed surface height power spectrum and approximately gaussian probability distribution of height

1. Introduction

Investigation of any rough surface or inhomogeneous medium might involve scattering an incident wave of a wavelength appropriate to the scale of the sample. Several theoretical predictions have been made about this behaviour [1, 2], invoking a so-called fractal model to describe the scatterer. Fractal is a term used to describe the geometry of many natural phenomena and it has come into frequent use since it was coined by Mandelbrot in 1977 [3]. In the present application it implies the existence of many length scales, self-similarity over a wide range of these scales and a gaussian distribution of surface heights.

Experimental tests of these predictions are few. Work has been reported [4] on a diffuser produced by randomly roughening the surface of a piece of germanium with an engraving tool. It exhibited the characteristics required by the fractal model, having a power law spectrum and a gaussian distribution of surface heights as deduced from experiments at wavelength $10.6\ \mu\text{m}$. However this method is necessarily random since there can be little control over the parameters.

The method described here allows the index for the inverse power of the spectrum to be chosen. There is also some measure of control over the spatial frequencies in the finished product.

2. Choice of technique

The method chosen to produce the surfaces is based on the technique of Gray [5] using photoresist and recently used successfully by Levine and Dainty [6]. When photoresist is exposed to light a photochemical reaction occurs which changes its physical properties. An example is Shipley Microposit 1375, which is a positive type, i.e. it becomes soluble in a developer on exposure. With appropriate development the response may be linear, i.e. the etch depth is proportional to the integrated intensity of the incident light. An incident laser speckle pattern containing a full range of tones would be reproduced as a relief pattern on the resist.

3. Theoretical background

The model commonly chosen to describe a multiscale rough surface [2] is a gaussian random function $h(x)$ for the surface height with the power law power spectrum,

$$G(u) = \frac{A}{|u|^\alpha} \quad (1)$$

where α lies between 1 and 3. Since speckle patterns are to be used to produce the surface we review some appropriate features. The first-order statistical properties of a far-field speckle pattern are independent of the diffuser which produces it as long as there is a large number of scatterers, no depolarization and the path differences introduced are sufficiently greater than the wavelength to ensure that the phase has circular statistics. However, the shape of the scattering area controls the second order statistics; this property is used here.

For the non-imaging geometry shown in figure 1, where the intensity of the field incident on the scattering area is represented by $|P(\xi, \eta)|^2$, the power spectrum of intensity in the observation plane is given by [7],

$$G(u, v) = \langle I \rangle^2 \left\{ \delta(u, v) + \frac{\iint_{-\infty}^{\infty} |P(\xi, \eta)|^2 |P(\xi - \lambda zu, \eta - \lambda zv)|^2 d\xi d\eta}{\left[\iint_{-\infty}^{\infty} |P(\xi, \eta)|^2 d\xi d\eta \right]^2} \right\}, \quad (2)$$

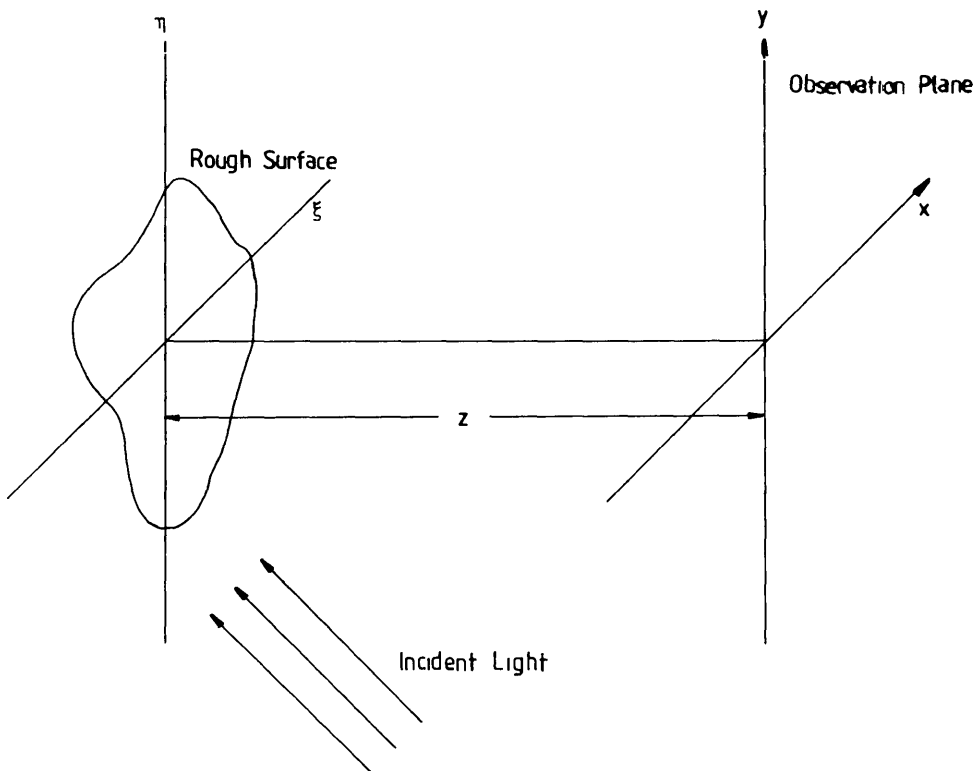


Figure 1 Geometry for speckle formation in free space

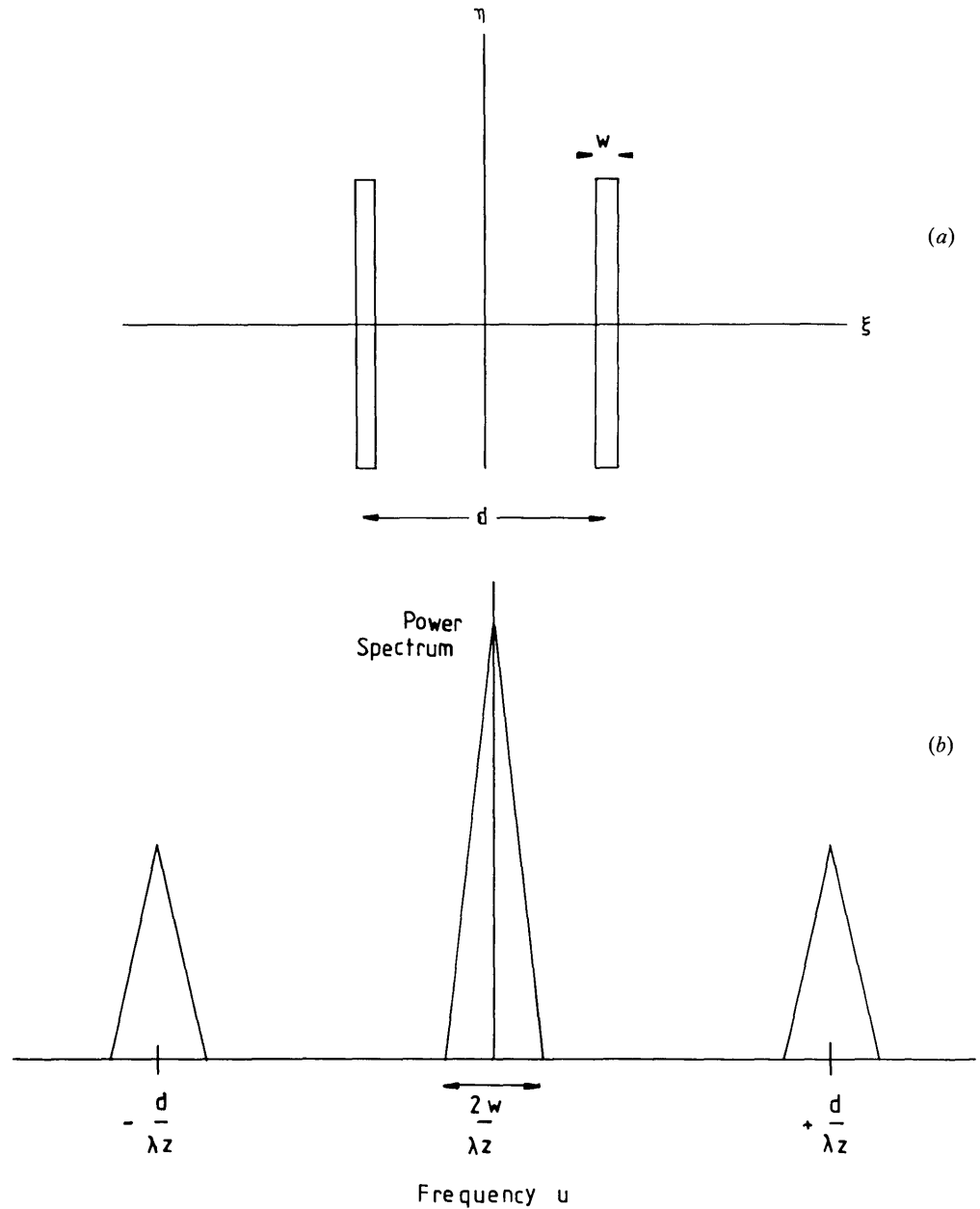


Figure 2 (a) Shape chosen as scattering area, (b) Power spectrum of speckle from the shape in (a).

where z is the slit-to-plate distance. Thus in addition to the delta function component $\delta(u, v)$ which appears in the power spectrum at zero spatial frequency, there is a component which is dependent on the intensity distribution incident on the scattering area, $|P(\xi, \eta)|^2$, and has the shape of the autocorrelation of $|P(\xi, \eta)|^2$. Consider then a scattering area in the shape of a pair of slits as shown in figure 2(a). The power spectrum for the u component of spatial frequency

must be as in figure 2(b). In order to build up a power spectrum with a power law dependence on spatial frequency we superimpose patterns of this kind for several different slit separations, such that the individual triangular spectra sum to give the required resultant curve. This is shown in figure 3, where the triangles are arranged so that the peak of one is at the same abscissa as the feet of the two on either side. The width of each triangle is $2w/\lambda z$. When arranged as in figure 3 the triangles are spaced at intervals in frequency of $w/\lambda z$, corresponding to changes in slit separations of w . The height of each triangle is proportional to the square of the intensity of the corresponding speckle pattern, from the $\langle I \rangle^2$ of equation (2). In practice, when recording on photoresist, the height of each triangle is controlled by the length of exposure. The exposures must be chosen so that the peaks of the triangles which make up the power spectrum decrease with increasing spatial frequency according to the required inverse power law with the index α of equation (1) having a value between 1 and 3. Thus the actual exposure times follow an inverse power law with index $\alpha/2$.

Ideally, the fractal surface should have a gaussian probability distribution of surface height. I do not expect to obtain this exactly by the method but the approximation will be good. If N totally uncorrelated speckle patterns are summed incoherently the probability density of total intensity is [7]

$$P(I) = \sum_{k=1}^N \frac{\lambda_k^{N-2}}{\prod_{\substack{p=1 \\ p \neq k}}^N (\lambda_k - \lambda_p)} \exp\left(-\frac{I}{\lambda_k}\right) \quad (3)$$

for $I \geq 0$, and zero for negative I . The λ_k are the mean intensities of the individual speckle patterns (assumed all different), each with the usual negative exponential distribution of intensity. The shape of this distribution for several values of N is shown in figure 4, the λ_k are chosen to give $\alpha = 2$. For larger N the shape is reasonably close to gaussian. The variance of the total speckle pattern can be given by

$$\sigma_I^2 = \sum_{k=1}^N \frac{2\lambda_k^{N+1}}{\prod_{\substack{p=1 \\ p \neq k}}^N (\lambda_k - \lambda_p)} - \left\{ \sum_{k=1}^N \frac{\lambda_k^N}{\prod_{\substack{p=1 \\ p \neq k}}^N (\lambda_k - \lambda_p)} \right\}^2. \quad (4)$$

From this expression we can compute a gaussian distribution of the same variance and mean, and compare the contrasts C [7] as given by

$$C = \sigma / \langle I \rangle \quad (5)$$

Figure 5 shows the drop in contrast from adding patterns. The loss is not too severe when the components are chosen to give the power index $\alpha = 2$.

4. Practical considerations

The arrangement for the variable slit separation is as in figure 6. The cube beam splitter is set so that the incident beam is split into two emerging parallel beams, so that only a single slit diffuser is needed. The effective separation of the pair of virtual slits is changed by translating the cube as shown in figure 6. The

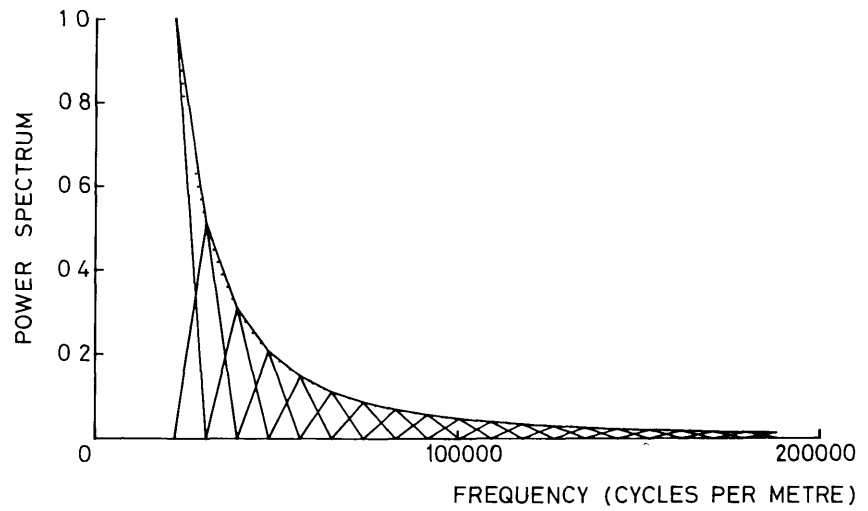


Figure 3 A composite specimen designed to have a square law power spectrum, the components shown correspond to speckle patterns formed at a distance of 0.5 m from 2 mm wide slits. The dotted line is the inverse square curve, and the solid line the composite curve from the triangles.

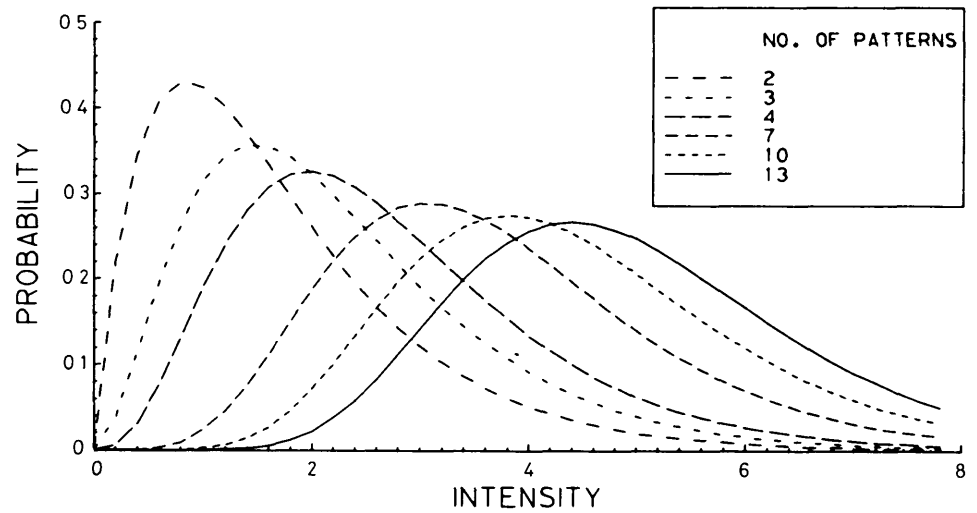


Figure 4 The probability distribution from equation (3) for several values of N , the number of superimposed patterns.

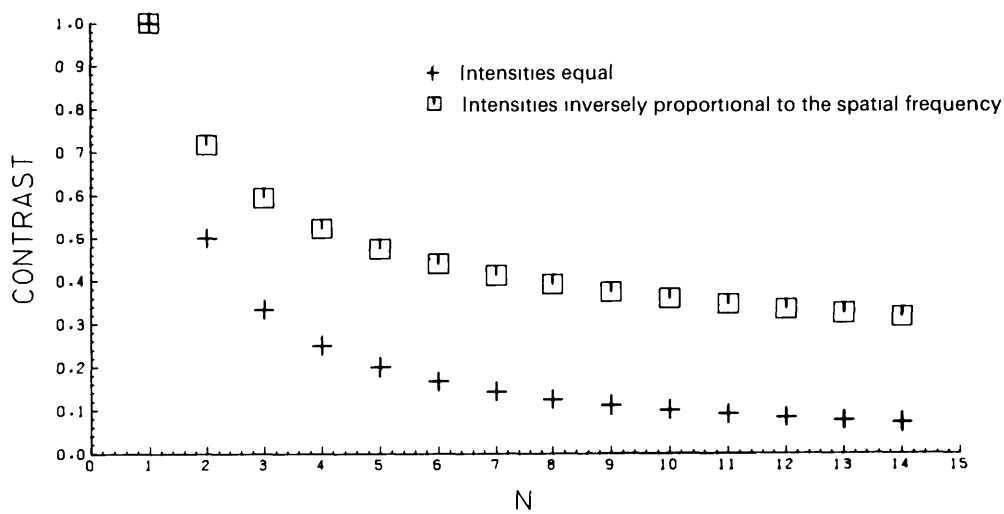


Figure 5 Contrast of N incoherently added speckle patterns

cube mount has rotation and translation movements; the rotation sets the parallelism of the beams and the translation adjusts the separation, which calibrated by means of a travelling microscope.

The available range of slit separation is limited by two factors. First, there is a limit to the number of exposures we can make on the photoresist and thus on the number of 'triangles' available to build up the power spectrum. Too many exposures produce a decrease in contrast in the summed speckle pattern. Also there is a practical limit to the thickness of resist which can be used. Secondly, the cube dimensions (40 mm by 45 mm by 45 mm in this case) define the range of slit separations. Exposures are made at separations differing successively by the width of the slit, and thus the total number depends on the slit width. Here a slit width of 2 mm was chosen, giving a maximum of twelve exposures.

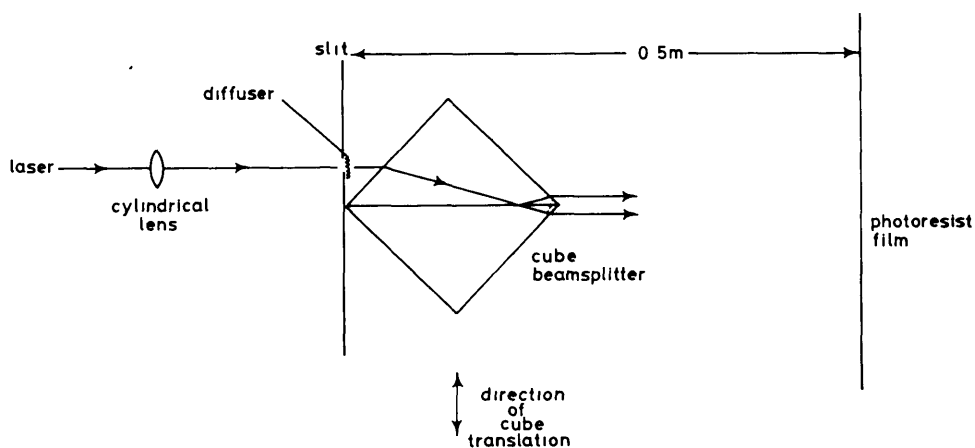


Figure 6 The experimental arrangement

The distance between the slit and the recording plane was determined by the output of the laser used; 0.5 m with wavelength 457.9 nm was used as a reasonable compromise on exposure time. The spatial frequencies obtained range from $2.2 \times 10^4 \text{ m}^{-1}$ to $1.2 \times 10^5 \text{ m}^{-1}$, corresponding to slit separations between 5 and 28 mm.

If the resist would tolerate more exposures greater spatial frequencies could be obtained by reducing the slit to plate distance z while retaining the maximum slit separation. The exposure times would have to be calculated allowing for this change in distance.

5. Discussion

Work is in progress to produce surfaces by this technique. The photoresist plates are 50 mm square glass slides coated by spinning at 1800 r.p.m., and baked for 30 min at 90°C to ensure complete evaporation of the solvent. This gives a smooth coating about 5 μm thick, which, as mentioned above, is a factor limiting the number of exposures before bottoming on the glass support. Refinements to allow an increase in the range of spatial frequencies are being considered, as is a method involving extra background illumination to produce a gaussian distribution of surface heights from a single exposure.

Acknowledgments

Thanks are due to Professor W. T. Welford for suggestions and encouragement and to the Science and Engineering Research Council for financial support.

References

- [1] BERRY, M. V., 1979, *J. Phys. A*, **12**, 781
- [2] JAKEMAN, E., and MCWHIRTER, J. G., 1981, *Appl. Phys. B*, **26**, 125
- [3] MANDELBROT, B. B., 1977, *Fractals*, (San Francisco: Freeman)
- [4] JORDAN, D. L., HOLLINS, R. C., and JAKEMAN, E., 1983, *Appl. Phys. B*, **31**, 179
- [5] GRAY, P. F., 1978, *Optica Acta*, **25**, 765
- [6] LEVINE, B. M., and DAINITY, J. C., 1983, *Optics Commun.*, **45**, 252
- [7] GOODMAN, J. W., 1984, *Laser Speckle and Related Phenomena*, second edition, Topics in Applied Physics, Vol. 9, edited by J. C. Dainty (Berlin: Springer Verlag)

A TECHNIQUE FOR PRODUCING CHARACTERISED MULTISCALE RANDOM ROUGH
SURFACES FOR USE IN SCATTERING EXPERIMENTS.

Elizabeth Berry, Optics Section, The Blackett Laboratory, Imperial
College, London, SW7 2BZ, England.

Abstract

A technique is proposed which would allow the production of multiscale random rough surfaces with a prescribed power law spectrum of surface heights, and approximately Gaussian probability distribution of height. Mechanical measurements from a surface so produced are presented, showing that the height distribution is approximately Gaussian; and the spectrum has a power law index within 5 % of the value sought.

1 Introduction

Scattering of coherent radiation by a random phase changing screen results in the formation of speckle. The statistics of such speckle can be predicted by using a simple scattering model. Such models have been applied to various problems from the ionospheric scattering of radio waves, to dynamic scattering liquid crystals. These models, however, generally assume that the scatterer contains fluctuations of similar characteristic size, corresponding to a Gaussian spectrum (1). The use of a model having a power law spectrum, and a range of scale sizes, was introduced in the study of propagation through turbulent media (2). Such models predict low contrast scattering, as opposed to smoothly varying screens whose scattered wave field is often dominated by geometrical caustics and singularities.

Experimental tests of these predictions are few (3), as no well characterised phase screen has been available. A method of producing a multiscale random rough surface, with a prescribed power spectrum of surface heights, and an approximately Gaussian probability distribution of height, is described here.

2 Technique

The method chosen to produce surfaces is based on that of Gray (4). Relief patterns are produced in photoresist by exposing it to laser speckle patterns.

To describe a multiscale random rough surface, the model chosen (5,6) is a Gaussian random function $h(x)$ for the surface height, with a power law spectrum given by

$$G(u) = \frac{A}{|u|^\alpha} \quad (i)$$

Thus any speckle pattern or patterns used in exposing the photoresist should result in this form of power spectrum. Goodman (7) gives an expression for the power spectrum of intensity for a speckle pattern which has propagated in free space. The spectrum has a component which has the shape of the autocorrelation of the intensity distribution incident on the scattering area. So, if the scattering area is made in the shape of a pair of slits, the resulting spectrum, in one dimension, will have a triangular feature at frequency $d/\lambda z$, of frequency width $2w/\lambda z$; where λ is the recording wavelength, d the slit separation, z the slit to plate distance, and w the slit width.

Patterns of this kind for several different slit separations, and different exposure times are superimposed such that the triangular spectra sum to give the required power law spectrum (8). In these experiments a value of $\alpha = 2$ was required.

When N totally uncorrelated speckle patterns are summed incoherently, the expected probability distribution of total intensity $P(I)$ is given by (7)

$$P(I) = \sum_{k=1}^N \prod_{\substack{p=1 \\ p \neq k}}^N \frac{\lambda_k^{N-2}}{(\lambda_k - \lambda_p)} \exp\left\{-\frac{I}{\lambda_k}\right\} \quad (ii)$$

for $I \geq 0$, zero for negative I . λ_k are the mean intensities of the individual speckle patterns. For the parameters used in this case, this approaches a Gaussian distribution.

3 Measurements

Ten Talysurf traces were made across a surface produced by this method

Each trace consisted of 1024 readings taken at separations of two μm .

Intensities in the recording speckle patterns correspond to depths in the resulting plate. The curve given by equation (ii) is shown in figure 1, together with a Gaussian curve and the probability distribution of surface depths from 10240 readings. Each plot has been normalised to have zero mean, unit area, and unit variance.

Figure 2 shows the power spectrum from a single Talysurf trace plotted in log-log form. The mean slope of ten such plots yields a value for α in equation (i) of 2.0 ± 0.1 , for spatial frequencies between 30000 and 138000 cycles per metre.

4 Discussion

The technique has resulted in the production of a surface which has the required inverse power law spectrum, and an approximately Gaussian distribution of surface height. The range of spatial frequencies covered is, however, very limited. Improved techniques are under consideration.

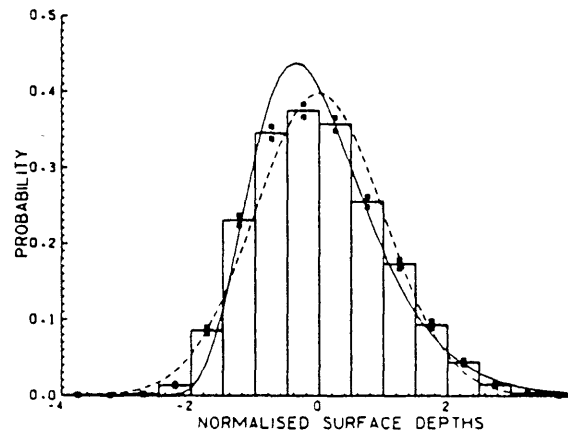


Figure 1. Histogram of surface depths from 10240 Talysurf readings. The solid curve is calculated from equation (ii); the dashed curve is a Gaussian curve; the points show the range of statistical uncertainty for each block.

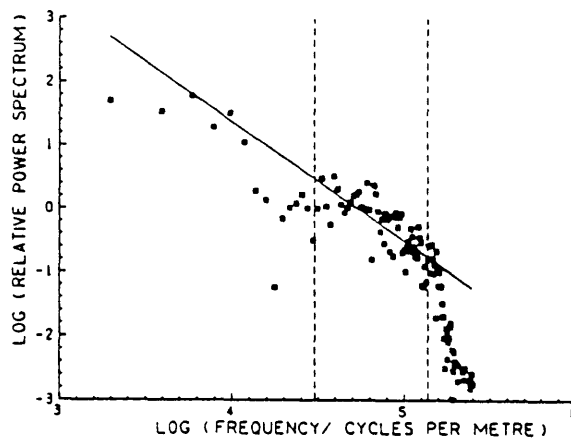


Figure 2. Spatial power spectrum of the surface profile from a single Talysurf trace of 1024 readings. The dashed lines indicate the range of spatial frequency used in recording; the solid line is a least squares fit to the points within this range.

References.

- (1) Jakeman E (1980) "Speckle statistics with a small number of scatterers", Applications of speckle phenomena, Proc. SPIE 243, 9-19.
- (2) Gochelashvily K S, Shishov V I (1975) Opt. Quant. Electron. 7, 524-536.
- (3) Jordan D L, Hollins R C, Jakeman E (1983) Appl. Phys. B. 31, 179-186.
- (4) Gray P F (1978) Optica Acta 25, 765-775.
- (5) Berry M V (1979) J. Phys. A. 12, 781-797.
- (6) Jakeman E, McWhirter J G (1981) Appl. Phys. B. 26, 125-131.
- (7) Goodman J W (1984) in Laser Speckle and Related Phenomena Second Edition. Topics in Applied Physics Vol. 9 edited by J C Dainty (Berlin: Springer Verlag)
- (8) Berry E (1985) Optica Acta 32 (To be published)

Acknowledgements

My thanks to Professor W T Welford for advice, and to the Science and Engineering Research Council for financial support.