University of Warwick institutional repository: http://go.warwick.ac.uk/wrap

A Thesis Submitted for the Degree of PhD at the University of Warwick

http://go.warwick.ac.uk/wrap/4156

This thesis is made available online and is protected by original copyright.

Please scroll down to view the document itself.

Please refer to the repository record for this item for information to help you to cite it. Our policy information is available from the repository home page.

AN IN-PROCESS, NON-CONTACT SURFACE FINISH SENSOR FOR HIGH QUALITY COMPONENTS GENERATED USING DIAMOND TURNING.

by

Ir. JAN HENRICUS RAKELS

This thesis is submitted for the degree of Doctor of Philosophy.

University of Warwick

Centre for Metrology and Micro-Engineering

Department of Engineering

Date of submission: April 1987

Contents.

1.	Introduction.	page	1
2.	Why measure ?	page	5
2.1.	Control of the manufacturing process.	page	6
2.2.	Performance of the component.	page	9
2.2.1.	Fatigue life.	page	10
2.2.2.	Bearing properties.	page	11
2.2.3.	Wear and friction.	page	11
2.3.	Machine tool condition monitoring.	page	11
3.	Parameters.	page	13
3.1.	Surface roughness.	page	14
3.2.	Surface roughness parameters.	page	15
3.2.1.	Summary of surface roughness parameters.	page	16
3.2.2.	Reference lines.	page	20
3.2.3.	Cut-off.	page	21
3.2.4.	Contact profiling instruments.	page	22
3.2.5.	Instrument calibration.	page	26
4.	Parameters obtained via powerspectral density functions.	page	28
4.1.	Characterisation of random surface profiles.	page	29
4.1.1.	The autocorrelation function ACF.	page	29
4.1.2.	The power spectral density function PSDF.	page	30
4.1.3.	Moments of the PSDF.	page	31
4.2.	Separation of the components of the surface roughness.	page	33

5.	The generation of surfaces and their micro-topographical properties.	page	35
5.1.	Turning.	page	36
5.2.	Relationship between surface finishes obtained by the various machining processes.	page	45
5.3.	Surface slopes and surface wavelengths expressed in optical wavelength.	page	52
5.4.	The progress of accuracy in machining.	page	55
6.	Techniques for measuring surface finish.	page	58
6.1.	Contacting profilometry (stylus technique).	page	59
6.2.	Contacting parametric techniques.	page	62
6.2.1.	Tactile testing.	page	62
6.2.2.	Thermal comparator.	page	64
6.2.3.	Friction dynamometer.	page	65
6.2.4.	Rolling ball.	page	66
6.2.5.	Flowing drop.	page	67
6.2.6.	Thetameter.	page	67
6.3.	Non-contacting non-optical techniques.	page	68
6.3.1.	Electrical methods.	page	68
6.3.2.	Fluid methods.	page	71
6.4.	Comparison of non-optical measurement techniques.	page	73
6.5.	Non-optical in-process measurements.	page	74
6.6.	Optical profilometry.	page	75
6.6.1.	Interferometry.	page	75
6.6.2.	Geometrical optics.	page	86
6.7.	Optical parametric techniques.	page	83

6.7.1. Total Integrated Scatter. page 87

6.7.2. Angular Resolved Scatter. page 100

6.7.3. Ellipsometry. page 104

6.7.4. Speckle. page 107

6.7.5. Interferometric technique. page 117

6.8. Comparison of the optical techniques and their suitability for in-process application. page 117

7. Comparison of the various scattering techniques. page 120

7.1. Various scattering theories. page 120

7.1.1. Geometrical optics. page 121

7.1.2. Scalar diffraction theory. page 122

7.1.3. Vector diffraction theory. page 124

7.1.4. Numerical methods. page 127

7.1.5. Comparison of the scattering theories. page 129

8. The Kirchhoff diffraction integral. page 136

8.1. History of diffraction. page 136

8.2. Derivation of the Kirchhoff scalar difraction formula. page 138

9. Application of the Kirchhoff diffraction integral to the inverse scattering problem for periodic surfaces. page 150

9.1. Derivation of the phase modulation model. page 150

9.2. Relation between the light intensity and the optical field. page 155

9.3. Relationship between the light intensity and the PSDF of surface profiles with small amplitudes. page 157

9.4. Calculation of the intensity for various diffraction patterns produced by periodic surfaces. page 158

9.5. Derivation of the algorithms for extracting the Rq parameter of periodic surfaces from their diffraction patterns. page 165

9.6. Relationship between the second and fourth moment of the diffraction pattern produced by a sinusoidal profile and the constants defining this profile. page 169

9.6.1. The second moment. page 169

9.6.2. The fourth moment page 170

9.7. Validity range of the Kirchhoff diffraction integral for sinusoidal profiles. page 173

9.8. Conclusions. page 177

10. Computer simulation of the optical diffraction system. page 179

10.1. Introduction. page 179

10.2. The computer program. page 180

10.3. Results obtained by the computer simulation program. page 186

10.4. Application of the simulation program. page 197

10.5. Conclusion. page 205

11. Deconvolution, page 206

11.1. Introduction. page 206

11.2. Definition of the convolution problem. page 207

11.3.1. Linear deconvolution methods. page 208

11.3.2. Constrained non-linear methods. page 215

11.4. Tests on the non-linear deconvolution algorithm. page 221

11.5. Discussion of the results. page 232

11.6. Conclusion. page 233

12. Practical embodiment. page 234

12.1. The sensor-head. page 234

12.1.1. The laser diode. page 237

12.1.2. The photo-diode array. page 239

12.1.3. The optics. page 241

12.1.4. Difference between the theoretical and practical system. page 245

12.2. The computer. page 250

12.2.1. Hardware. page 250

12.2.2 Software page 254

12.3. Practical results. page 257

12.4. Conclusions. page 260

Final conclusions. page 262

References page 264

Appendices.

A0 Machining processes.

A1 Summary of vector identities and vector operations.

A2 Mirror effect of a perfect, plane conductor.

A3 Calculation of the amplitudes of the diffraction pattern for even periodic surfaces.

A4 Derivation of the algorithms to extract Rq/ from the diffraction pattern.

A5 Accuracy of the algorithms for the determination of Rq/ from the value of the zeroth order of the diffraction pattern.

A6 Calculation of the second and fourth moments of the diffraction pattern for a sinusoidal surface profile.

A7 Data sheets on the collimated laser diode and the photo diode array.

A8 Data sheets on the photo-diode array processor driver assembly.

A9 Listing and flowcharts of the sensor software.

List of tables.

Table 5.1 Typical roughness values obtained with ordinary materials and common production processes.[Thwaite 1984]

Table 5.2 Surface texture of common machining processes, using the symbols from figure 5.5.

Table 6.1 Comparison of non-optical measurement techniques.

Table 6.2 Comparison of optical techniques.

Table 9.1 Fourier coefficients of various periodic surface profiles.

Table 9.2 Approximation of $J_o(Z)$.

Table 9.3 Number of diffraction orders containing a certain percentage of the diffracted light.

Table 10.1 Theoretical and simulated intensity amplitudes.

Table 12.1 The relative amplitudes of the spectra 12.13a-d.

Table 12.2 Results obtained with the algorithms 9.42 ans 12.7.

Table 12.3 Test results obtained by Rank Taylor Hobson Ltd.

List of illustrations.

figure 2.1 Why measure surfaces.

figure 3.1 Influence of direction of measurement.

figure 3.2 2RC filter network.

figure 4.1 Powerspectrum showing periodic and random contributions.

figure 5.1 Typical workpiece surfaces produced by in-feeding.

figure 5.2 Basic structure of a lathe.

figure 5.3 Frequency bands as related to the turning process.

figure 5.5 Symbols for the direction of lay. [BS 1134:Part 1:1972]

photo 5.1 Turned surface.

graph 5.1 Surface profile of the turned surface taken across the lay. Magnification: Horizontal X20, vertical X2000.

photo 5.2 Ground surface.

graph 5.2a Surface profile of the ground surface taken across the lay. Magnification: Horizontal X20, vertical X2000.

graph 5.2b Surface profile of the ground surface taken with the lay. Magnification: Horizontal X20, vertical X2000.

photo 5.3 Milled surface.

graph 5.3a Surface profile of the milled surface taken across the lay. Magnification: Horizontal X20, vertical X2000.

graph 5.3b Surface profile of the milled surface taken with the lay. Magnification: Horizontal X20, vertical X2000.

figure 5.6 Surface topography in terms of optical wavelength.

figure 5.7 Modern trends in high precision machining.

figure 6.1 Effectiveness of tactile comparison tests (Schlesinger).

figure 6.2 Michelson two-beam interferometer.

figure 6.3 Mirau interferometer.

figure 6.4 Optical paths in the dove prism interferometer.

figure 6.5 Fizeau interferometer.

figure 6.6 Fabry-Perot interferometer.

figure 6.7 Detail of the ac-interferometer.

figure 6.8 Profile obtained by SEM and contact stylus method.

figure 6.9 Detection of the surface slope.

figure 6.10 Wavelength dependence of TIS.

figure 6.11 Schematic diagram of an ellipsometer.

figure 6.12 Illumination and recording configurations for speckles.

figure 6.13 Variation of speckle contrast.

figure 7.1 Grating configuration

figure 7.2 Two surfaces with the same slope distribution but differing in R_a and R_q.

figure 7.3 Angular dependence of scalar and vector scattering.

figure 7.4 Validity range for the scalar diffraction theory.

figure 8.1 Geometry involved for the evaluation of integral 8.22.

figure 8.2a Geometry for solving integral 8.30.

figure 8.2a Geometry for solving integral 8.31.

figure 9.1a Three-dimensional scattering geometry.

figure 9.1b Two-dimensional scattering geometry.

figure 9.2 Phase modulation model.

figure 9.3a Sinusoidal surface profile.

figure 9.3b Approximated triangular surface profile.

figure 9.3c Approximated rectified sinusoidal surface profile.

figure 9.3d Approximated parabolic surface profile.

figure 9.4 Idealised light intensity pattern in the diffraction plane.

figure 9.5a Intensity amplitudes for a sinusoidal surface profile.

figure 9.5b ,, ,, ,, triangular ,, ,, .

figure 9.5c ,, ,, ,, rectified sinusoidal ,, .

figure 9.5d ,, ,, ,, parabolic surface ,, .

figure 9.6 Ratio of the accuracy of R_q/λ and the accuracy of I_o.

figure 9.7 Relationship between the second moment and Δ_q.

figure 9.8 Relationship between R_q and equation (9.50).

figure 9.9 Maximum angle of diffraction and maximum surface slope.

figure 9.10 Validity range of the Kirchhoff diffraction approach.

figure 10.1 Geometry of the optical diffraction system.

figure 10.2 Diagram of the diffraction system showing the variables used in the computer program.

figure 10.3 Flow diagram of the simulation program.

figure 10.4a Diffraction pattern produced by an uniform light intensity.

figure 10.4b Diffraction pattern produced by a Gaussian light intensity.

figure 10.4c Enlargement of fig. 10.4a.

figure 10.4d Enlargement of fig. 10.4b.

figure 10.4e Real diffraction pattern using a mirror.

figure 10.5 Diffraction patterns by profiles $y(x)=h\cdot\cos(2\pi x/D)$.

figure 10.6 Computer generated diffraction patterns by insertion of real profiles.

figure 10.7 Actual diffraction patterns obtained from diamond turned drums.

figure 10.8 The surface profiles used to produce the actual and simulated optical spectra.

figure 10.9 The digitally computed power spectral density functions of the profiles.

figure 10.10 Artificial surface profiles exhibiting certain periodic machining errors.

figure 10.11 Simulated diffraction patterns of surface profiles with certain periodic machining errors.

figure 10.12 Power spectral density functions of surface profiles with certain periodic machining errors.

figure 10.13 Modulation signal with which profile 10.10a has been multiplied to produce profile 10.10e.

figure 10.14 Power spectral density function of the modulation signal 10.13.

figure 10.15 Simulated diffraction patterns of the surface profiles, plotted on logarithmic scale.

figure 11.1 Convoluted input signal.

figure 11.2 "Ideal" diffraction pattern.

figure 11.3 Blur-function.

figure 11.4 Graphical representation of the relaxation functions A-G.

figure 11.5A Results obtained by the relaxation function A

figure 11.5B Results obtained by the relaxation function B

figure 11.5C Results obtained by the relaxation function C

figure 11.5D Results obtained by the relaxation function D

figure 11.5E Results obtained by the relaxation function E

figure 11.5F Results obtained by the relaxation function F

figure 11.5G Results obtained by the relaxation function G

figure 12.1 Physical lay-out of the sensor head.

figure 12.2 Initial and final lay-out of the transform system.

figure 12.3 Collimated laser diode.

figure 12.4 Electrical connections of the diode array electronics.

figure 12.5 Range of focal lengths obtainable with various combinations of lenses indicated.

figure 12.6 Influence of the shadowing effect.

figure 12.7 Conditions for a space-invariant transform system.

figure 12.8 Determination of SW_{max}, caused by the illumination.

figure 12.9 Pin connections of the i.c.'s used in the interface.

figure 12.10 General resistor network for unipolar operation of the A/D converter i.c. ZN427.

figure 12.11 Electronic circuit diagram for interfacing the BBC computer and the optical sensor-head.

figure 12.12 Flowchart of the computer program of the sensor.

photo 12.1 Complete optical surface finish measurement system.

photo 12.2 Internal components of the sensor-head.

figure 12.13 Optical power spectra, which are used to calculate R_{qop}

Acknowledgements.

I would like to acknowledge with gratitude the help of the following people during the course of research described in this thesis.

Professor D.J. Whitehouse for his continuous encouragement and constructive criticism.

Mr. R.A. Bridgland and Mr. J.S. Pillier, who managed to transform "engineering drawings" into hardware.

Mr. H.T. Hingle for having problems with non destructive measurements of diamond turned components, and thereby necessitating this research

Declaration.

This thesis is presented in accordance with the regulations for the degree of Doctor of Philosophy. It has been composed by myself and has not been submitted in any previous application for a higher degree. The work described has been undertaken by myself. All sources of information are specifically acknowledged.

J.H. Rakels.

Summary.

The object of this Ph.D. project was to design and construct an in-process, non contact surface finish sensor for high quality components generated using diamond turning. For this application the instrument must have the following properties:

- i rapid acquisition of data.
- ii capability of measuring translating and or rotating surfaces.
- iii ruggedness for in-process use.
- iv insensitivity to moderate vibrations.
- v remoteness from the surfaces to be measured.

The remoteness requirement virtually excludes the otherwise ubiquitous stylus instrument, while the rapid gathering of data from rotating surfaces excludes other profiling techniques. The above mentioned properties strongly suggest an optical method. An optical diffraction technique has been chosen, since it produces an optical Fourier Transform of the surface. This transform is produced at the speed of light, since the optical system has the property of parallel data processing, unlike a typical electronic computer. With the aid of a microprocessor various surface finish parameters can be extracted from the optical transform. These parameters are respectively the rms surface roughness, slope and wavelength.

The actual sensor consists of a measuring head and a mini-computer. It fulfils the above mentioned requirements. Its only limitations are:

- i limited to surface finishes upto 100nm.
- ii presence of cutting fluids has to be avoided, although certain modern lubricating fluids can be tolerated.

The algorithms devised to extract the surface finish parameters from the optical transforms have initially been tested on optical spectra produced by Thwaite. Comparison of the optical roughness values and the values quoted by Thwaite show close agreement. Thwaite's values are obtained by a stylus instrument.

Rqopt (um)	Rqstylus (um)
0.16	0.156
0.38	0.37
0.44	0.40

In addition a computer program has been devised which simulates the optical sensor head. The input data can be obtained by a profiling instrument, or generated by a computer program. This last option enables the creation of surface profiles with "controllable" machining errors. This program can be utilised to create an atlas, which maps optical diffraction patterns versus machine-tool errors.

1. Introduction.

The present revolution in manufacturing technology has led to searching questions as to future expectations. To answer these questions E. Merchant [1971] led a Delphi type study for C.I.R.P. in order to forecast certain events concerning the future manufacturing capability. Certain of these results are anticipated to impact directly on the future of surface finish measurement. For example, this study predicts that by about 1980 "complete in-process inspection of parts as they are machined will become a reality". In-process measurements made while the part is being machined or formed are distinguished from on-line measurements, i.e. real time, high speed, automated measurements that can be performed routinely and without disruption between stages of manufacturing. By about 1985, "automated assembly will be extended to the greater part of mass production operation by development of measuring, sorting and selfselecting techniques", and by about 1990, "automation of quality control by in-process inspection devices connected on-line with a process-computer forming part of a larger manufacturing system will be reality".

The above predictions clearly suggest that in-process measurements must be expected. This in-process gauging could be concerned with component dimensions, condition of the cutting fluids, surface finish etc.. That these measurements will, if possible, be used for real time, closed loop control of manufacturing processes, and that information gathered from these measurements will be employed to detect malfunctions such as tool wear and breakage. For

example, slow attainment of the required surface finish might signal the need for tool replacement. It is also valuable to terminate the machining process as soon as the required finish is obtained and thus increasing the throughput per machine. Only by considering the developing measurements needs engendered by automation can the surface finish measurement discipline adequately fulfil its new role in the manufacturing process. So the principal reason for in-process measurement is that there are no ideal machine and tools which are capable of producing in series surfaces of consistent quality and therefore the workpiece should really be subjected to continuous control. The common method today is to check the surface with a stylus instrument, this however requires a great deal of time and is only suitable for spot checks. A further disadvantage is that this instrument contacts the surface and only allows a linelike scanning, which produces general roughness parameters but which is also often unable to show the relation to the actual function of the surface. The two clear advantages of the stylus technique are an exact calculation of the geometric profile and a large measuring range.

The new in-process instrumentation must have very different properties from the stylus instrument and the microinterferometer. In order to supply the computer and control system with the necesssary information in sufficient time to alter the process or to stop a malfunctioning machine or to terminate a successful process, and to do so in a machineshop environment, the instrument must have the following properties:

i rapid measurement and processing of data.

ii provision of an electrical output signal of some kind.

iii gathering of many types of surface information in parallel.

iv richness in information, so the computer can select and process the necessary information.

v capability of measuring moving surfaces.

vi ruggedness.

vii suitability for a wide range of surfaces.

viii remoteness from the surface being measured.

ix relative insensitivity to environmental changes like temperature and vibrations.

x capability of measuring in the presence of cutting or lubrication fluids.

Furthermore it should be fairly small in order that it can be accommodated in the machine-tool.

The remoteness requirement virtually excludes the otherwise ubiquitous stylus instrument and strongly suggests an optical instrument.

In this thesis the development of an optical sensor for the measurement of surface finish is described which fulfils most of the aforementioned requirements. Its only limitations are:

i it is limited to surface finishes up to 100 nm.

ii presence of cutting fluids has to be avoided. However, some modern cutting lubricants can be accommodated.

The surface finish limit of Rq=100 nm applies to the usage of a semiconductor laser with a wavelength of 870 nm. This roughness range is normally achieved in diamond turning. It might appear that this only represents a small fraction of the totality of manufactured surfaces, but this fraction is not insignificant. Examples of these surfaces are photocopier drums, hard discs for computers but also germanium lenses. It is also interesting to note that these surfaces cannot be measured by stylus techniques because of possible damage to the components. The minimum surface wavelength to which this technique applies is approximately 10 λ. Where λ is the illumination wavelength.

The operation of the surface sensor is based upon the Kirchhoff scalar diffraction theory, which is a refinement of the Huygens-Fresnel approach. This theory forms also the basis of a computer program which simulates the sensor.

The surface finish values obtained with the sensor are in close agreement with results obtained with stylus instruments. The computer program generates optical diffraction spectra which are consistent with theoretical patterns and spectra obtained with the actual sensor. From this it can be concluded that the scalar diffraction theory is adequate for this particular application and that the more rigorous vector theory can be omitted.

2. Why measure?

Surfaces and their measurement provide a vital link between the manufacture of engineering components and their suitability for intended performance [Whitehouse 1978].

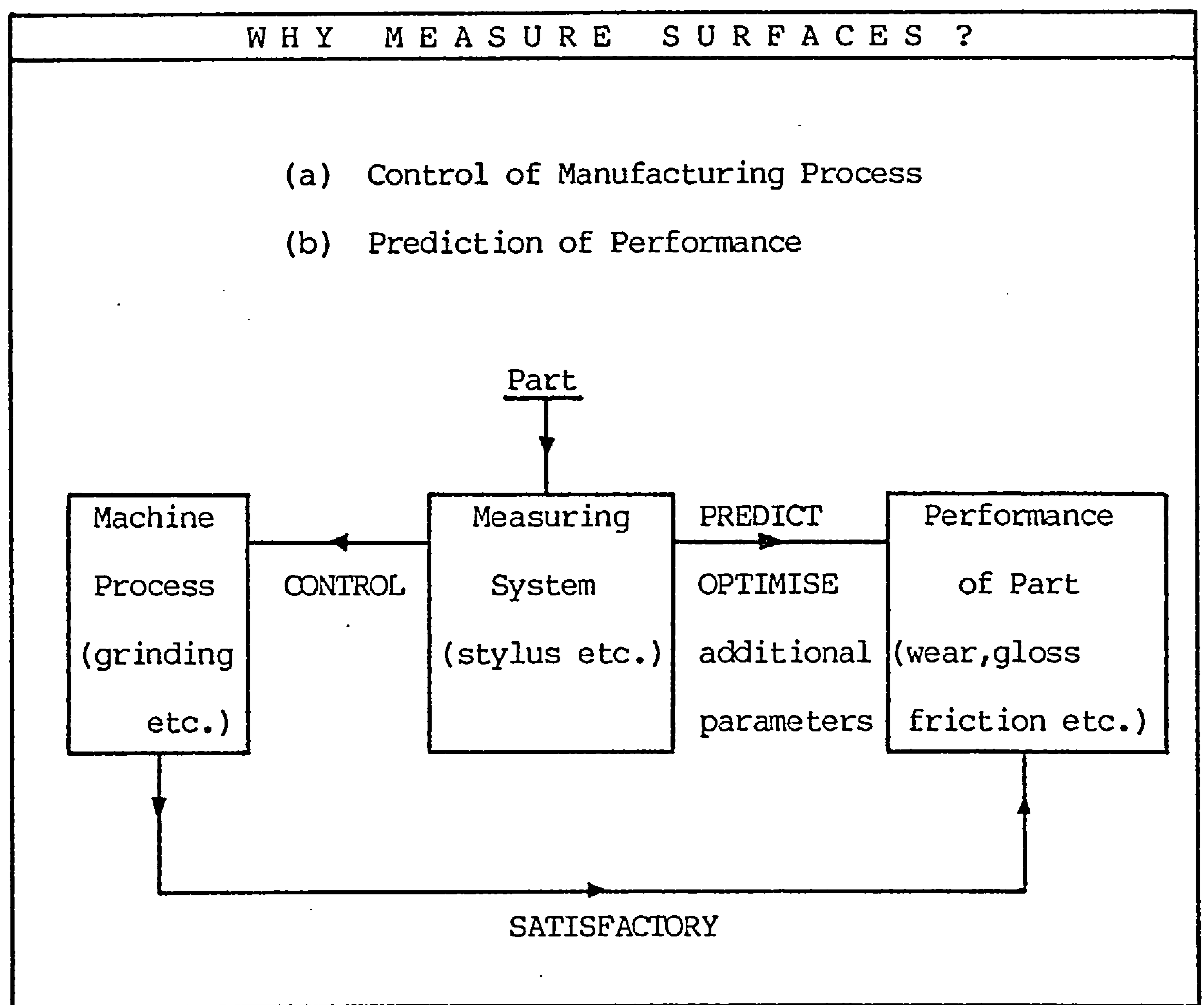


Fig. 2.1.

Figure 2.1 shows where the block representing surfaces and their measurement can be placed relative to the blocks representing

manufacture and function. The measurement of surface texture is not purely a means of quality control, but forms an essential part of the manufacturing process, used correctly it facilitates control of:

i the manufacturing process.

ii the performance of the component.

iii the machine tool condition.

2.1. Control of the manufacturing process.

Historically surface measurement was employed to control the manufacturing proces. In practice what happened was that a component was made and tested. If it functioned satisfactorily the same manufacturing conditions were maintained to produce the subsequent components. It soon became apparent that the control of the surface finish was being employed as an effective go not-go gauge for the manufacturing process. The reason for this possibility is that the texture is in effect a fingerprint of the complete manufacturing process. It is a very sensitive end product of a long sequence of operations. Any deviation in the manufacture reflects itself in the texture [Chetwynd et al. 1982]. So, simply making certain that the surface texture is within set limits ensures similar manufacture, and subsequently a similar performance of the workpiece.

This control was effectively an overall control, independent of any specific monitoring of process parameters. It proved to be quite satisfactory. Under these conditions the blocks in figure 2.1. were

balanced and thus an equilibrium existed.

Unfortunately, the control equilibrium can be broken either by changing the measurement parameters, the production process or even the function of the component. This may seem trivial, but it is the cause of every day problems in engineering.

Taking the measurement parameters first, there is much confusion at present, resulting in a "parameter rash" [Whitehouse 1982]. In principle any one of a number of parameters would be adequate for simple control. Some examples are R_a, R_t, R_q etc., definitions of which follow later. All these and many more besides are now appearing on production drawings, despite that they are nominally height parameters. Which one is the best to use in industry depends only on whether the criterion of acceptability is ease of measurement or reliability. The whole situation has been brought about by various attempts in different countries and large firms to improve control of the process without fully understanding the causes behind loss of control. The principal cause of this weakening of control is that of changing the manufacturing process parameters in some way. This cannot always be avoided because flexibility in the use of plant may dictate it or a new more economic process may have to be introduced. What is rarely done under these circumstances, is to perform a few tests to ensure that the existing control parameters are still valid. Often the process is left off the drawing completely. This violates the control philosophy and invariably leads to a weaker grip on the functional quality of the components. A similar situation arises if the function is changed. The loading on a bearing may be altered, the

speed of an engine increased or the oil formula changed, any of which will be changing the working environment.

In all these cases control values for textures should only be specified after suitable tests have been carried out. The problem in engineering today is that economics force changes in manufacture and function and yet apparently preclude adequate testing.

One way to allow freedom of production is to replace the single parameter texture control by a more stringent measurement of surface properties which may include physical (hardness) as well as geometric parameters.

In this method a comprehensive list of surface parameters is entered on the drawings to ensure that the workpiece is satisfactory in performance; in effect the processs parameters are replaced by surface parameters. This is not always a simple task because embodied in the process is a wealth of functionally significant properties of the surface which are not necessarily all geometric. The difficulty here, as before, is that sufficient tests have to be undertaken to establish the type of, number of and values of the new surface parameters. The bare minimum of parameters has to be sought because any correlation between parameters is wasteful and a large number of parameters will simply not get put on the drawings. On the other hand, if too few parameters are specified the control is weakened. To determine which are the best geometric parameters of the surface to use, statistical information techniques should be applied.

Statistically the complete profile can be condensed into the height density function and the auto-correlation function [Whitehouse et al.1974c]. The height density leaves out the order in which the

asperities occur, whereas the autocorrelation function (ACF) leaves out the amplitude of the profile.

The size of the height distribution can be represented by the R_a, R_q and the R_t parameters. Its shape can be condensed in the third and fourth central moments of the profile: the Skewness and Kurtosis [Whitehouse 1971].

The ACF can be condensed in a few values, by splitting it up in additive or modulated random and deterministic components [Peklinek 1967].

In order to characterize the deterministic component of the profiles, the amplitude spectrum has to be given. This spectrum can be condensed by the ratios of its harmonics to the fundamental and the presence of broad or narrow band noise. Theoretically it should be completed by the relative phases between the harmonics.

2.2. Performance of the component.

Another reason for surface finish measurement is the functional behaviour of the component. There are three important effects which are a function of the surface finish; namely fatigue life, bearing properties and wear. Other effects are electric contact, heat transfer etc.

The microtopography is studied so that functional behaviour may be explained in terms of parametric descriptions of surfaces. Measured parameters are correlated with the results of experiments or models of the physical behaviour constructed using values of the

parameters. A detailed knowledge of the texture is required in establishing and testing new designs. Flow in pipes and heat transfer are traditional examples of this which are of continuing interest, while the effects of the roughness on the efficiency of solar collectors is a more recent example.

2.2.1. Fatigue life.

If a component is subjected to repeated reversals of stress, it undergoes fatigue and its life is considerably shorter than it would be if the part carried an equivalent constant load [Reason 1960]. The number of stress reversals it can withstand at a given stress is called the fatigue life. Failure due to fatigue seems to start at a sharp corner, where stress concentrations occur, such as the root of a surface irregularity, even on a non-working surface.

It has been demonstrated that in certain engines the gudgeon pin, which carries cyclic loads, has a significantly longer life if its bore is highly finished. The bore of the gudgeon pin is not in contact with any other part of the engine, in fact a gudgeon pin is usually hollow simply to reduce the reciprocating weight. So polished finishes may be specified to minimize the occurrence of stress concentrations which limit fatigue strength.

2.2.2 Bearing properties.

A perfect surface, i.e. one with no irregularities whatsoever and therefore perfectly smooth is not a good bearing. In fact seizure will probably occur due to the difficulty of maintaining a lubricating film and causing a metal to metal contact. The same applies to engine cylinder walls. Probably the best form of surface for a bearing is one which has a large contact area to reduce the load and deep valleys to retain the lubricant. Plateau honing is a suitable process for producing such a surface

2.2.3. Wear and friction.

It is a well known law of physics that friction does not depend on contact area. However the rate of wear is dependent on the area in contact, the larger the area the lower the load per unit area and hence the lower the rate of wear.

According to Myers [1962] an important parameter for the prediction of the interaction between two metal surfaces is the slope of the flanks of the surface profile.

2.3. Machine tool condition monitoring.

Surface data can be used to monitor machining trends with a view to forstalling deterioration in production processes, and subsequent

substandard manufacture of components, which results in wastage of material. By analysing the surface finish, it is possible to monitor the condition of the machine tool [Raja and Whitehouse 1984].

The philosophy behind this is that the component produced by the machine is a fingerprint of the machine and tool combination. Therefore any changes in the condition of the machine or the tool should be reflected in the component, specially in the nature of the component surface generated and its dimensions. Since the condition of the machine is diagnosed through the part which has been produced, the information obtained is more relevant to the performance of the component.

3. Parameters.

All industrial nations, through their industrial standards organisations have developed standards documentation which specifies conditions governing the measurement of surface roughness. This has been done principally for the purpose of process control but also to provide a consistent basis for technical roughness measurement. This approach is necessary because the assessment of topography is dependent upon the means used. The British Standards Institution for example has published a comprehensive standard on the subject, called "Method for the Assessment of surface texture." BS 1134: part 1 and 2.

There is a large measure of agreement between the national standards related to surface texture, which has been brought about by the activities of the International Standards Organisation (ISO). There are a number of ISO standards which define nomenclature and measuring methods and which have been accepted as the basis for national standards.

National standards pay almost exclusive attention to the stylus technique of measurement. This is because although optical microscopy, light scattering, electro-optics, capacitive, pneumatic, and other methods can give information on a variety of surface properties, stylus profiling techniques are still currently superior for producing comprehensive quantitative descriptions of surface topography and are easily traceable.

3.1. Surface roughness.

When referring to a rough surface it is necessary to define the scale of roughness as otherwise it could equally well be related to structures at the atomic level or gross features several millimetres in size. Every machined part will deviate from its ideal shape. These deviations can be expressed by errors of FORM, WAVINESS and SURFACE ROUHGHNESS. This convention in nomenclature is fairly universally accepted.

Generally, the error of form refers to the long wavelength components of the surface, and it is caused by setting the machine tool for the wrong geometrical dimensions of the part to be manufactured. Usually waviness is considered an error of form due to an incorrect geometry of the machine tool producing the component and has an intermediate spatial wavelength and is as such outside the field of surface roughness. Surface roughness may be defined as the irregularities which are an inevitable consequence of the process if carried out on a geometrically perfect machine [Reason 1960].

To be able to measure a surface in quantitative terms the boundary between roughness and waviness must be defined with sufficient precision to achieve the level of uncertainty required in the measurements. This can be done graphically as indicated in the national standards but for convenience a filter is used with the characteristics of a high-pass filtering process which excludes the long spatial wavelength components in a controlled manner. This filtering process is the most important of the essentially arbitrary conditions which have to be imposed to make surface roughness

measurable. It will be discussed in more detail later.

There is no generally accepted definition of the boundary between waviness and form though the subject is being actively pursued at the present time.

3.2. Surface roughness parameters.

There is an extensive range of possible parameters that can be used to typify surface roughness. Many of the quantities which occur in the literature of mathematical statistics or the theory of random processes are possible candidates. Only a small number of these are in common use.

There are strong reasons for restricting the number of parameters used for process control. The need to avoid overcomplicating shop-floor communication and to restrict proliferation of instrumentation are probably the most important. On the other hand, there is a continuing search for parameters that will correlate with the functional properties of surfaces.

The choice of parameters used in process control has arisen to some extent for historical reasons. Originally surface finish values were obtained from profile graphs with a planimeter. This is the major reason why the arithmetic average, R_a, is used rather than the standard deviation, R_q, as the principal parameter in current practice, although it was always possible to obtain the rms value electrically. A summary of the most encountered surface parameters are listed in section 3.2.1., with their definitions as published in

ISO 4287/1-1984. These parameters are expressed in terms of sampling length, l, and profile departures y(x). The sampling length is the length of the reference line used for identifying the irregularities characterising the surface roughness. The profile departure is the distance between a profile point and the reference line in the direction of measurement, i.e. perpendicular to the reference line. The various reference lines are discussed in section 3.2.2..

3.2.1. Summary of surface roughness parameters.

Symbol:	parameter name:	definition in words and mathematical expression:
R_a	Arithmetical mean deviation of the profile	The arithmetical mean of the absolute values of the profile departures within the sampling length l. (usually averaged over several consecutive sampling lengths) $R_a = \frac{1}{l}\int_0^l \lvert y(x) \rvert dx$
R_q	Root-mean-square deviation of the profile	The root-mean-square value of the profile departures within the sampling length l. $R_q = \sqrt{\frac{1}{l}\int_0^l y(x)^2 dx}$

continued

Δ_a	Arithmetical mean slope of the profile	The arithmetical mean of the absolute values of rate of change of the profile departures within the sampling length. $$\Delta_a = \frac{1}{\ell}\int_0^{\ell}\left\|\frac{dy}{dx}\right\|dx$$
Δ_q	Root-mean-square slope of the profile	The root-mean-square value of the rate of change of the profile departures within the sampling length. $$\Delta_q = \sqrt{\frac{1}{\ell}\int_0^{\ell}\left(\frac{dy}{dx}\right)^2 dx}$$
λ_a	Mean wavelength of the profile	2π times the ratio of the arithmetical mean deviation of the profile to the arithmetical mean slope of the profile $$\lambda_a = 2\pi\frac{R_a}{\Delta_a}$$
λ_q	Root-mean-square wavelength of the profile	2π times the ratio of the root-mean-square deviation of the profile to the root-mean-square slope $$\lambda_q = 2\pi\frac{R_q}{\Delta_q}$$

continued

S_k	Skewness	A measure of asymmetry of the distribution density of profile deviations given by the formula: $$S_k = \frac{1}{R_q^3} \cdot \frac{1}{\ell} \int_0^{\ell} y(x)^3 \, dx$$
t_p	Profile bearing length ratio	The ratio of the bearing length to the sampling length. (usually expressed in percent)

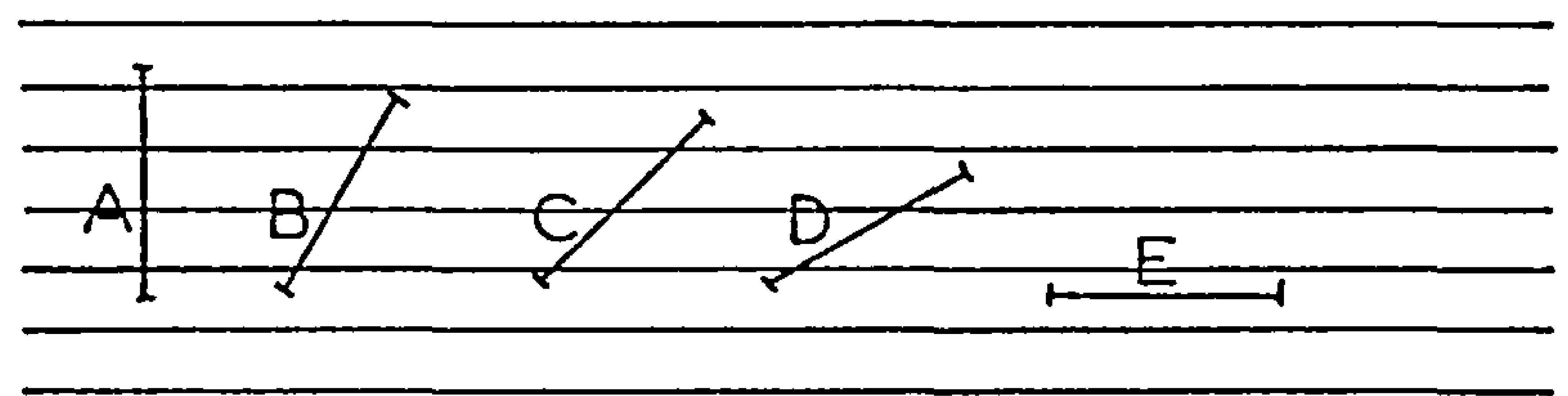
Since R_q is equal to one standard deviation of the profile about the mean line, a number of theoretical calculations of surface properties yield results which are naturally expressed in terms of R_q rather than R_a. R_q is therefore a more significant parameter in theoretical work. R_q is always greater than or equal to R_a.

The assessment length consists normally of several cut-off or sampling lengths, five consecutive cut-off lengths are taken as standard.

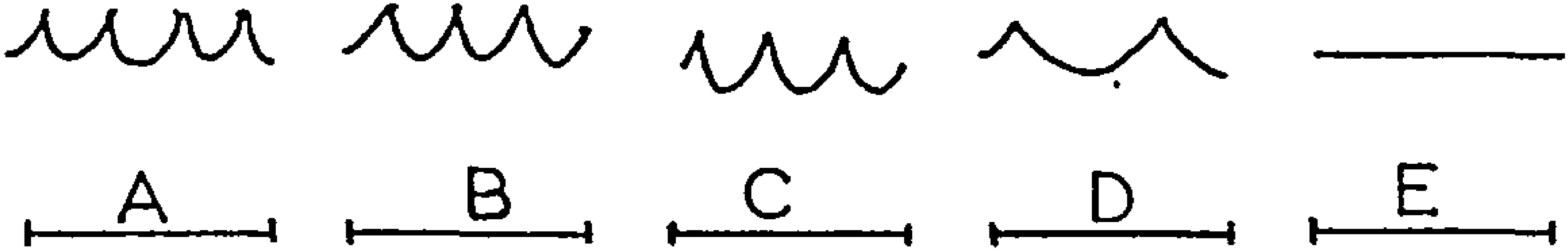
It should be noted especially that these parameters are defined in terms of profiles, and not surfaces, and either use a reference or are specifically related to a cut-off length. The relevance of this statement is that the statistics of a profile are not the same as the statistics of a surface. It is necessary to distinguish a peak on a profile from a summit on the surface [Williamson and Hunt 1978/68]. A profile will more often than not pass over the shoulder of an asperity on the surface instead of its summit. The shoulder will, nevertheless appear as a peak on the profile, though one of a reduced

height. Thus the profile indicates the presence of far fewer high peaks than actually exist on the surface. A similar error occurs in the determination of the mean surface gradient; it is not the same as the mean slope on a profile. Even if results are obtained by two-dimensional sampling the results are still dependent on the sampling method [Whitehouse and Philips 1985].

Furthermore the direction of measurement relative to the lay of the surface is important. On surfaces where the texture is multi-directional in character, the direction chosen for the purpose of measurement is usually immaterial but on turned, ground and other surfaces having definite directional scratches or toolmarks forming the texture, it will readily be recognised that totally different readings could be obtained from measurements made in different directions see fig. 3.1..



a. Showing various directions of measurement on a surface.



b. Profiles of the surface in the different derections shown in a.

Fig. 3.1. Influence of direction of measurement.

Even when all precautions are taken during surface finish

measurements, some considerable variations of the roughness parameters can be expected [Thomas & Charlton, 1981]. They suggest to make the measurements at the lowest possible range of the instrument, and a cut-off below 0.8 mm should not be used if possible.

The term "lay" is applied to the directional marks constituting the topography of the surface and it is generally accepted that the measurements are made approximately at right angles to the predominant direction of the lay.

The way in which the reference line and the cut-off are defined is of paramount importance to the realisation of the measurements.

3.2.2. Reference lines

Formally the reference line is given as the curve or straight line defined by convention in relation to the nominal profile or other datum, which serves as a reference for the definition of parameters. In practice there are three reference lines in use:

Electrical reference line: a line established by the filtering action of an instrument on the electrical signal which represents the measured surface profile.

Centreline: a reference line representing the form of the traced profile and parallel to its general direction throughout the cut-off length, such that the sums of the

areas contained between it and those parts which lie on each side of it, are equal.

Least-squares centre line: a reference line having the form of the nominal profile within the limits of the cut-off, and placed so that within the cut-off the sum of the squares of the deviations of the profile from the reference line is a minimum.

The electrical reference line is usually the appropriate reference since it is built into most electronic stylus instruments. The centreline is defined because it can be used simply on graphical profile records with the use of a planimeter. The least squares reference line is readily implemented digitally.

3.2.3. Cut-off

In usage cut-off has two meanings:

i the value of the wavelength , defined by convention as the longest wavelength transmitted by an instrument, i.e. in the signal used to assess the roughness.

ii that length of the profile over which the assessment of the surface texture is a reliable estimate.

3.2.4. Contact profiling instruments.

As mentioned in section 3., national standards pay almost exclusive attention to the stylus technique of measurement [Whitehouse 1974d]. Contact profiling devices using a sharp stylus attached to a transducer to trace profiles of a surface are by far the most common instruments. Magnifications in commercial instruments go as high as 500 000, and there are also low magnification instruments. In their simplest form, stylus instruments give a measure of roughness in terms of the arithmetic mean deviation , R_a, of the profile from a reference line. They may also have provision for producing a graphical record of the trace.

There is a class of stylus instruments which can be used to obtain profiles with very high vertical resolutions. With these, magnifications of 2 000 000 are obtainable, giving resolutions of better than 2 nm. They have specialised uses such as the measurement of optical finish and can also be used for measuring the thickness of thin films.

To ensure reproducibility of results and consistency, it is necessary to standardise a number of features of instruments in an arbitrary way. These features are the stylus radius and force, the minimum traversing length and the all-important bandwidth or cut-off. These features are specified in the national standards. The essential instrument characteristics are dealt with in the following sections.

Measuring traverse length.

This needs to be specified to ensure that the signal is present for sufficient time for the filtering circuits to operate without significant switching disturbance. An additional requirement is that the longer the signal the more representative it will be. The longest wavelength that can be analysed in the signal is also related to this length. The measuring traverse length of an instrument can be readily measured.

Transmission characteristics.

The instrument bandpass or cut-off is of critical importance and is closely specified in the standards. The standard instrument long wavelength filter has characteristics corresponding to the attenuation properties of two independent RC networks of equal time constant. The circuit of this type of filter is illustrated in fig. 3.2..

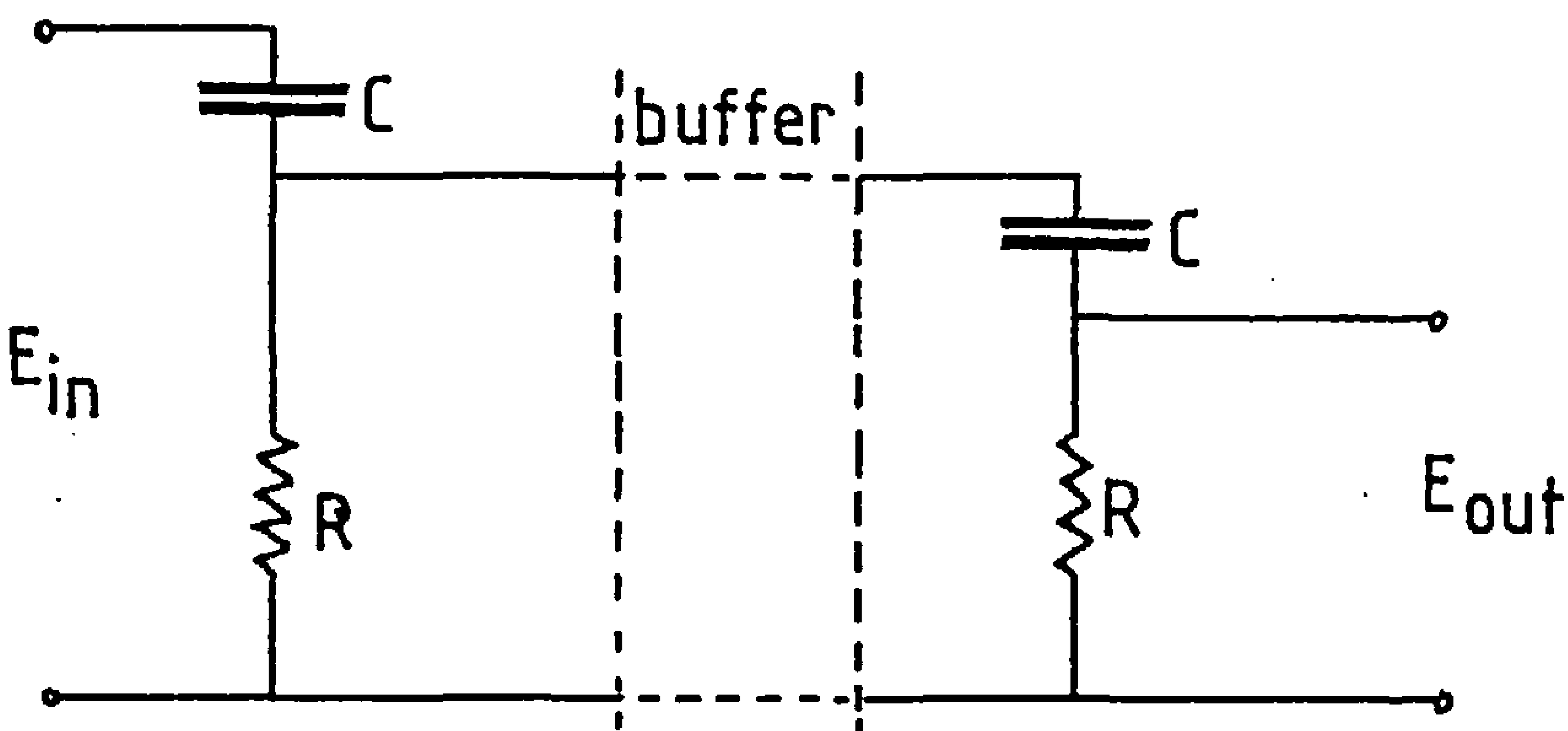


Fig. 3.2. 2RC filter network.

The long wavelength roughness cut-off (B_{max}) is assessed at an amplitude transmission of 75% of the maximum transmission and is chosen from the following range of values: 0.08, 0.25, 0.8, 2.5 and 8.0 mm.

An instrument is considered to have satisfactory cut-off if the specified value lies within 70 to 80% of the maximum transmission. A full description of the other filtering requirements for instruments can be found in section 2 of BS1134: part 1.

While it is imperative that the filtering meets the requirements of the specifications, it is not an easy matter to fully test them in assembled commercial instruments from basic principles, however producing filters within specification is not normally considered a problem. The filters in standard instruments are technically primitive and by their nature introduce major phase distortions in the region of the cut-off. While they could be replaced by more suitable active filters giving phase compensation [Whitehouse 1967/68], the process of obtaining international agreement on instrument specifications has been arduous and lengthy and the need to ensure continued uniformity has inhibited the adoption of more advanced designs. If filtering other than the standard RC type is used, say digital filtering, it needs to be carefully specified and the results interpreted accordingly.

Stylus shape and force.

It is important that the stylus shape be within the tolerances laid down and since the stylus is subject to wear, it needs to be checked from time to time. Hillmann working at the Physikalisch-Technische Bundesanstalt (PTB), FRG, has explored stylus tip geometry with great thoroughness [Hillmann 1980].

The checking is usually performed using a shadow projector or a visual microscope. In both cases high resolving powers are required. With projectors, even the best quality instruments at high magnification are just adequate, while with microscopes a magnification of at least 100 is necessary and a high numerical aperture is required.

An alternative method of checking the stylus is to use a specially fine grating, a "stylus check", which gives a specified value of a roughness parameter only when the stylus is not too badly worn. When the stylus is worn it does not penetrate the structure of the grating fully and a low reading results. This method has been advocated for many years but as yet is not widely used because suitable gratings are not readily obtainable.

Scanning electron microscopes have been used to examine the stylus but the surface need normally to be coated with a conductive layer and this is a major draw-back. If the stylus arm cannot be fitted into the microscope a replica in gold has been used.

If the force of the stylus is too great it will scratch the surface of quite hard materials, if too low it may bounce. Since the tip radius is small very high pressures can be obtained at small

loads. The specified forces range from 0.7 mN for a radius of 2 um, to 16 mN for a radius of 10 um with additional restrictions on the rate of change of the measuring force. A beam balance or a torsional balance or other sensitive force measuring device is required to make a check of these forces. Because of the type of construction used it is unlikely that the values specified will be exceeded.

Skid dimensions.

Very often stylus instruments are used with a skid or shoe fitted to the stylus arm. The skid acts as a support of the stylus arm and greatly reduces the sensitivity of the instrument to vibration. Provided the radius on the skid surface is not too small, the use of a skid will not drastically affect the transmission characteristics of an instrument. For this reason the skid radius is required to be at least 50 times greater than the cut-off being used. Also important is skid stylus distance, especially for periodic surfaces.

3.2.5 Instrument calibration.

Profilometers are essentially displacement measuring devices and require calibration at regular intervals to maintain their accuracy. The two functions, profile recording and parameter measurement, are often calibrated separately even though both derive their signals

from the stylus transducer.

The means used to perform this calibration depends upon the magnification and the accuracy required. In most cases, calibration is performed on one or two ranges and the instrument attenuators are relied upon for the other ranges.

In most cases the calibrations are done using calibrated grooves and R_a standards which are often supplied with the instrument. In some cases the instrument manufacturers advocate calibration of the transducer only and rely on the stability of the processing electronics to give correct parameter values.

The grooves and R_a standards in turn need calibration and this is achieved using a calibrated instrument or alternatively in the case of grooves by optical interference using an interference microscope.

The instrument transducer may also be calibrated using gauge blocks at the low magnifications or gauge blocks and a lever arm at higher magnifications. Further details relating to calibration techniques will be found in BS1134: part 2. R_a standards are calibrated by taking the direct readings from a calibrated master instrument evaluating graphical records of profiles by planimetry, or digitally recording profiles and computing the appropriate parameter value.

4. Parameters obtained via powerspectral density functions.

In the previous chapter the surface roughness parameters were obtained by direct processing of surface profiles. There is, however, an alternative method to obtain some of the surface parameters, namely by using the power spectral density function (PSDF) of the surface profile. At first glance it might seem a detour to transform the profile data into a PSDF and then to extract from this the surface roughness parameters, since it is possible to calculate the parameters directly from the surface profile data. Under certain conditions, however, it is only possible to obtain the PSDF of the surface by optical Fourier transform techniques, without being able to produce a surface profile. From there the necessity to be able to extract the roughness parameters from the power spectrum. Most of the reported relationships apply to random Gaussian surfaces. Deterministic surfaces can be dealt with in a simpler fashion. The results are obtained by using the techniques of random process theory. The height of a rough surface may be considered to be a two-dimensional random variable, with the Cartesian coördinates in a reference plane being the independent variables. This approach was first used by Longuet-Higgins in 1957 in studies of random ocean surfaces [Longuet-Higgins 1957a,b]. Some of his results are reported in a later section.

4.1. Characterisation of random surface profiles.

4.1.1. The autocorrelation function ACF.

Consider a surface profile whose height above a reference line is y(x), where y(x) is a random variable of the ordinate x on the reference line. The mean line of the surface profile is choosen as the reference line. It is assumed that the surface is homogeneous, that is, its statistical description is invariant with respect to translation along the reference line. The ACF is then defined by:

$$R(\alpha) = \lim_{L \to \infty} \frac{1}{L} \int_0^L y(x)\, y(x+\alpha)\, dx \qquad (4.1)$$

This function $R(\alpha)$ describes the general dependence of the heights of the surface profile at one position on the heights at another position. The quantity $R(\alpha)$ is always a real-valued even function with a maximum at $\alpha=0$, and may either be positive or negative. In equation form:

$$R(\alpha) = R(-\alpha) \quad \text{and} \qquad (4.2a)$$

$$R(0) \geq |R(\alpha)| \quad \text{for all } \alpha \qquad (4.2b)$$

From equation (4.1) it follows that the mean square value of y(x) is expressed by:

$$R_q^2 = R(0) \qquad (4.3)$$

That is, the square of the rms roughness parameter $,R_q,$ equals the ACF with zero displacement.

4.1.2. The power spectral density function PSDF.

The Fourier transform of $R(\alpha)$ is called the PSDF $P(\omega)$. It is defined by:

$$P(\omega)=\frac{1}{2\pi}\int_{-\infty}^{\infty} R(\alpha)\, e^{-i\alpha\omega}\, d\alpha \tag{4.4a}$$

The quantity $P(\omega)$ is always a real valued, non negative even function. Since $R(\alpha)$ is an even function, (4.4a) can be rewritten as:

$$P(\omega)=\frac{1}{2\pi}\int_{-\infty}^{\infty} R(\alpha)\cos(\alpha\omega)\, d\alpha \tag{4.4b}$$

The inverse Fourier relation holds:

$$R(\alpha)=\int_{-\infty}^{\infty} P(\omega)\, e^{i\alpha\omega}\, d\omega \tag{4.5}$$

From the equations (4.3) and (4.5) it follows that

$$R_q^2=R(0)=\int_{-\infty}^{\infty} P(\omega)\, d\omega \tag{4.6}$$

Equation (4.6) indicates that $P(\omega)$ is a decomposition of R_q into contributions from various spectral components, which are waves with a spatial wavelength $\lambda=2\pi/\omega$.

4.1.3. Moments of the PSDF.

The moments of the PSDF $P(\omega)$ of a surface profile are defined as follows:

$$m_n = \int_{-\infty}^{\infty} \omega^n P(\omega)\, d\omega \tag{4.7}$$

To obtain a large number of useful statistics of the surface profile via the PSDF, its zeroth, second and fourth moment have to be calculated. As shown previously, the zeroth moment equals the mean square or variance of the roughness of the profile. The second and fourth moment of the PSDF can be related to the standard deviation of the profile slopes (σ_s) and profile curvatures (σ_k) respectively, as demonstrated in the following. Starting from the relationship between the ACF and PSDF:

$$R(\alpha) = \lim_{L\to\infty} \frac{1}{2L} \int_{-L}^{L} y(x)\, y(x+\alpha)\, dx = \int_{-\infty}^{\infty} P(\omega)\, e^{i\alpha\omega}\, d\omega \tag{4.8}$$

Taking the second and fourth derivatives with respect to α of this equation yields:

$$-\int_{-\infty}^{\infty} \omega^2 P(\omega)\, e^{i\alpha\omega}\, d\omega = \lim_{L\to\infty} \frac{1}{2L} \int_{-L}^{L} y(x) \frac{\partial^2}{\partial\alpha^2}\, y(x+\alpha)\, dx \tag{4.9a}$$

$$\int_{-\infty}^{\infty} \omega^4 P(\omega)\, e^{i\alpha\omega}\, d\omega = \lim_{L\to\infty} \frac{1}{2L} \int_{-L}^{L} y(x) \frac{\partial^4}{\partial\alpha^4}\, y(x+\alpha)\, dx \tag{4.9b}$$

Using $\frac{\partial^n}{\partial\alpha^n} y() = \frac{\partial^n}{\partial x^n} y()$ and subsequently substitution of $\alpha=0$ leads to:

$$-m_2 = -\int_{-\infty}^{\infty} \omega^2 P(\omega)\, d\omega = \lim_{L\to\infty} \frac{1}{2L} \int_{-L}^{L} y(x) \frac{\partial^2}{\partial x^2} y(x)\, dx \quad (4.10a)$$

$$m_4 = \int_{\infty}^{\infty} \omega^4 P(\omega) d\omega = \lim_{L\to\infty} \frac{1}{2L} \int_{-L}^{L} y(x) \frac{\partial^4}{\partial x^4} y(x) dx \quad (4.10b)$$

By integration by parts and under the condition that y(x) and its first four derivatives have finite values over the interval $-L<x<L$, the equations can be replaced by:

$$m_2 = \lim_{L\to\infty} \frac{1}{2L} \int_{-L}^{L} \left(\frac{dy}{dx}\right)^2 dx = \sigma_d^2 \quad (4.11a)$$

$$m_4 = \lim_{L\to\infty} \frac{1}{2L} \int_{-\infty}^{\infty} \left(\frac{d^2y}{dx^2}\right)^2 dx = \sigma_k^2 \quad (4.11b)$$

The relationships (4.6), (4.11a) and (4.11b) are valid for deterministic and random surfaces. From the equations (4.6) and (4.11a) the rms wavelength parameter can derived. By definition, the mean square angular frequency of a waveform is given by [Spragg & Whitehouse 1970/71]:

$$\omega_{rms} = \int_{\infty}^{\infty} \omega^2 P(\omega) d\omega \Big/ \int_{-\infty}^{\infty} P(\omega) d\omega = (\lambda_q / 2\pi)^2 \quad (4.12)$$

from this together with (4.6) and (4.11a) it follows that

$$\lambda_q = 2\pi R_q / \Delta_q \quad (4.13)$$

For random, isotropic Gaussian surface profiles, however, Longuet-Higgins has shown that the densities of zeroes ,D_z, and extrema ,D_{ex}, (maxima and minima) along the profile can be expressed

as functions of the first three even moments of the PSDF, namely:

$$D_2 = \frac{1}{\pi}\left(\frac{m_2}{m_0}\right)^{\frac{1}{2}} \tag{4.14}$$

$$D_{ex} = \frac{1}{\pi}\left(\frac{m_4}{m_2}\right)^{\frac{1}{2}} \tag{4.15}$$

In passing, it can be noted that the density of peaks along the profile is, due to symmetry, one half of the number of extrema:

$$D_{peak} = \frac{1}{2\pi}\left(\frac{m_4}{m_2}\right)^{\frac{1}{2}} \tag{4.16}$$

4.2. Separation of the components of the surface roughness.

Most surfaces are not purely random and can be regarded as an additive composition of periodic and random components [Staufert 1979]. This is based on an assumption concerning the physical process of surface generation. Periodicity is introduced by copying the tool on the workpiece or by vibration of the tool. Randomness is introduced by the detachment of material. If both processes can be regarded as physically independent, the assumption of additivity is valid. Although the random and periodic components can be regarded as an additive combination, they cannot be separated by linear filtering because their wavelengths fall in the same band. Inspection of an amplitude spectrum of a turned profile shows that a pattern can be

recognised (fig 4.1). There always exists a decreasing relatively broad band of amplitudes fluctuating around a mean line. This band of amplitudes is the spectrum of the random part of the profile. If periodicity is present sharp peaks higher than the level of random fluctuations appear in the spectrum. These peaks are the harmonics of the periodic spectrum.

From this it follows that the amplitude or powerspectrum is a powerful method of analysing the nature of the surface profile. It also allows the specification of the periodic component by the amplitude and wavelength and the random component by bandwidth and R_q values.

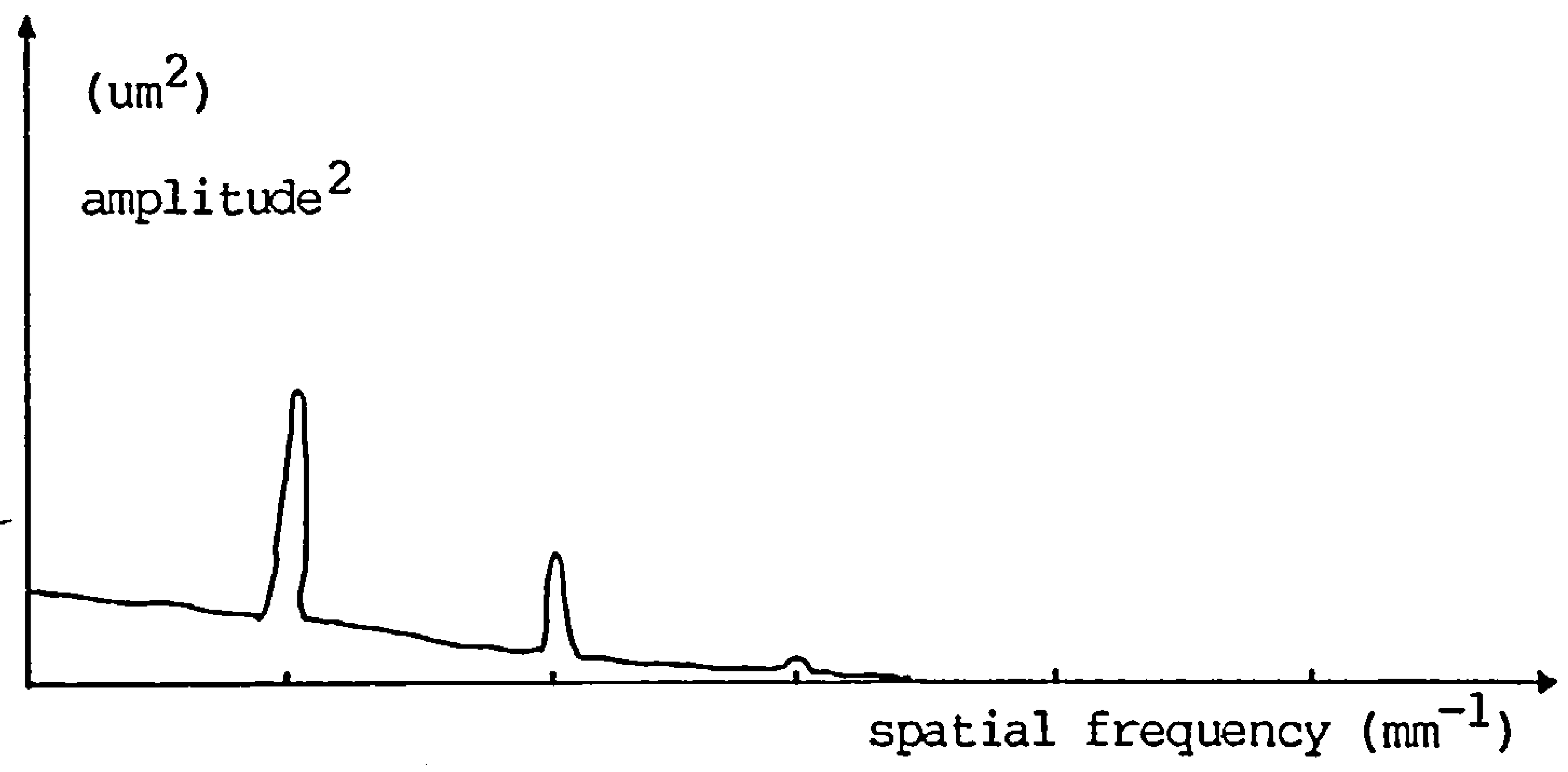


Fig. 4.1. Powerspectrum showing periodic and random contributions.

5. The generation of surfaces and their microtopographical properties.

Surfaces are generated by manufacturing processes. The manufacturing operations can be catalogued under the following headings:

1 refining and general preparation of the raw material.

2 shaping of the material:

i casting.

ii forming by plastic deformation and shear cutting.

iii machining.

3 fabrication- welding, adhesive bonding etc.

4 assembly.

5 cleaning, polishing, plating, coating etc.

The operations 2 and 5 are the processes which will produce engineering surfaces . From these processes the machining operations are claimed to be the most important . This is based on the fact that some machining is involved in the production of almost any item one can think of. Even when machining is not directly involved, it is virtually always a necessary process in making the dies or fixtures used in production operations.

Since the late 1940's, the whole approach to the shaping process has undergone a number of revisions. Casting to final shape has become reasonably common place, accurate forming has been accepted,

with considerable development in processes such as spinning, extrusion and high rate forming. These developments in the so-called chipless production processes have been said to herald the end of material removal processes. In fact this does not seem likely. There are situations where the non chip forming applications are economically desirable, but for a wide range of conditions the simplicity of the machining process will make it an economical method for years to come. In keeping with the forming processes great developments have been made in material removal processes in recent years. Tool materials have been improved. The size and speed of machine tools have increased considerably. Machine tool control has been improved. Many new processes have been introduced, such as ultrasonic machining, electrospark cutting, electrolytic machining, chemical milling and cutting with electron,ion and laser beams.

A brief description of these machining processes [Yankee,1979], except for turning, can be found in appendix A0.

5.1. Turning.

Turning may be defined as a machining process by which cylindrical, conical, or irregularly shaped external or internal surfaces are produced on a rotating workpiece. The cutting action is generated by one or more stationary single-pointed cutting tools which are held at an angle to the axis of the workpiece rotation. Turning operations are most commonly associated with a machine tool known as a lathe.

Description of the process.

Figure 5.1. depicts how the shape of the desired workpiece surface is determined by the shape and size of the stationary cutting tool, which is being fed inwardly but without longitudinal feed.

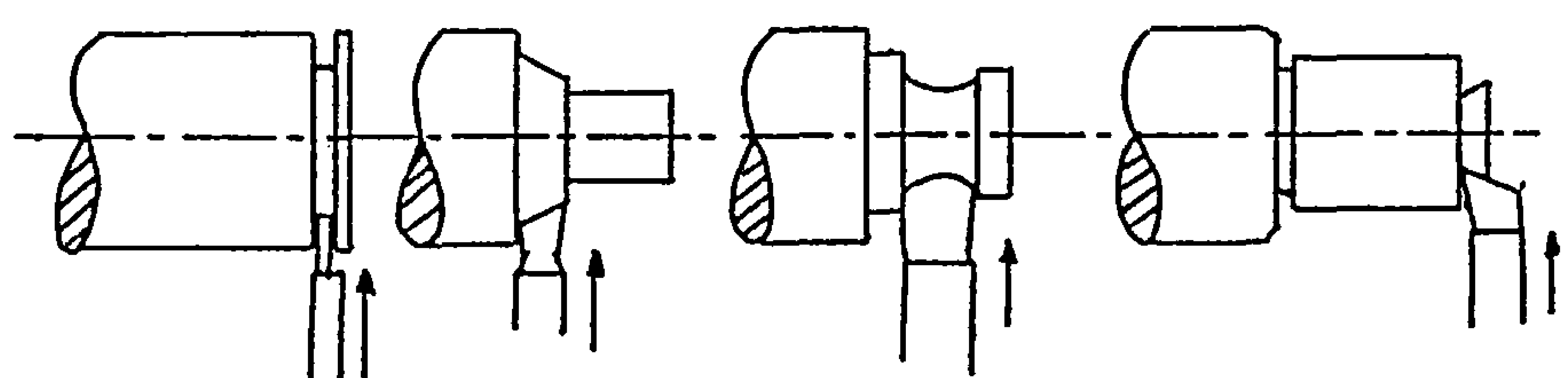


Figure 5.1. Typical workpiece surfaces produced by in-feeding.

Speed and feed are important cutting factors. Speed is a term that describes the cutting speed or velocity of the rotating workpiece as it moves past the cutting edge of the tool. The measurement of velocity is in units of surface meters per minute or (smpm). Spindle speed, on the other hand in units of revolutions per minute (rpm) and may range from 10 to 2000 rpm. Turning feed is the rate of advance of the cutting tool per revolution of the spindle. Feeds may range from 0.03 to 2.0 mm per revolution. Considerable study has been devoted to the determination of appropriate turning feeds and speeds, with the result that there now are tables available in manufacturer's catalogues and in various machining handbooks which list recommended speeds and feeds for practically any classification of work. Such tables usually relate such variables as workpiece hardness, use of coolant, type of operation, and tool-bit material. In some plants selections of speed and feed is often heavily

influenced by experience with similar parts or from previous production runs.

Depth of cut is a term that denotes the distance to which the tool enters the work. The thickness of the metal chip thus removed is equivalent to the depth of cut.

Tool bits.

Most metal lathe-turning operations employ single-point cutting tool bits which are ground to cut in only one direction. For proper cutting action to take place, the cutting edge must contact the workpiece before any other parts of the tool bit do. For high quality turning, diamond tooltips are used which are relapped after dulling.

Lathe operations

Lathes are the principal machine tools used for the vast majority of turning operations. Many other machining operations are possible on lathes. These include facing, parting, necking, drilling, boring, undercutting, reaming, chamfering, threading, tapping, and, by the use of special attachments, additional operations, such as taper turning and boring, grinding, and milling can be performed. It should be noted that two or more similar or dissimilar lathe operations may be simultaneously performed. For example, the outside diameter of a workpiece may be turned while a hole is being drilled or reamed in its centre.

Types of turning machines.

Since the optical sensor for surface finish was initially developed for diamond turning processes, a more detailed discription of a lathe will be given.

There are many kinds of workpiece-rotating machine tools. Close examination reveals a remarkable similarity in the manner by which all lathes operate.

The basic structure of a lathe may be coveniently divided into five main sections: the headstock, tailstock, bed, carriage and the gearbox. These sections are shown in figure 5.2. and will be

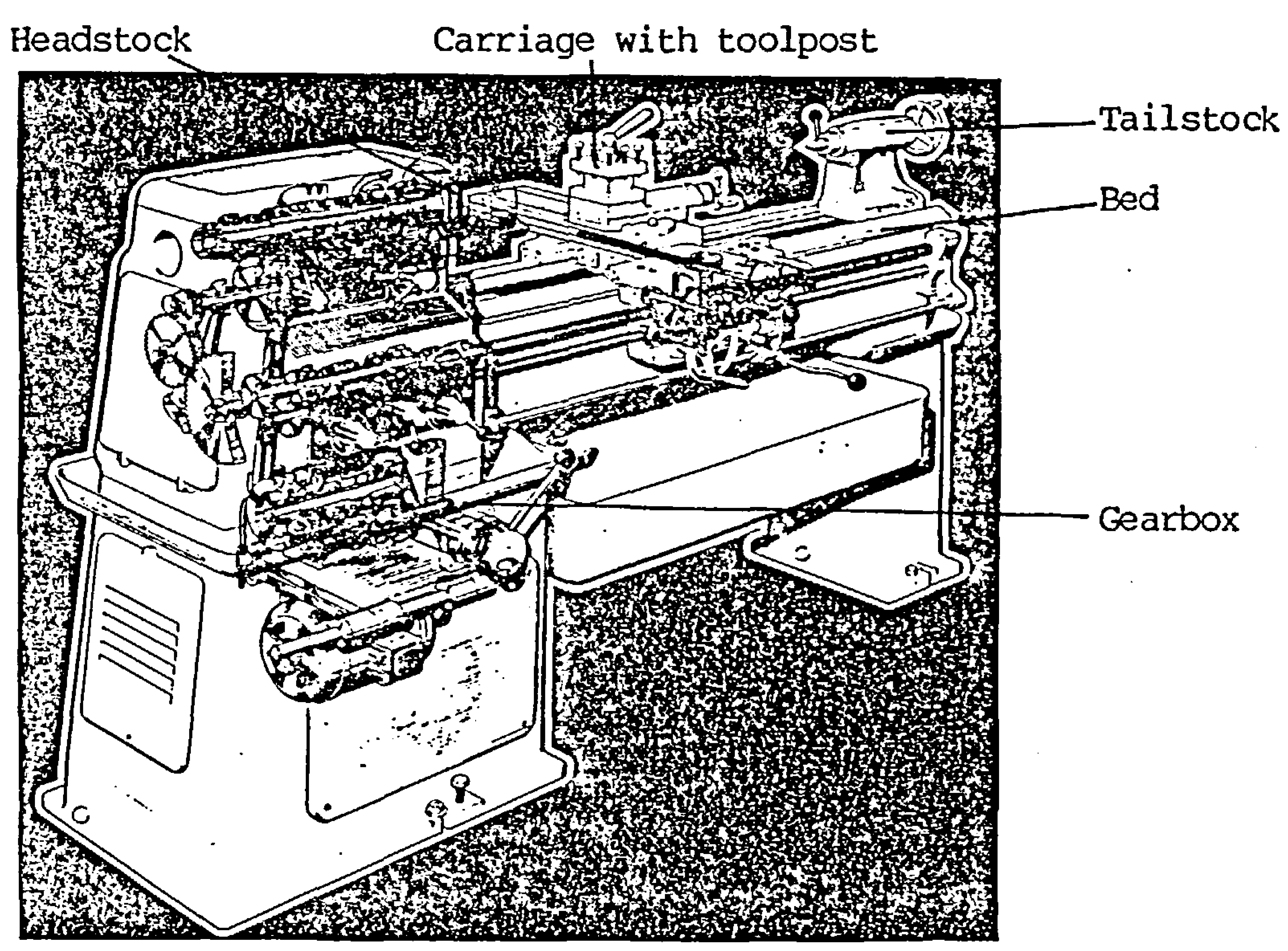


Figure 5.2. Basic structure of a lathe.

discussed together with certain related effects on surface finish.

The headstock.

The headstock is mounted in a fixed position. It houses and supports the driving mechanism which drives the spindle. The headstock spindle is hollow to permit the loading of tubular or shaftlike workpieces. Workholding attachments such as a driving plate, faceplate, or various types of chucks may be mounted on the threaded spindle nose.

Any motion float of the headstock spindle will cause variations in the feed rate, which will manifest itself in the powerspectrum of the surface profile as sub-harmonics of the feed frequency.

Spindle bearing wear produces a random and a low frequency component in the normally periodic profile.

The tailstock.

The tailstock can be adjusted along the bed of the lathe and clamped in any desired position to accomodate different length of stock. A hardened steel part known as a "dead centre", or a rotating part known as a "live centre" is inserted into the tapered hole in the hollow tailstock quill. Alignment is made with a corresponding point on another centre, which is similarly held in the headstock spindle. Work held in this manner is said to be in position "between centres". Tools such as drills and reamers fitted to the tailstock are controlled by rotating the tailstock handwheel.

Any errors in the tailstock will produce similar effects in the surface profile as mentioned for the headstock.

The bed.

The bed is the central foundation upon which the lathe is built. This rigid structure serves as a mounting frame for the headstock, the tailstock, the carriage and the gearbox. Two sets of accurately machined parallel slideways on the top of the bed guide the carriage and align the headstock and tailstock.

Inaccuracies of the machine bed will cause the tool to deviate from a rectilinear motion, thereby causing waviness in the surface profile. This is represented by a low frequency in the powerspectrum.

The carriage.

The carriage supports and controls the action of the cutting tool, which is firmly held in a tool post. The carriage can be moved longitudinally by manual or power feed. A cross slide, mounted on top of the carriage, provides cutting tool motion which is at right angles to the axis of the workpiece. The cross feed movement may also be controlled by manual or by power feed. The compound rest which is mounted above the cross slide can be swivelled in a horizontal plane. In this manner, the cutting tool can be positioned and set at any desired angle to the axis of the workpiece. In and out feed adjustments may be made by manually turning the compound feed-rest knob, which has a graduated dial.

Toolpost instabilities cause the tool to move in the axial, radial and tangential direction, thereby producing a combined effect of small changes in the feedrate and the depth of cut. This modifies

the ideal profile into a frequency and amplitude modulated profile.

The gearbox.

There are two gearboxes, one controlling the feed rate and the other one controlling the spindle speed.

Errors in the feed gearbox will produce a cyclic variation in the cutting tool motion, thereby producing a subharmonic in the power spectrum.

The headstock gearbox can introduce vibrations in the machine-tool causing chatter marks on the surface.

Analysis of the power-spectrum of turned surfaces.

It is interesting to observe that:

i Malfunctioning of the machine can contribute only to those errors in the surface profile which occur at a maximum spatial frequency equal to the feed frequency. The feed frequency is defined as the reciprocal of the feed mark spacing.

ii The surface profile produced by a perfect machine and tool combination occupies a limited range of frequencies with the feed frequency as the lower bound. This can be explained by decomposing the ideal surface profile y(x)

in its Fourier components as expressed by:

$$y(x) = \sum_{n=1}^{\infty} a_n \cdot \cos(2\pi n x/D + \phi_n) \tag{5.1}$$

where D is the feed mark spacing,

a_n is the amplitude of the nth component with a spatial frequency (n/D).

Since the toolshape is generally smooth, the amplitudes a_n will decay rapidly with increasing n (frequency).

iii Since the toolshape is reproduced every feed mark spacing, the mark caused by a tool chip will also occur at the feed frequency. Because chip marks have sharp edges, its Fourier decomposition will contain frequency components ranging from the feed frequency up to a higher limit range than the toolshape. Since the amplitudes of the chipmarks are smaller than the toolshape ones, the toolchipping Fourier components will be dominant in the high frequency range where the toolshape have decayed to zero.

Consequently, the amplitude- or power-spectrum can be divided in three frequency ranges. Each range contains information about one aspect of the turning process. The low frequecies up to the feed frequency is associated with the machine condition. The mid-frequencies contain information about the nominal toolshape and the high frequencies about the tooltip condition. This is graphically

represented in figure 5.3..

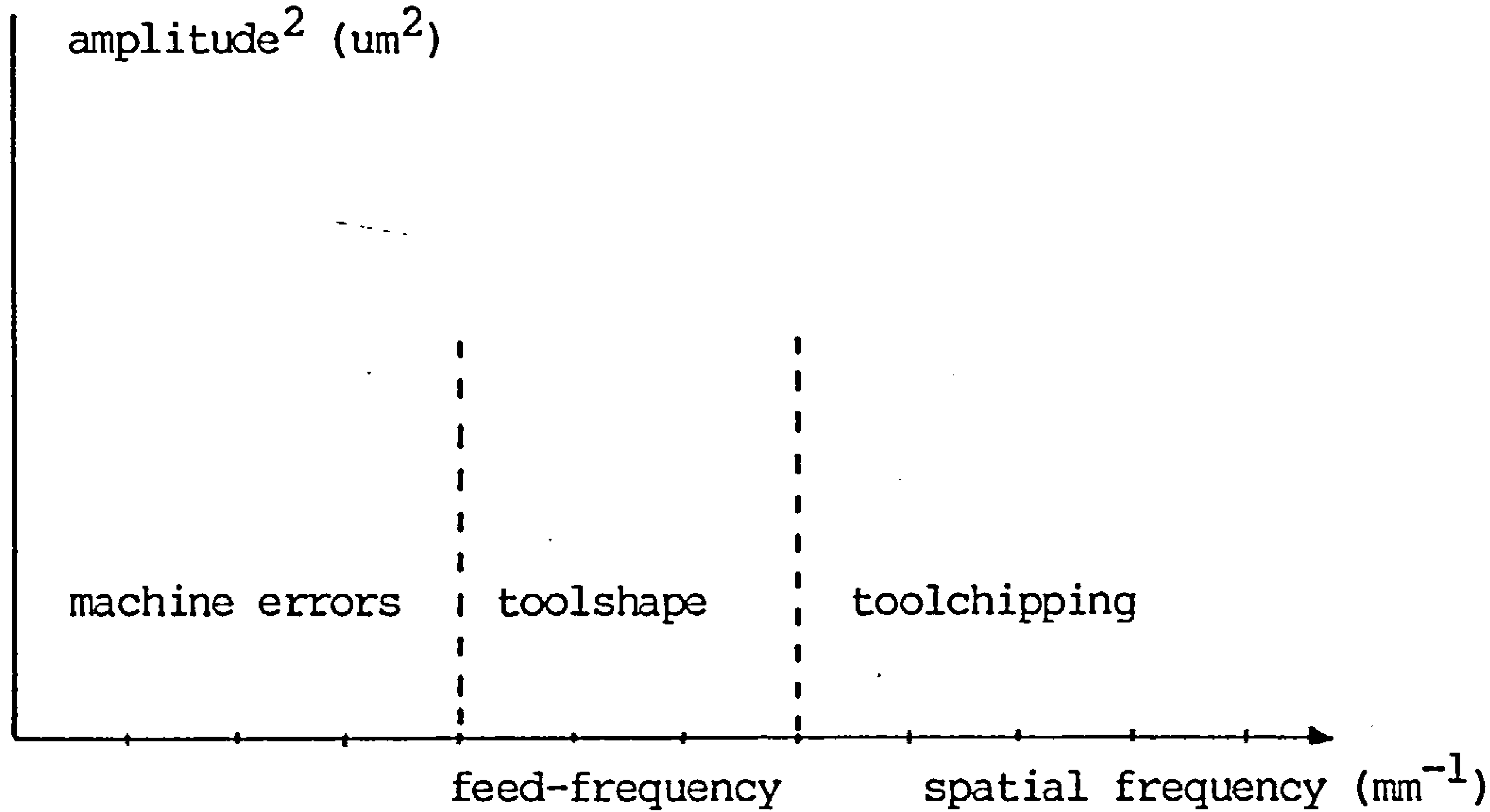


Figure 5.3. Frequency bands as related to the turning process.

5.2. Relationship between surface finishes obtained by the various machining processes.

Table 5.1. shows surface roughness values that can be expected from various common place production processes. In general each finishing process has its characteristic surface texture. These types of surface textures can be classified as indicated in BS 1134 part 1, see figure 5.5.. A further distinction can be made between the nature of the surface profiles. They can be periodic or random. In Table 5.2. the machining processes are listed together with the surface characteristics, using the symbols as defined in figure 5.5.. Also the type of the surface profile is stated. The profiles are assumed to be taken across the surface lay. The machining process may be stated in conjunction with the R_a value and the sampling length, in order to define more fully the required surface. It can be concluded from Table 5.1. , that the manufacturing methods consisting of material removal, produce the better finishes. On the pages 49,50 and 51 are photographs of actual surfaces together with their surface profiles. These surfaces are produced by turning, grinding and milling respectively.

Table 5.1 : Typical roughness values obtained with ordinary materials and common production processes [Thwaite 1984]

MANUFACTURING METHOD		Arithmetic average surface roughness R_a in µm (1 µm = 0.001 mm)													
		0.006	0.012	0.025	0.05	0.1	0.2	0.4	0.8	1.6	3.2	6.3	12.5	25	50
Casting	Sand casting														
	Shell moulding														
	Die casting														
	Pressure die casting														
	Investment casting														
Reforming	Drop forging														
	Plain rolling														
	Deep-drawing of sheet														
	Impact extruding, extrusion moulding														
	Coining														
	Rolling of shapes														
Separating	Cutting														
	Longitudinal turning														
	Face turning														
	Recessing														
	Planing														
	Slotting														
	Shaving														
	Drilling														
	Boring														
	Countersinking, Counterboring														
	Reaming														
	Milling														
	Broaching														
	Filing														
	Plain longitudinal grinding														
	Plain surface grinding														
	Plain plunge grinding														
	Cylindrical grinding														
	Face grinding														
	Polishing														
	Tumbling														
	Long-stroke honing														
	Short-stroke honing														
	Cylindrical lapping														
	Surface lapping														
	Liquid honing														
	Polish lapping														
	Blasting														
	Barreling														
	Flame cutting														

Symbol	Interpretation
=	Parallel to the plane of projection of the view in which the symbol is used. (Direction of lay)
⊥	Perpendicular to the plane of projection of the view in which the symbol is used. (Direction of lay)
X	Crossed in two slant directions relative to the plane of projection of the view in which the symbol is used. (Direction of lay)
M	Multi-directional.
C	Approximately circular relative to the centre of the surface to which the symbol is applied.
R	Approximately radial relative to the centre of the surface to which the symbol is applied.

Figure 5.5. Symbols for the direction of lay. [BS 1134:Part 1: 1972]

Table 5.2. Surface texture of common machining processes, using the symbols from figure 5.5.

Process	surface texture	surface profile
Face milling	M	periodic
Slab milling	=	periodic
Circular pocket milling	R	periodic
Shaping	=	periodic
Planing	=	periodic
Broaching	=	random
Turning	=	periodic
Face turning	C	periodic
Grinding	= or X	random
Chemical milling	isotropic	random
Electrical discharge machining	isotropic	random
Electron beam machining	isotropic	random
Ion beam machining	isotropic	random

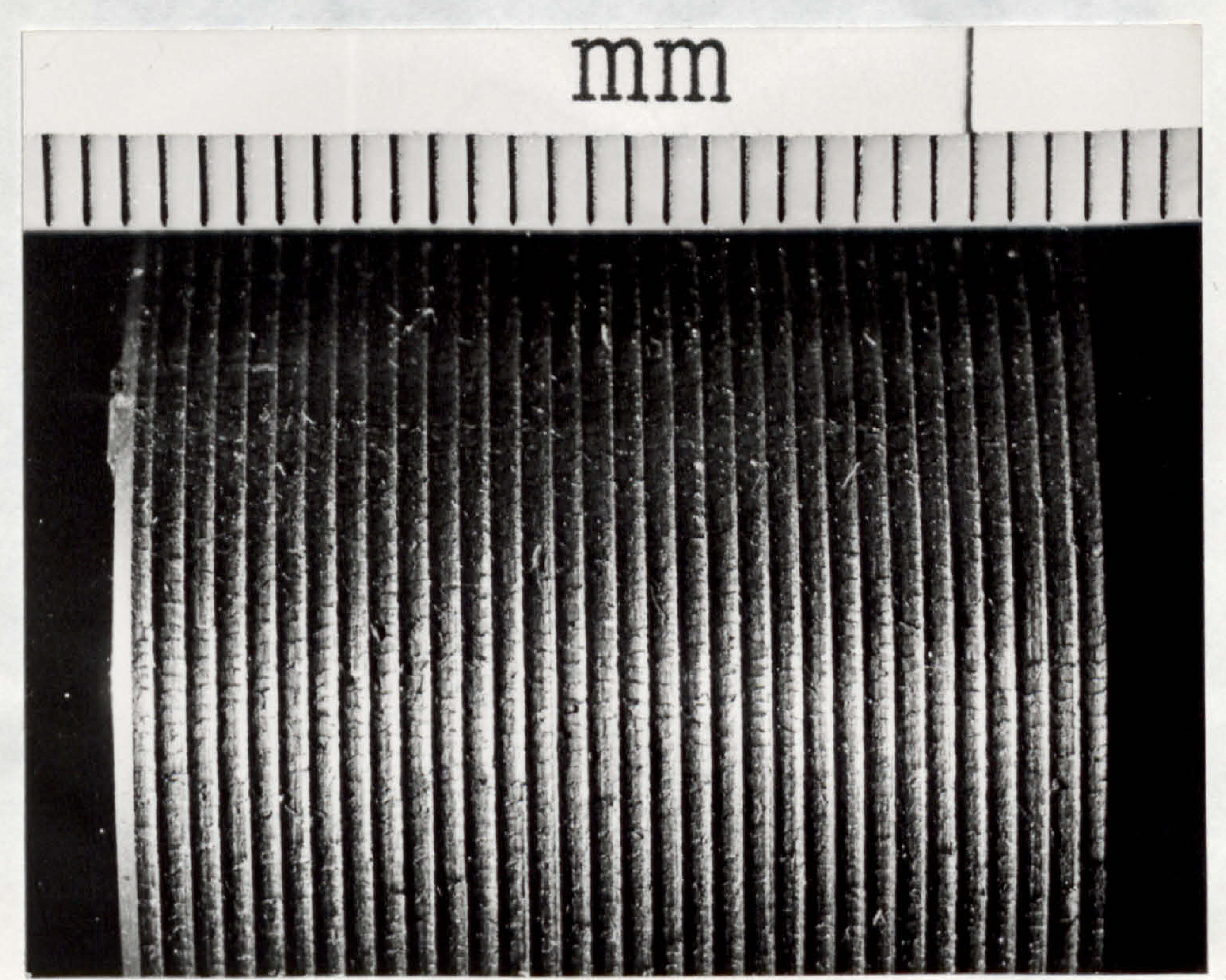
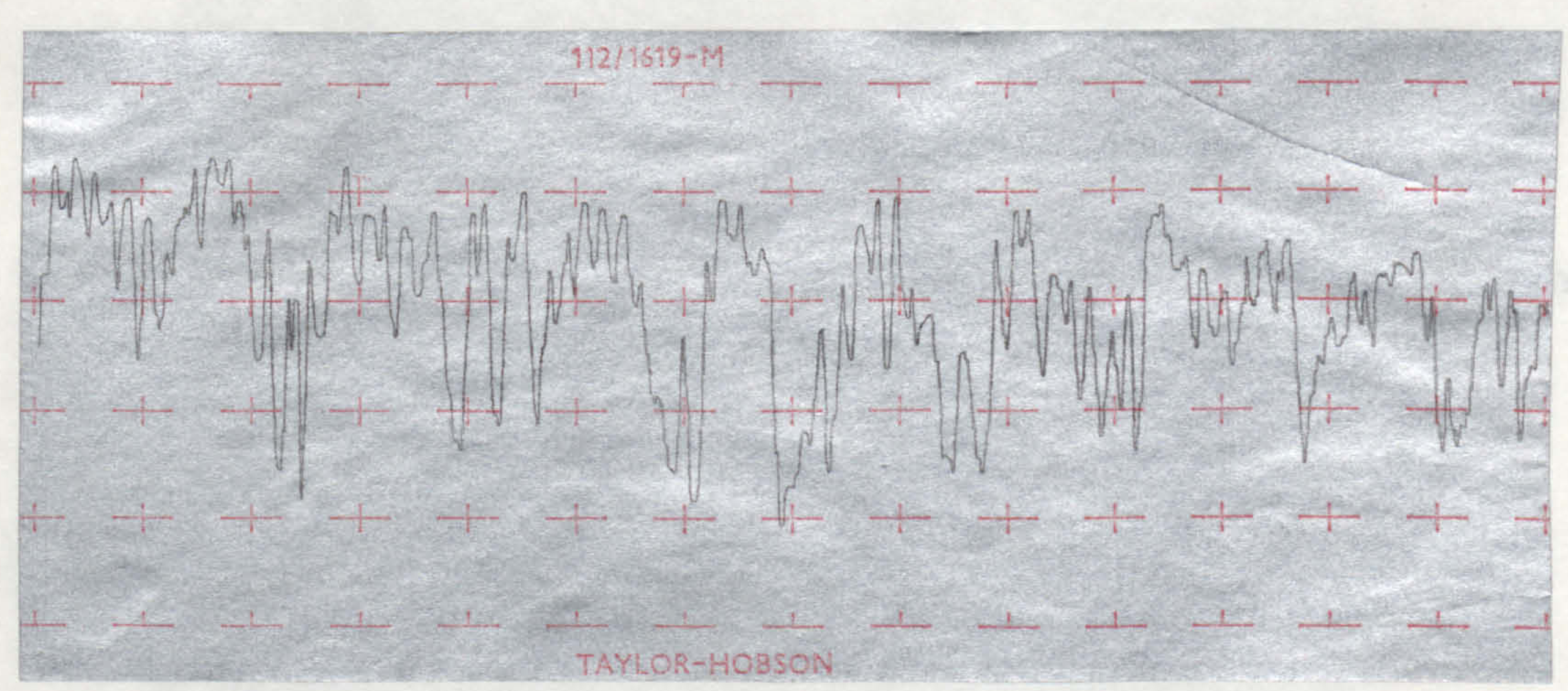


Photo 5.1. Turned surface.



Graph 5.1. Surface profile of the turned surface taken across the lay.

Magnifications: horizontal X20, vertical X2000

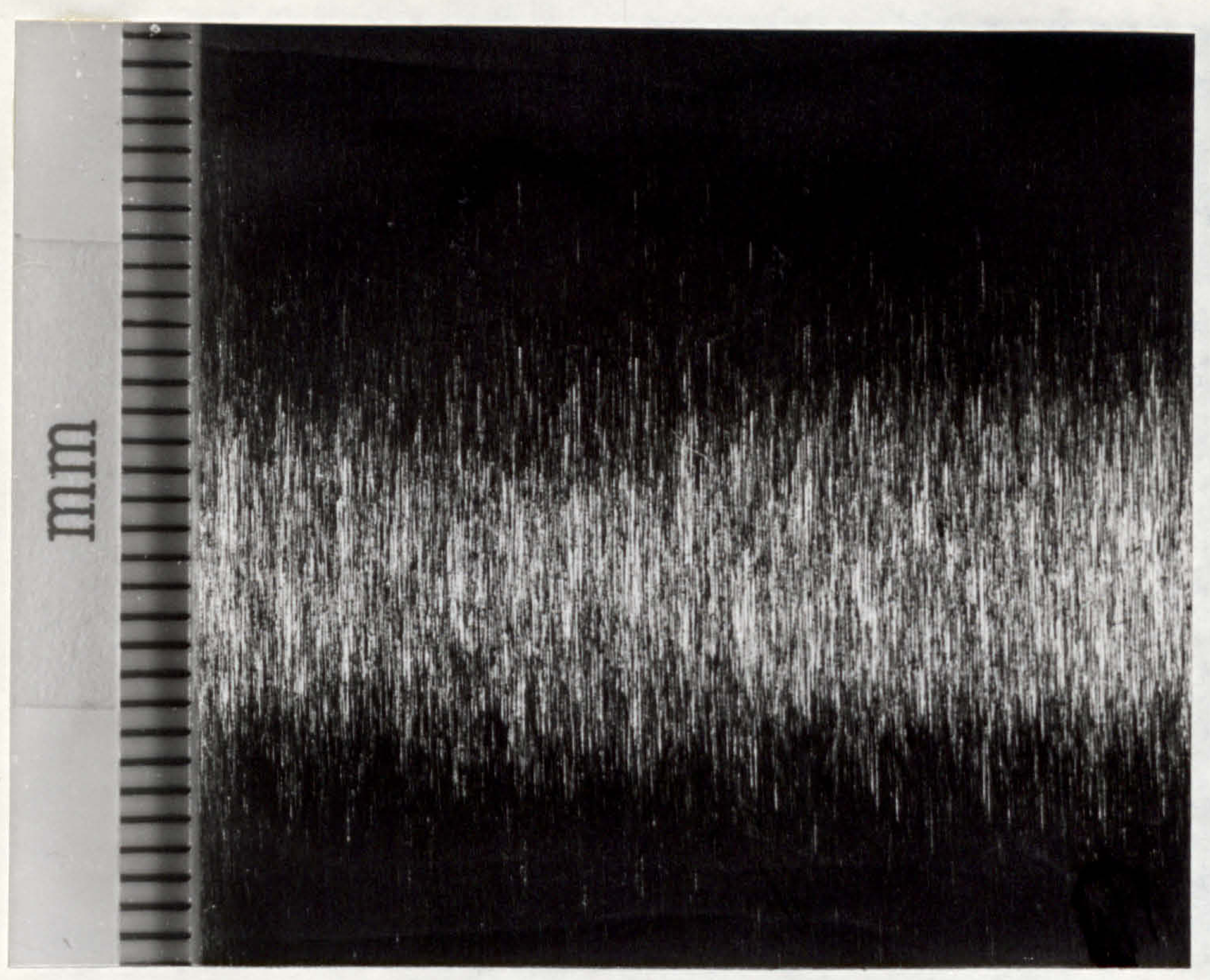
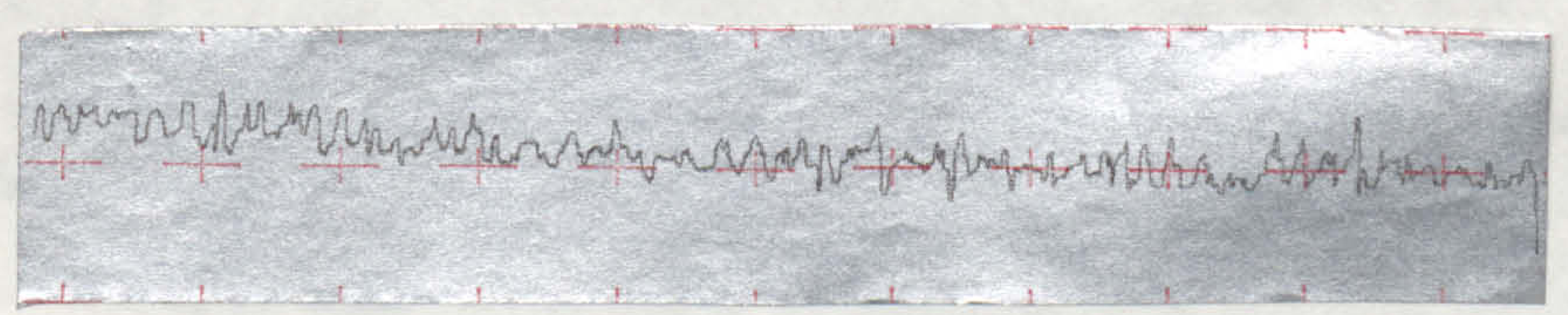
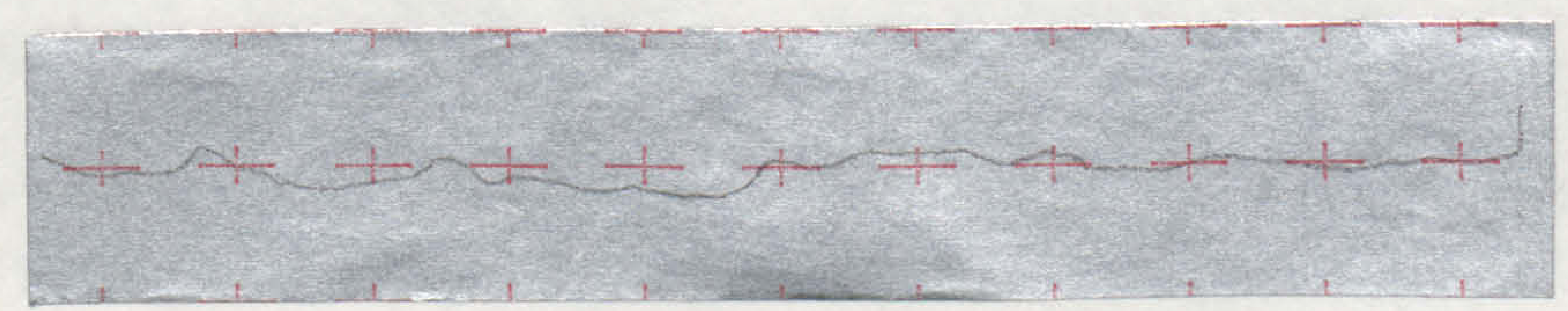


Photo 5.2. Ground surface.



Graph 5.2a. Surface profile of the ground surface taken across the lay. Magnifications: horizontal ×20, vertical ×2000.



Graph 5.2b. Surface profile of the ground surface taken with the lay. Magnifications: horizontal ×20, vertical ×2000.

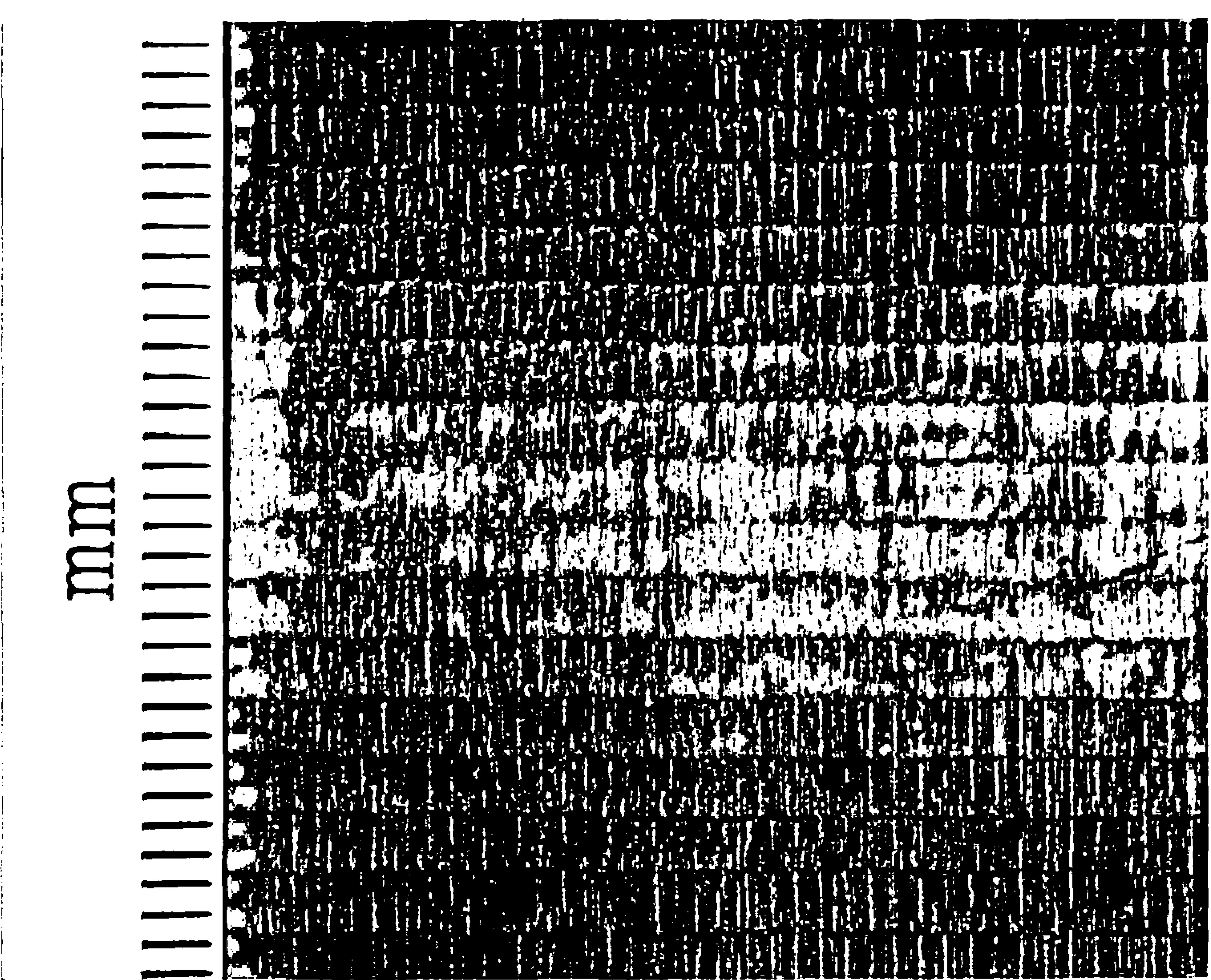
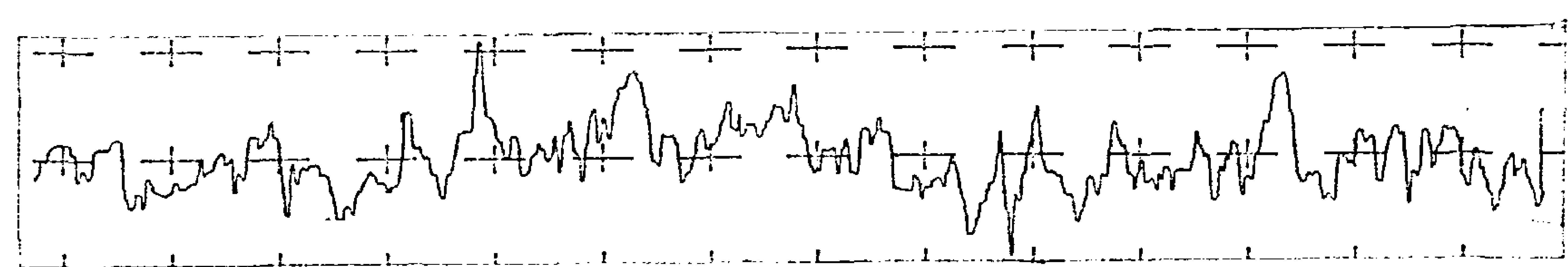
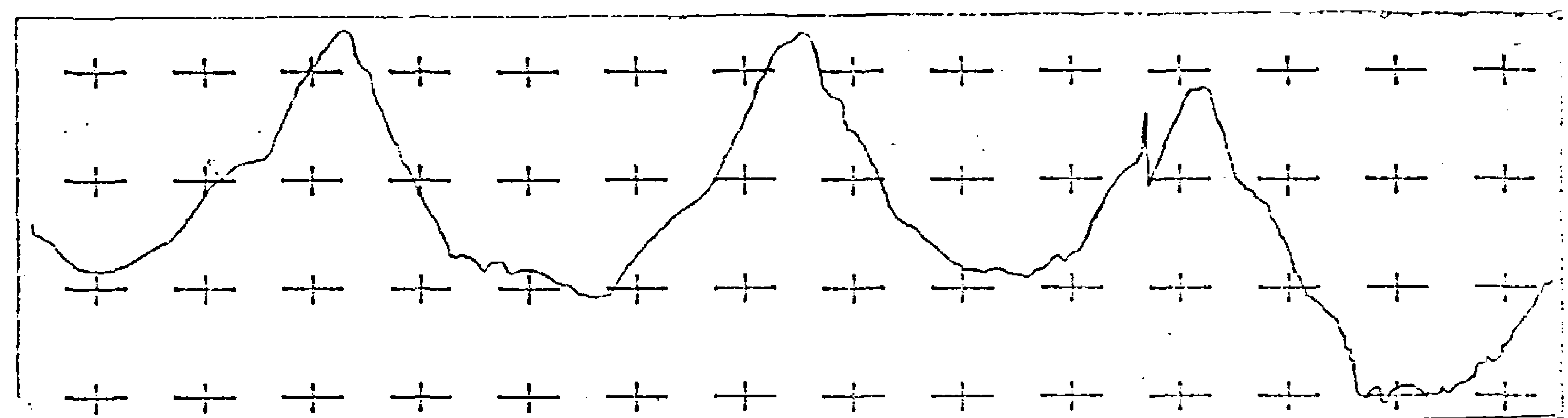


Photo 5.3. Milled surface.



Graph 5.2a. Surface profile of the milled surface taken across the lay. Magnifications: horizontal X20, vertical X2000.



Graph 5.2b. Surface profile of the milled surface taken with the lay. Magnifications: horizontal X20, vertical X2000.

5.3. Surface slopes and surface wavelengths expressed in optical wavelengths.

For defining the needs of light scattering techniques for the measurement of surface finish, Teague et al. [1981], have represented the surface topography in terms of two parameters: surface slope, H/D, and surface wavelength, D, H is the height of the surface. These two parameters were chosen as those most natural of major importance in theoretical models of light scattering and most convenient in describing the capabilities of measurement techniques. The two parameters also satisfy the widely recognised requirement that at least two descriptors are necessary to characterise the surface topography [Whitehouse and Archard 1970,Whitehouse and Philips 1978, Sayles and Thomas 1979,Thomas 1981, Peklinek 1961] : one descriptor for the height variations or the behaviour of a profile in a direction normal to the surface and one descriptor for the wavelength variations or the behaviour of a profile in a direction parallel to the surface. By normalising the surface wavelength to the optical wavelength of light used for the measurement (HeNe laser light at a wavelength of ca. 600nm), two dimensionless parameters are obtained for the construction of comparison charts or graphs. A third descriptor is needed for non-isotropic surfaces to account for the variations of profile properties with profiling direction. Light scattering techniques can be very useful for obtaining such a third descriptor [Young et al. 1980]

Figure 5.6. is a graph of the regions occupied by typical manufactured surfaces in the slope/relative-wavelength space. The

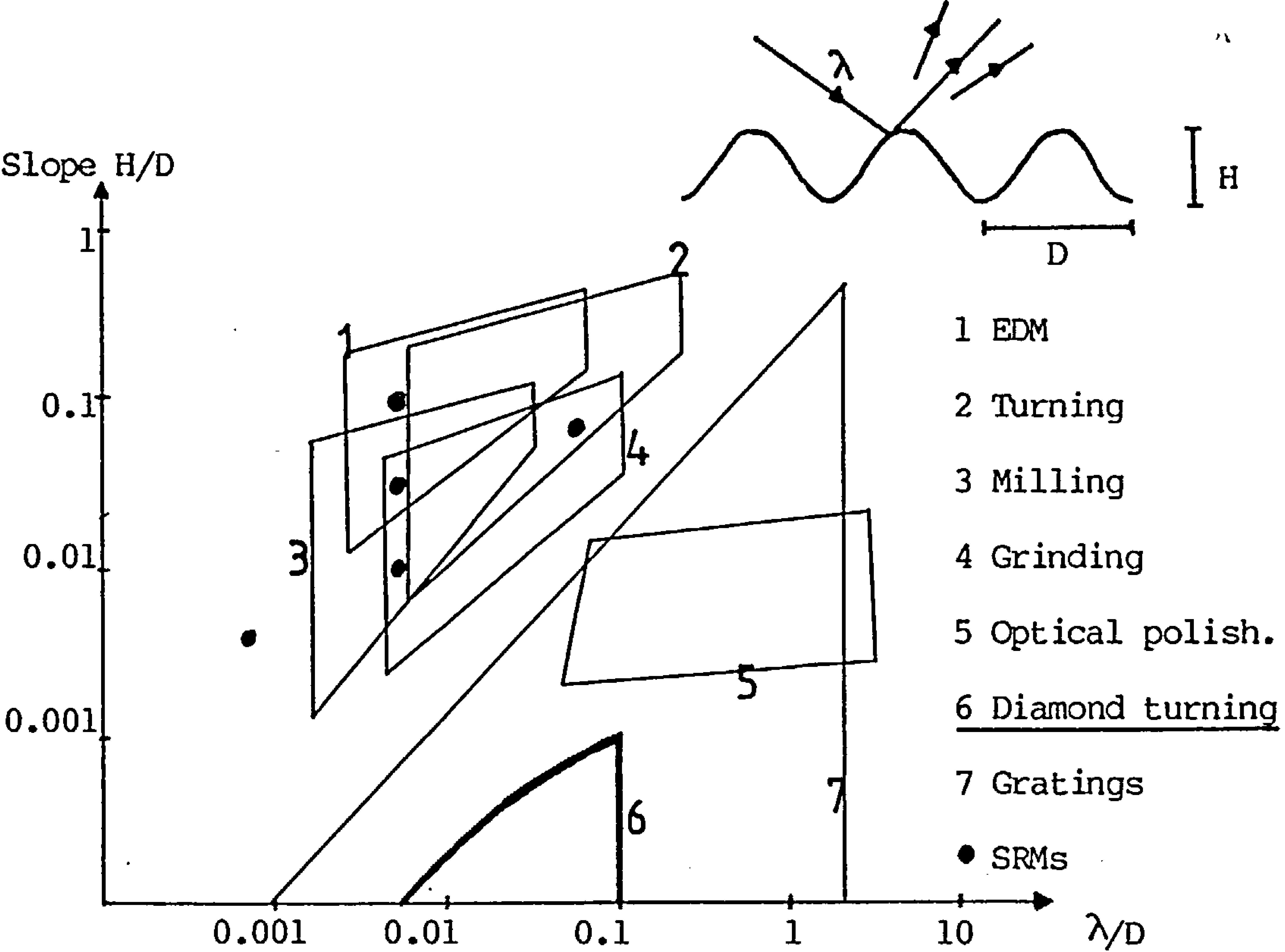


Figure 5.6. Surface topography in terms of optical wavelength.

indicated regions are determined by estimating the Fourier component amplitude and wavelength properties available from literature. Components of the Fourier decomposition of an arbitrary surface are assumed to be of the form:

$$y(x) = H/2 \sin(2\pi x/D) \tag{5.2}$$

No attempt is made to represent the relative position , or phase, of the different components or of their distribution. The estimates of slope range and relative wavelength range from electron discharge machined, turned, milled, and ground surfaces are based upon results given in papers by Peters et al [1979] and Spragg and Whitehouse

[1972], those for optical polishing and diamond turning upon results given by Bennett and Dancy [1981] and Church et al. [1977] and those for gratings on the experience of Maystre. The NBS standard reference materials (SRMs) are sinusoidal profile precision turned roughness specimens [Teague et al. 1982].

The process for turned surfaces is detailed, to illustrate how the regions are estimated from the usual surface parameters. Peters et al give R_t values for turned surfaces as ranging from 7 to 18 um and crest radii as ranging from 10 to 2 um, with some extensions to smaller radii. A range of H_{max} values for sinusoidal components making up a R_t value is assumed equal to the R_t range or $7<H_{max}<18$ um. This range is in reasonable agreement with the R_a values given by Spragg and Whitehouse and other available sources. A corresponding range of H_{min} values is then assumed to be one-tenth of the H_{max} range, i.e. $0.7<H_{min}<1.8$ um.

A wavelength range is estimated from the crest radii range using the radius of curvature equation

$$R_c = \frac{(1 + y'(x)^2)^{3/2}}{y''(x)} \qquad (5.3)$$

applied to the waveform in equation 5.2. Assuming large radii of curvature are associated with large wavelengths and largest amplitudes, a D_{max} and D_{min} were calculated. A range of $2.6<D<60$ um was obtained. Since in this instance Spragg and Whitehouse quote an average of 100 um, the range for D is extended to 100 um. In all other cases the bounds computed for D include their average wavelength as an intermediate value.

Normalising with respect to an optical wavelength of 0.6 um then results a λ/D range of 0.006 to 0.23. A slope range for the upper value of λ/D is computed as extending from (min H_{min})/D_{min} to (min H_{max})/D_{min}; the slope range at the minimum value of λ/D is computed as (min H_{min})/D_{max} to (max H_{max})/D_{max}.

5.4. The progress of accuracy in machining.

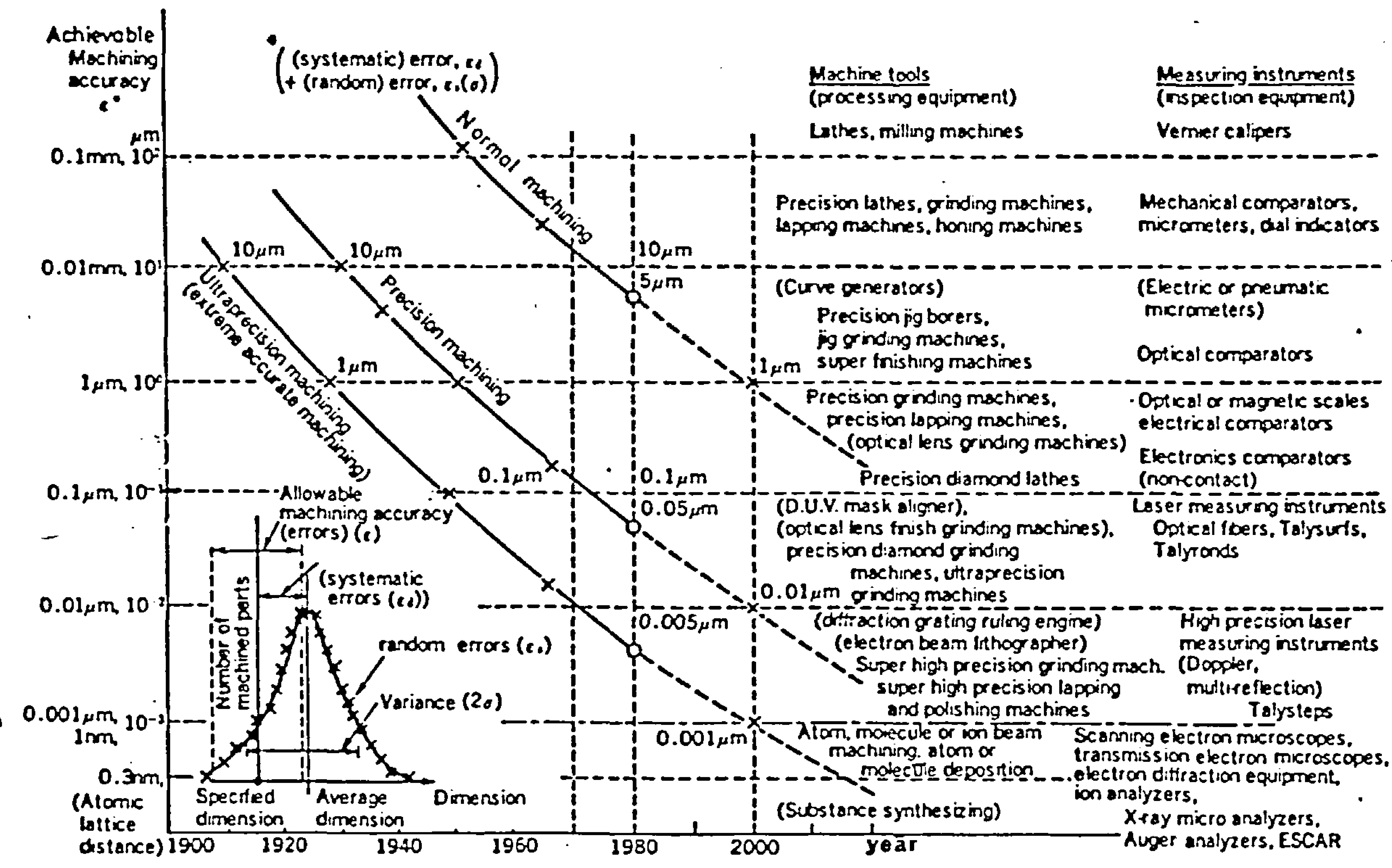


Figure 5.7. Modern trends in high precision machining.

The development of achievable machining accuracy over the last seventy years is shown in figure 5.7. [Taniguchi 1983] under the generalised classification of:

i normal machining.

ii precision machining.

iii ultraprecision machining.

It also lists the machines, processing equipment, and dimensional measuring equipment (and displacement transducers) by which the indicated resolution can be achieved.

By "ultra precision machining" are meant those processes and machines by which the highest possible dimensional accuracy is, or has been achieved at a given point in time. Today ultraprecision machining means the achievement of dimensional tolerances in the order of 10 nm and surface roughness of 1 nm. The dimensions of parts or elements of parts produced may be as small as 1 um and the resolution and repeatability of the machines used must be in the order of 10 nm.

These accuracy targets for present ultraprecision machining cannot be achieved by simple extension of conventional machining processes and techniques. In fact, a development and introduction of practice of a whole new range of materials processing technologies for the manufacture of parts to this order of accuracy has been taken place. These technologies together with the order of accuracy capability they offer are listed under "machining tools" in figure 5.7. Progressing from the microtechnology region (1 um accuracy capability) to the nanotechnology region, the systems engineering demand rapidly increases in stringency and complexity. Machine tools/processing equipment and dimensional and surface quality measuring instruments have to be integrated into closed loop

control systems which will provide high accuracy acceleration, velocity and position control between tools and workpieces, frequently at very high speed. The elements of machine tools, e.g. guideway bearings, displacement transducers/measuring equipment and servopositioning techniques broadly appropriate to levels of machine accuracy targets, are also shown in figure 5.7.

"Nanotechnology" is the term used to classify the integrated manufacturing technologies and machine systems which provide ultraprecision machining capability in the order of 1 nm. Nanotechnology, perhaps the most advanced manufacturing technology present, might also be called "extreme technology" because the theoretical limit of accuracy in machining of substances must be the size of the atomic lattice separation (0.2 to 0.4nm). The thin film technology required for the next generations of semiconductors demands the study of extreme technology problems and techniques; individual atoms have to be controlled and positioned where required. In this sense, ultraprecision machining technology is already approaching the extreme or ultimate limits.

6. Techniques for measuring surface finish.

Surface roughness measuring techniques can be broadly classified in two groups: (1) contacting types of instruments (tactile and stylus method), and (2) non-contacting types such as pneumatic, capacitive and optical instruments. These groups can be further separated into profiling and parametric techniques, whereby the latter can be devided in comparative and absolute measurements.

With profiling techniques the surface finish information is derived from a point by point scan of the surface height y as a function of the linear distance x along a straight line on the surface. The profile can subsequently be analysed mathematically to derive various surface finish parameters.

Parametric techniques are those in which some other property of the surface, such as friction, reflectivity or polar scatter, is measured. The required roughness parameter is then computed using an appropriate theory. Alternatively, the parametric measurement may be directly obtained in terms of the roughness parameter by the use of standard comparison surfaces with known properties, themselves calibrated by other means. These calibration specimens would have to be identical to the workpiece in every other way which might affect the result of the measurement.

Many of the non-optical techniques have been reviewed by Thomas [1982] and the optical ones by Vorburger and Teague [1981].

6.1. Contacting profilometry (stylus technique).

By far the most widely used method of measuring surface roughness is the stylus instrument. This has also influenced the standards. Most national and international standards are based upon stylus instruments, in which the path traced out by a sharp point drawn across the surface at constant speed is amplified and recorded. In the earliest stylus instruments the amplification was mechanical [Reason 1960, Way 1969, Abbott et al. 1938] or optical [Abbott & Firestone 1933], but in modern instruments it is electronic.

Originally the stylus was a phonograph needle [Abbott & Goldschmidt 1937] but sewing needles have also been used [Gray & Johnson 1972]. Today diamonds are universally employed: either 60° cones [Williamson 1947] or truncated 90° pyramids. The tip of the pyramid is about 3 um X 8 um in area, with the short edge parallel to the direction of travel [Jungles & Whitehouse 1970].

The finite dimensions of the stylus tip imply that the path it traces will not exactly follow the profile of the real surface: it may not enter steep valleys, and may record sharp peaks as cusps. These and other errors have been treated by a number of workers [Schlesinger 1942, Reason & al. 1944, Williamson 1947, Reason 1951, 1954, 1956, Peres 1953, Nakamura 1966, Radhakrishnan 1970, Whitehouse 1974, Tsukada & Anno 1974, 1975]. It is generally agreed that they are not important for most machined surfaces. A lateral resolution of the order of .1- .2 um and a height sensitivity of .1 nm rms can be achieved [Elson & Bennet 1979]. This instrument is a modified Talystep with a special 1 um radius diamond stylus. Other sources of

error such as the dynamic response of the stylus [Ajioka 1966, Damir 1973] its lateral deflection [Agullor & Pages-Fita 1974] have also been investigated and shown to be negligible.

Some researchers have suggested that plastic deformation under the stylus can be a source of serious error [Estill & Moody 1966, Schwartz & Brown 1966, Quiney et al. 1967, Guerro & Black 1972]. This has been refuted convincingly [Reason et al. 1944, Reason 1944 , Williamson 1968a,b]. Also some instruments , like the Talystep have the facility to adjust the stylus loading continuously from 0 to 10 mg and incremently in 10 mg steps up to 50 mg. However, for optical surfaces this damage may not be tolerable and a non-contact measuring technique is called for [Kranz 1980].

As with the cartridge of a high-fidelity phonograph, the performance of a stylus instrument is only as good as its transducer. The cheaper instruments use velocity-sensitive transducers such as piezoelectric or moving-coil [Reason 1951, 1959] or viscous-coupled moveable-anode vacuum tubes [Underwood & Bidwell 1953, Bidwell 1953]. Most high-priced stylus instruments use a linear variable differential transformer [Reason 1951, 1960, Chinick 1968] which will measure low frequencies right down to a static stylus displacement.

It is usual to measure stylus position with respect to a skid travelling over the surface beside or behind it, thus avoiding setting-up errors. The skid itself may cause surface damage [Tucker & Meyerhoff 1969] and will in any case act as a high-pass filter, though the effect of this can be allowed for [Nara 1966; Whitehouse 1982]. With the skidless type the stylus moves relative to a precision surface built into the instrument; the resultant

information is a total non-distorted profile.

In some early instruments the amplified output drove a loudspeaker [Harrison 1931] or an oscilloscope [Abbott & Firestone 1933].

Now it is usual to record the signal on a chart recorder and also on an averaging meter of some kind; this may measure centre-line average or root-mean-square roughness or even the bearing area Oonishi 1966] or the average wavelength [Spragg & Whitehouse 1970/71, 1974]. Alternatively, the signal may be tape-recorded for subsequent replaying through signal analysing equipment [Tallian et al. 1964, Silin & Frederick 1967, Ber & Braun 1968, Hammond 1970 Osman & Sankar 1975]. This analysing equipment, which may also be used on-line, can vary in complexity from fairly simple analogue circuitry [Pesante 1955,1964, Wolff 1967/68, Nara 1969, 1971, Hammond 1970, Nayak 1971, Willn 1972, Selvam 1975], up to full-scale analogue computers [Tallian et al. 1964, Henry & Fenech 1964, Tallian & McCool 1966].

A logical development from this was was the use of the digital computer, a far more powerful and flexible device because of its ability to accept complex programs for analysis. This requires the analogue signal from the stylus instrument to be converted to digital form, an operation which must be carried out carefully if information is not to be lost [Kaliszer et al. 1971, Whitehouse 1974].

Fine resolution and wide range are important features of stylus instruments. The ultimate vertical resolution can be as small as .1 nm rms and the range can be as much as 100 um. The horizontal resolution is limited by the stylus tip width; it is typically several um but can be as small as .1 um, more than acceptable for

most topographic measurements. A newly introduced stylus instrument employs a laser interferometric transducer with the reflector attached to a pivoted arm which carries a stylus to contact the workpiece.

However, the stylus instrument has a number of shortcomings. The transducer and the stylus are fragile, so the instrument must be used in a fairly quiet, clean environment. Without proper use, the sharp stylus can damage the surface and, depending on set-up time, a surface profile measurement can take from one to ten minutes, which is too slow for on line or in process applications. Finally, conventional use of stylus instruments only generates one-dimensional surface profiles and therefore, the fraction of the surface area examined is quite small. Thus its usefulness is limited for such applications as nondestructive, post process evaluation.

6.2. Contacting parametric techniques.

6.2.1. Tactile testing.

The basic idea of a tactile test is that a probe of some kind is run across the surface to be measured and the friction between the surface and the probe is compared with that from a similarly machined surface of known roughness. Sets of machined specimens suitable for tactile comparison tests are commercially available [Rubert 1961]. The simplest and cheapest probe is the human fingernail, and it is surprisingly effective.

The human fingernail is more sensitive to some frequencies than others [Abbott & Goldschmidt], so there is presumably an optimum speed with which it should be drawn along the surface. Schlesinger [1942] performed some careful tests in which subjects were asked to differentiate between pairs of testpieces of increasingly different roughness. He found that for some finishes differences in roughness as little as 20% could be detected by the majority of his test panel (figure 6.1.). A similar test by Haesing [1961] showed that there was

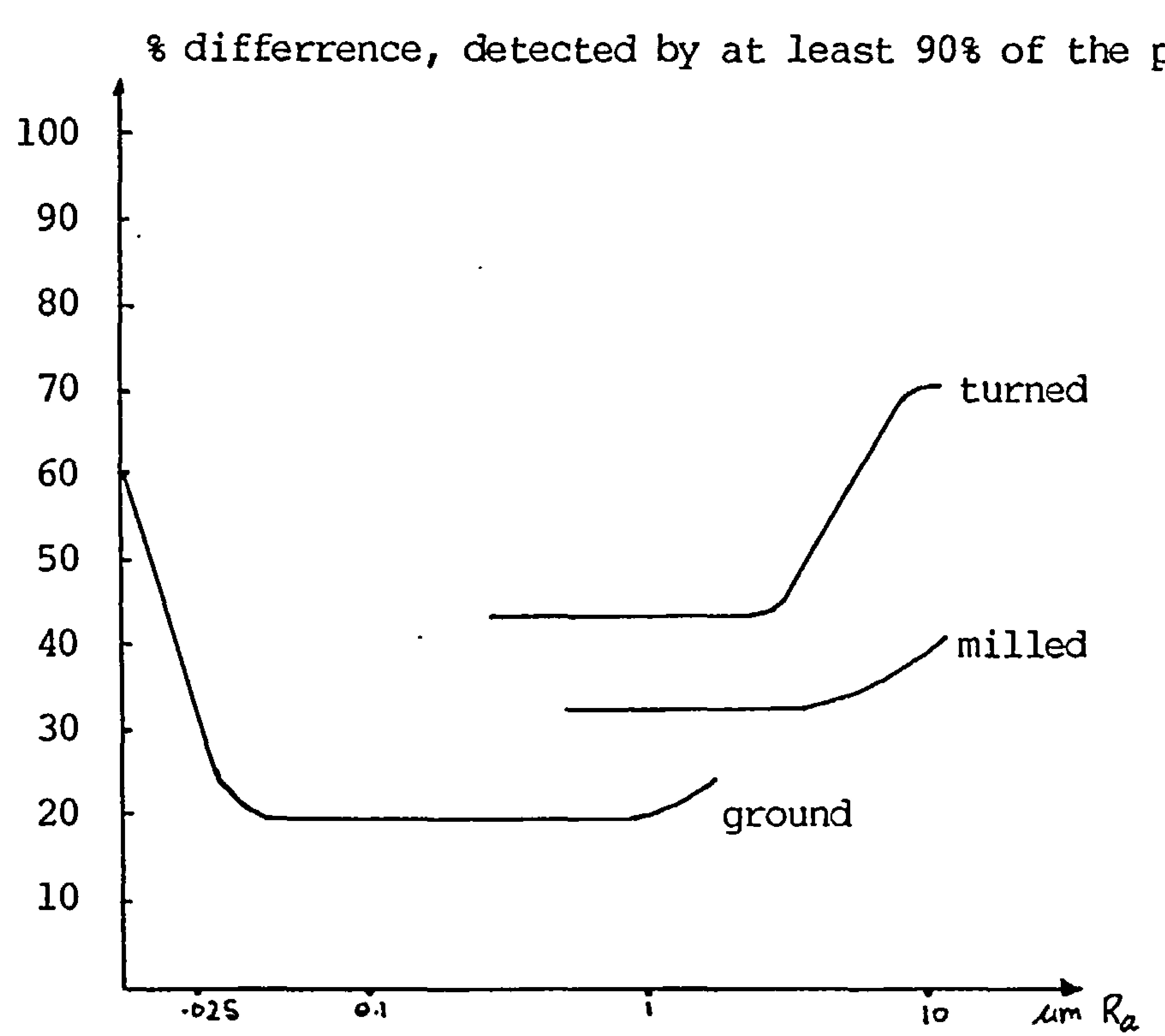


Figure 6.1. Effectiveness of tactile comparison tests (Schlesinger).

a higher correlation between the subject's assessments of R_p than there was with R_a. In a variation on this, however, where the tester draws the edge of a coin across the surface and assesses noise as well as friction, no better results than those from unaided visual

inspection were achieved [Schlesinger 1942].

This technique is quicker, cheaper and simpler than any other method providing a suitable range of reference specimens is available, but to produce reliable results a fair amount of experience is probably necessary on the part of the tester.

A commercially available development is the Mecrin tester, where a thin flexible steel blade is pushed along the surface at a gradually increasing angle till it buckles [Rubert 1967/68]. As the device presumably relies on just failing to overcome static friction it must be sensitive to the mean slope.

6.2.2. Thermal comparator.

In the thermal comparator two metal spheres are mounted in a block of thermally insulating material [Powell 1957]. One is completely recessed while the other protrudes slightly. They are connected by a differential thermocouple. The apparatus is placed in an oven till it attains constant temperature and then removed and placed, under a load of about its own weight, on the surface whose roughness is to be measured. The protruding sphere will cool faster because it is in contact with the test surface, and this is quantified as a voltage reading from the differential thermocouple after a certain time has elapsed from contact.

Clearly this technique relies on far too many parameters, most of them difficult to quantify, to be suitable for an absolute determination of roughness. The rate of cooling of the sphere not in

contact will depend on its initial temperature and on various material and atmospheric properties. The increased rate of cooling of the sphere in contact will depend on the thermal conductance of the contact, which in turn will depend on the thermal conductivities of the contacting materials and their elastic moduli or relative hardness depending on whether the contact is elastic or plastic, in addition to the surface properties and the load [Thomas 1979].

The sensitivity increases as the test surfaces become smoother. The useful upper roughness limit is probably about 4 um rms. The lower limit would probably be set by the roughness of the spheres themselves. The range of wavelengths measured depends on whether contact is plastic or elastic, but the long-wavelength cut-off for this partical instrument is probably about 10 um. An instrument based on this principle might be useful on a production line as a go not-go gauge possibly with a built-in heater; it should be relatively cheap and robust and very simple to operate. Thermal comparators are now commercially available.

6.2.3. Friction dynamometer.

Another friction method is based on the retardation of the swing of a pendulum due to friction between the tested surface and a smooth shoe attached to the pendulum. The pendulum is released from a position 30° from the vertical and the apparatus simultaneously begins to eject a paper tape at constant speed. The length of tape ejected before the pendulum comes to rest is taken as a measure of the finish

of the test piece [Jost 1944]. Dynamic friction is influenced by at least two surface parameters, rms roughness and mean slope [Hirst & Hollander 1974], so the instrument must be used with reservations even as a comparator. The long wavelength cut-off will depend on the nominal contact area, which in turn will depend on the load and the geometry of the contacting surfaces. The vertical resolution will depend ultimately on the finish of the shoe. The apparatus is easy and quick to operate but rather expensive to build; its measurements will also be affected by the cleanliness of the test surface.

6.2.4. Rolling ball.

In a method employing spherical contact a ball of radius r rolls down an inclined plane as soon the tilt angle Θ exceeds a value which increases with the roughness of the plane [Bikerman 1970]. If R_p is the peak to valley roughness,

$$R_p=r(1-\cos(\Theta)) \qquad (6.1)$$

This relation is stated to hold approximately for surfaces rougher than 1 um.

Although this method makes use of the phenomenon of static friction the equation shows that it is independent of material parameters. Assuming contact is elastic, the area measured is probably not larger than 10 um in diameter. The method gives an absolute measurement. It is very cheap, robust and simple to use, and

might be suitable for production-line gauging, though it might be easier to set up in the form of a rolling cylinder.

6.2.5. Flowing drop.

Using a somewhat similar principle, an oil drop is timed as it flows down a fixed length of an inclined test piece [Kamnev 1966]. Both the mean slope and the rms roughness might be expected to affect the performance. This is confirmed by a difference in behaviour between surfaces finished in various ways. Roughness was found to be proportional to the 3.7 power of the time for filed and milled surfaces but to the power 1.74 for etched surfaces. Sensitivity falls off rapidly by the diameter of the drop. The method is simple, cheap and surprisingly reproducible, but rather slow.

6.2.6. Thetameter.

A smooth steel sphere is pressed into the test surface under a known load. The increase in load required to increase its penetration by a fixed amount is measured [Tornebohm 1936]. The "theta" of its title is the effective change in the Hertzian elastic modulus of the test piece due to its roughness. A rigorous theory of the elastic contact of the sphere with a rough plane was developed by Greenwood and Tripp [1967]. The effect of roughness prevails only at light loads, and then not in a simple relationship with load. Asperity

density and curvature are also involved, though for certain surfaces these could probably be replaced by the rms roughness. As the instrument reading is the result of a combination of at least two surface parameters it is probably not suitable even for a comparison without exhaustive calibration against a less ambiguous instrument using the range of surfaces to be tested.

On the other hand the thetameter is reasonably cheap, robust, quick and easy to use and reproducible. Its sensitivity will decrease with decreasing roughness and the limit of vertical resolution will be set by the load measuring device. The long wavelength cut-off will depend on the Hertzian contact area. No dimensions are given in the reference, but an estimate of 10 um diameter seems plausible.

6.3. Non-contacting non-optical techniques.

6.3.1. Electrical methods.

Capacitive.

The capacitance between two conducting elements is directly proportional to their area and the dielectric constant of the medium between them and inversely proportional to their separation. If a rough surface is regarded as the sum of a number of small elemental areas at different heights it is fairly easy to work out the effective capacitance between it and a smooth surface for various deterministic surface models [Sherwood & Crookall 1967/68;ten Napel &

Bosma 1970/71].

Unfortunately, real surfaces are rarely deterministic. The capacitance of a capacitor, one of whose plates has a roughness expressed by f(x,y), is proportional to

$$C = k \iint_A \frac{dx\,dy}{d + f(x,y)} \tag{6.2}$$

where d is the separation of the two mean planes.

The difficulty is at once apparent: the integrand tends to infinity as soon as the separation of the two planes becomes comparable with the maximum peak height of the surface, an infinite capacitance will be the result. Some numerical solutions for a Gaussian height distribution truncated at various arbitrary heights indicate that the capacitance is very sensitive to the mean plane separation and to the height of the highest point on the surface. An instrument for measuring surface roughness based on a capacitance principle has become available [Fromson et al. 1976; Brecker et al. 1977]. Fig. 6.2. suggests that its reading may vary considerably with the finishing process. On the other hand, it may prove to be a sensitive detector of running-in.

Inductive.

The reluctance between two magnetic surfaces will also be a function of their roughness, again because reluctance falls off with increasing separation. Radhakrishnan [1977] has measured the

inductances between a magnetic recording head and a number of rough surfaces, and compared his results with stylus measurements. Useful correlation with average roughness was obtained only when the comparison was restricted to a particular process. A stronger correlation was reported with peak density, again indicating a sensitivity to local maxima. As the measurement is quick and cheap this technique also has possibilities for quality control, though is restricted to magnetic materials.

Skin effect.

Alternating currents of high frequencies are shifted from the central to the peripheral annuli of a wire; the major part of the current flows in a surface layer, which, for copper, would be about 0.4um thick at 25 GHz [Bikerman 1970]. Thus, at high frequencies, the thickness of the actively conducting region is of the order of magnitude of the height of the surface hills. Consequently, the experimental resistivity of a wire deviates from that calculated under the assumption of no rugosity, and the degree of roughness can be deduced from this deviation.

The largest height difference which can be detected decreases as the frequency increases. The smallest detectable height difference will depend on the resolution of the resistance change. The long wavelength cut-off will be set by the frequency of the a.c. current and the velocity of propagation in the metal; for the above frequency it would be about 5 mm for a steel wire. The method is expensive and

suitable only for measurement of wire specimens, but for these it might well be the only practicable technique.

6.3.2. Fluid methods.

Pneumatic gauging.

The essential elements are a gauging nozzle in proximity to the test surface connected through an intermediate chamber to a source of air at constant pressure P. Air is admitted to the intermediate chamber through a control orifice of cross-sectional area a, and escapes between the gauging nozzle and the rough surface through an effective cross-sectional area A, thus lowering the pressure in the intermediate chamber to p. It can be shown that over a certain range the relation between p/P and A/a is linear.

A circular nozzle was originally used, but various other shapes have been tried [Radhakrisnan & Sagar 1970], including slits and cruciform sections. A sufficiently fine slit should provide something nearly like a profile measurement, with the high-pass cut-off set by its length. The horizontal and vertical resolutions presumably depend in principle on gas laws, but in practice the useful horizontal range of measurement is set by the range of linearity of the relationship given above. Measurements have been reported [Graneek & Wunsch 1952; Wager 1967] which correlate very well with Ra readings from a stylus instrument over a range from 0.1 to 5 um.

This pneumatic technique seems to be a practical method. With a

circular nozzle it is unduly sensitive to long wavelengths, but offers the compensating advantage of giving a reading independent of orientation of a surface with lay. Its reading depends on height parameters only. Up till now it has only been used as a comparator, but a more rigorous theory could be devised to give an absolute interpretation [Tanner 1981]. It is cheap, robust, simple quick and non-contacting , well suited to a production line.

Oil droplet.

A droplet of oil of volume v is placed on to a rough solid and squeezed with an optical flat [Bikerman 1970]. If the greatest area of the oil patch which can be achieved is A, then v/A is the average thickness of the patch. This is claimed to be approximately equal to half R_p, though this relationship almost certainly depends both on the finish of the surface and on the size of the droplet. The vertical resolution will depend on the waviness of the optical flat. The method is simple and cheap, but slow and not very reliable.

Stagnant layer.

A plate is covered with a non-volatile liquid and suspended vertically [Bikerman 1970]. The mass and thus the volume v of the liquid remaining on the plate after time t is determined from time to time. The experimental function v=f(t) differs from that predicted by

hydrodynamics in such a manner as if a stagnant layer were present. The deduced thickness of the layer is from one to two times the RMS roughness. The height resolution will depend on the resolution of the weighing equipment; a difference in roughness of 0.1 um on an area of 100 cm would cause a change of only about 1 mg in a total mass of perhaps 100 g. The method is not very simple and requires quite expensive equipment; it is slow and not very reliable; and it is only suitable for workpieces of the appropriate geometry.

6.4. Comparison of non-optical measurement techniques.

Table 6.1. summarizes the capabilities of various measurement techniques. From this table it is obvious that the stylus instrument is superior to the other instruments, however it does not score very well for relative ease of use and robustness.

Table 6.1. Comparison of non-optical measurement techniques.

Technique	Horizontal		Vertical		Ease of use
	resolution	range	resolution	range	and robustness
Stylus	0.1 um	100 mm	0.1 nm	10 mm	3-5
Tactile	1 um	1 mm	0.1 um	3 um	0
Thermal	0.1 um	1 um	1 nm	3 um	2
Capacitive	0.1 um	10 mm	30 nm	100 um	2
Pneumatic	1 um	10 mm	0.1 um	3 um	1

6.5. Non-optical in-process measurement.

Few of the techniques discussed in the foregoing will be suitable for in-process measurements. Pneumatic gauging looks promising, and it has been established [Wager 1967] that the dynamic effects of the moving workpiece are not serious. This is rather surprising, as the theory of the pnuematic gauge [Graneek & Wunsch 1952] assumes isothermal conditions whereas fluctuations at turning speeds are more likely to be adiabatic. The pnuematic gauge is robust and the air jet will help to clear unwanted surface fluid. It also measures a roughness integrated over the entire path of movement of the surface. Its disadvantages are that a nozzle fixed relative to the workpiece will be unduly affected by waviness, and that it is too insensitive to measure fine finishes.

The stylus instrument, which might be thought a priori unsuitable, has proved itself a serious contender in the measurement of turned and ground surfaces [Dutschke & Eissler 1978; Webster & Kaliszer 1980]. The measuring device is a steel cylinder which rotates in contact with the workpiece being machined. At every revolution of the cylinder a stylus, piercing a hole in its circumference, is deflected, producing an electrical signal which is read through telemetry. Surprisingly good correlations with orthodox roughness measurements are reported, which is surprising since this system is not really measuring across the lay as with a conventional stylus instruments but is measuring in a helical fashion, which can introduce some aliasing effects.

6.6. Optical profilometry.

Optical profilometry can be divided into two categories, (1) interferometric and (2) optical stylus techniques.

6.6.1 Interferometry.

Interferometry is an important technique for measuring the roughness and form of high quality, optical surfaces [Tolansky 1960, Koehler & White 1955, Baird & Hanes 1967, Bennett 1976, Dietz & Bennett 1967, Moody 1968].

The essentials of interferometry is that an optical wavefront is split into two coherent beams. One beam is reflected from a smooth flat reference surface R while the other is reflected from the surface under test S. The reflected beams are then recombined. With a small amount of tilt between either the reference or test surface, the interferometer will produce a set of generally parallel fringes. The undulations of each fringe reveal the peaks and valleys of surface in a manner similar to that of a roughness profile obtained from a stylus instrument. Interferometry enables one to measure surface heights directly in terms of optical wavelengths. Its ultimate resolution using multiple beams can be as small as .05 nm peak to valley [Lang & Scott 1968] but is generally about one nm rms [Bennett 1976].

There are, however, undesirable features which limit the applicability of interferometry for on-line measurement of a wide

range of surfaces. Firstly, since the interferometer normally produces an optical fringe, added instrumentation and data processing are required to convert the fringe into profile data. Secondly, the vertical range is limited. If the surface roughness is greater than $\lambda/2$, the interference fringes tend to merge into one another making it impossible with conventional interferometry to track one fringe at a time.

Michelson interferometry.

One important type of interferometer is the Michelson interferometer as depicted in figure 6.2. Here A is an extended

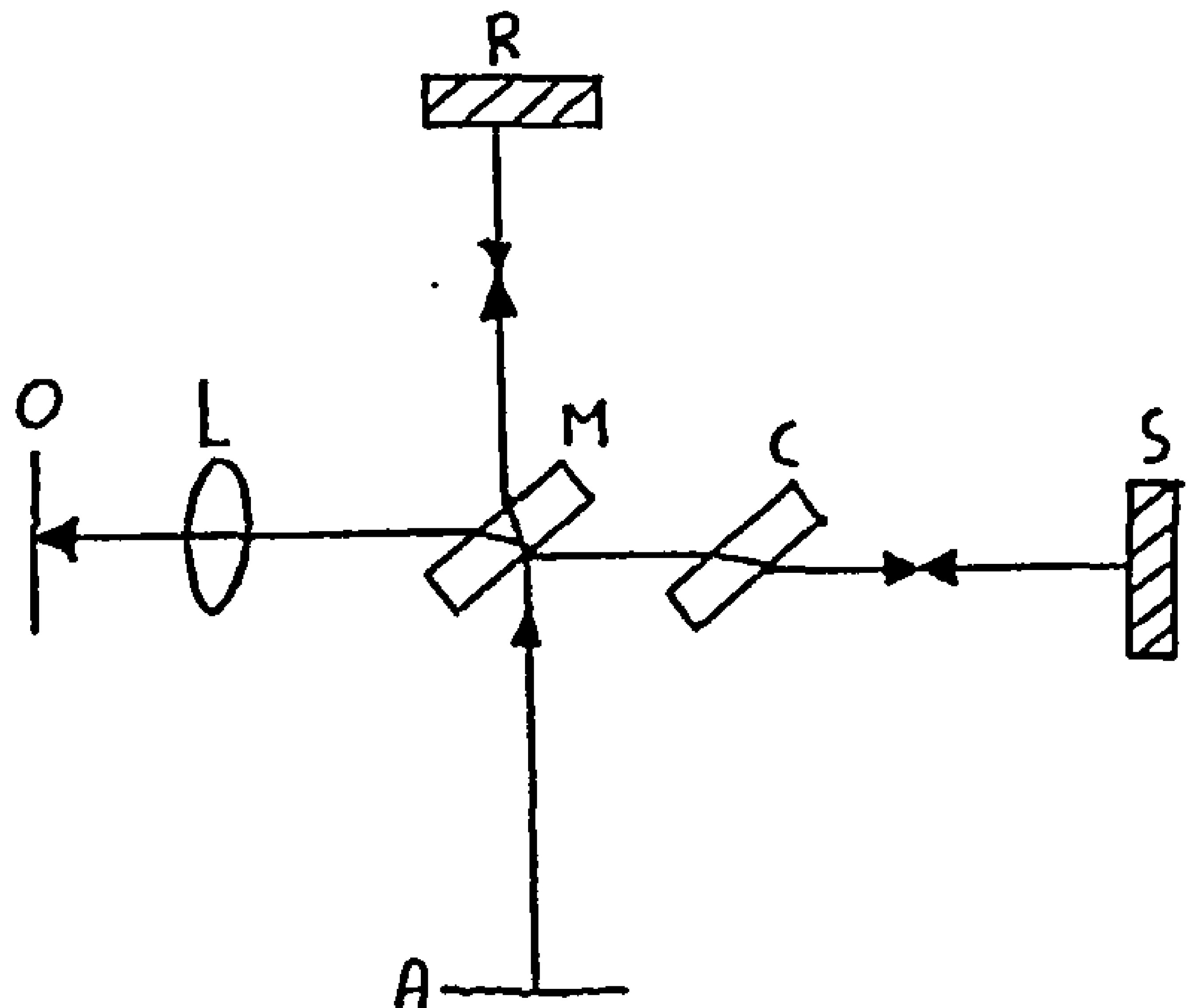


Figure 6.2. Michelson two-beam interferometer.

source of monochromatic light. The fringes are observed at O. C is a compensator positioned to keep the optical path lengths of the two beams equal.

To overcome the general vertical height limitation of $\lambda/2$, Pettigrew and Hancock [1979] have designed an ac-interferometer based upon the Michelson interferometer, but employing two illumination wavelengths simultaneously. Their system creates an effective wavelength $\lambda_{eff}=\lambda_1\lambda_2/(\lambda_2-\lambda_1)$. The vertical limit in this system is $\lambda_{eff}/4$. An effective wavelength of about 50 um has been generated.

Mirau interferometry.

Bhushan et al. [1985] have developed an instrument based upon a Mirau interferometer. A schematic of the interferometer is illustrated in fig. 6.3. The interferometer is an attachment to a

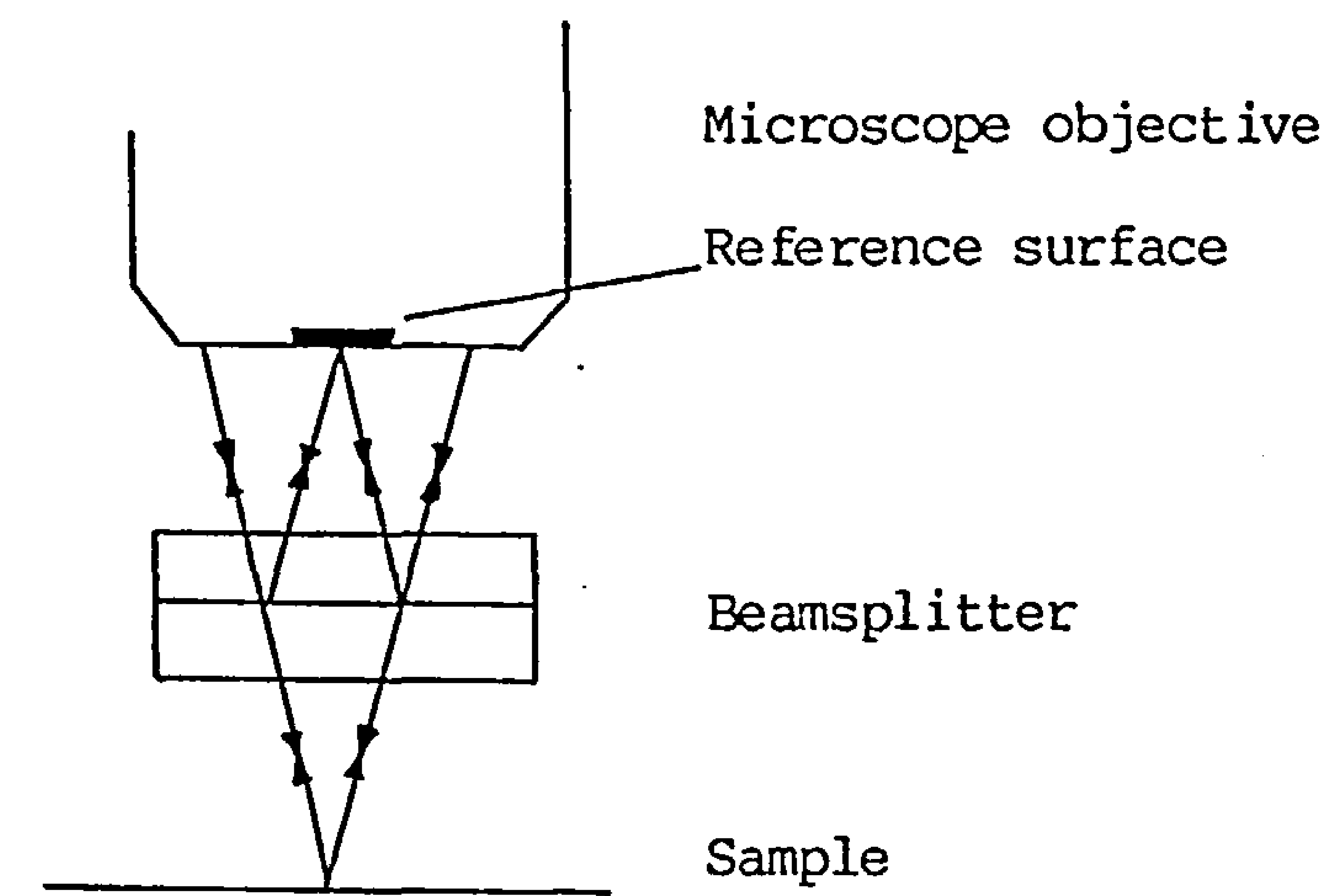


Figure 6.3. Mirau interferometer.

long working distance microscope objective. Part of the light goes to the sample under test and the rest of the light is reflected by the beamsplitter to the reference surface. The light reflected by the sample and the light reflected by the reference surface are combined

again at the beam splitter. These two light beams interfere. The resulting interference fringes give the difference between the sample surface and the reference surface.

A major advantage of the Mirau interferometer compared with other two beam interferometers is that the interferometer is a common path up to the beamsplitter plate. Beacause the two beams of the interferometer are created after the microscope objective, the same abberations caused by the microscope objective and all the optics before the beamsplitter are present in both arms and yield no net optical path difference when the two beams are combined after interference. Small perturbations in the optical path caused by the beamsplitter plate are averaged, because the interference in the Mirau interferometer is localised at the test and reference surfaces.

The interferometer does have the disadvantage that it requires a long working distance microscope objective to accommodate the two plates, which consist of the reference surface and the beamsplitter plate. This limits the numerical aperture to 0.6. The reference surface modifies the optical transfer function of the objective because the reference surface is a central obscuration in the beam. This has the effect of reducing the contrast for larger field of view optical systems.

The surface information is obtained by changing the optical phase of the reference beam in 3 steps of $\pi/2$. The optical phase $\phi(x)$ the light reflected by the surface can then be determined. This phase information is then converted into a corresponding height distribution by the equation:

$$y(x) = (\lambda/4\pi)\,\phi(x) \tag{6.3}$$

The system uses a solid state linear array of 1024 photodetectors, whose signals are fed into a computer for evaluation. The optical phase shifting is accomplished by moving the reference plate by a piezoelectric transducer which translates the reference plate.

A lateral resolution of 0.65um can be achieved and the repeatability is 0.5 nm rms. and an accuracy of 1 nm.

MacBean [1984] has proposed a modified Mirau interferometer, which operates under oblique incident illumination, see figure 6.4.

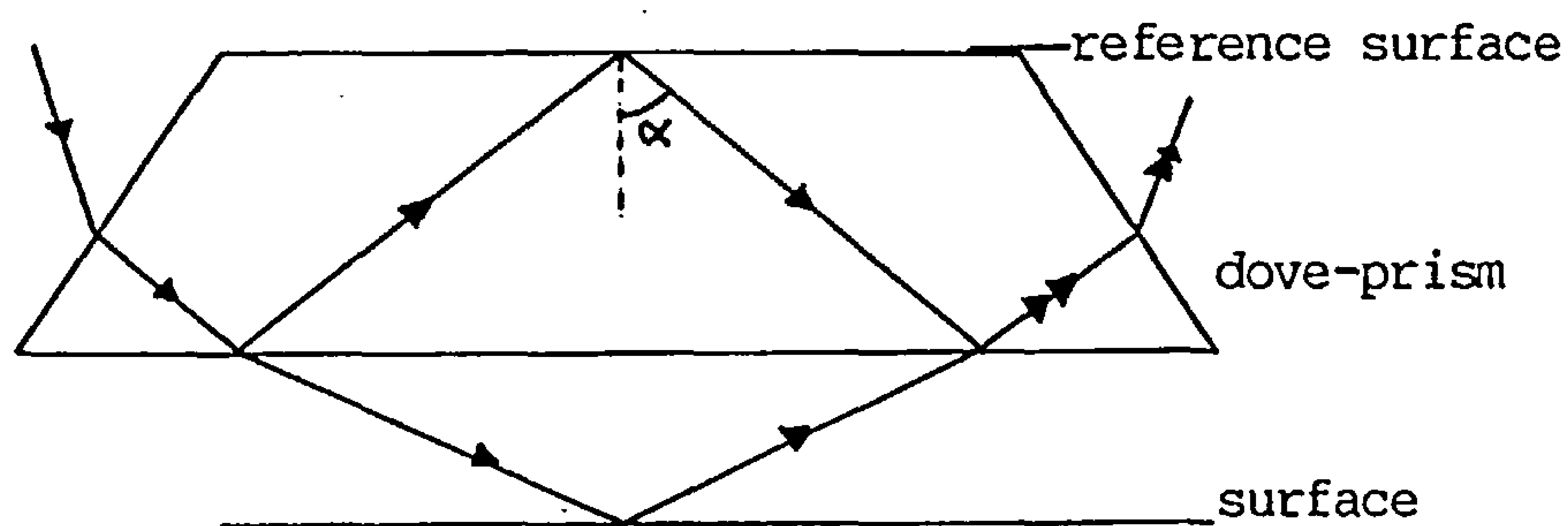


figure 6.4. Optical paths in the dove prism interferometer.

In this version the beamsplitter-plate and reference surface have been replaced by a dove-prism. This prism acts as an beamsplitter and one of its faces provides the reference surface. This changes the effective wavelength to $\lambda/\cos\alpha$. This increases the fringe spacing compared with normal incident light, thereby increasing the vertical range considerably.

Fizeau interferometry.

Two other interferometric techniques that are suitable for determining the roughness on optical surfaces are FECO and Fizeau interferometry Bennett & Bennett 1967, Tolansky 1948]. In Fizeau interferometry, a whole class of interference fringes exists, for which the optical thickness, $n_f d$, is the dominant parameter rather than the angle of incidence. These are referred to as fringes of equal thickness. Under white light illumination the iridescence of oil slicks (a few wavelengths thick) and even oxidized metal surfaces, are all the result of variations in film thickness. Interference bands of this kind are analoguous to the constant height contour lines of a topographical map. Each fringe is the locus of all points in the film for which the optical thickness is constant. In general, n_f does not vary, so that the fringes actually correspondent to regions of constant film thickness. As such, they can be quite useful in determining the surface features of optical elements. For example, a surface to be examined may be put into contact with an optical flat. These flats can be made better than $\lambda/200$. The air in the space between the two generates a thin film interference pattern. If the test surface is flat, a series of straight, equally spaced bands indicates a wedge shaped air film, usually resulting from dust between the flats.

When viewed at nearly normal incidence in the manner illustrated in fig 6.5., the contours arising from a nonuniform film are called Fizeau fringes. For a thin wedge of small angle the optical path length difference between two reflected rays may be approximated by

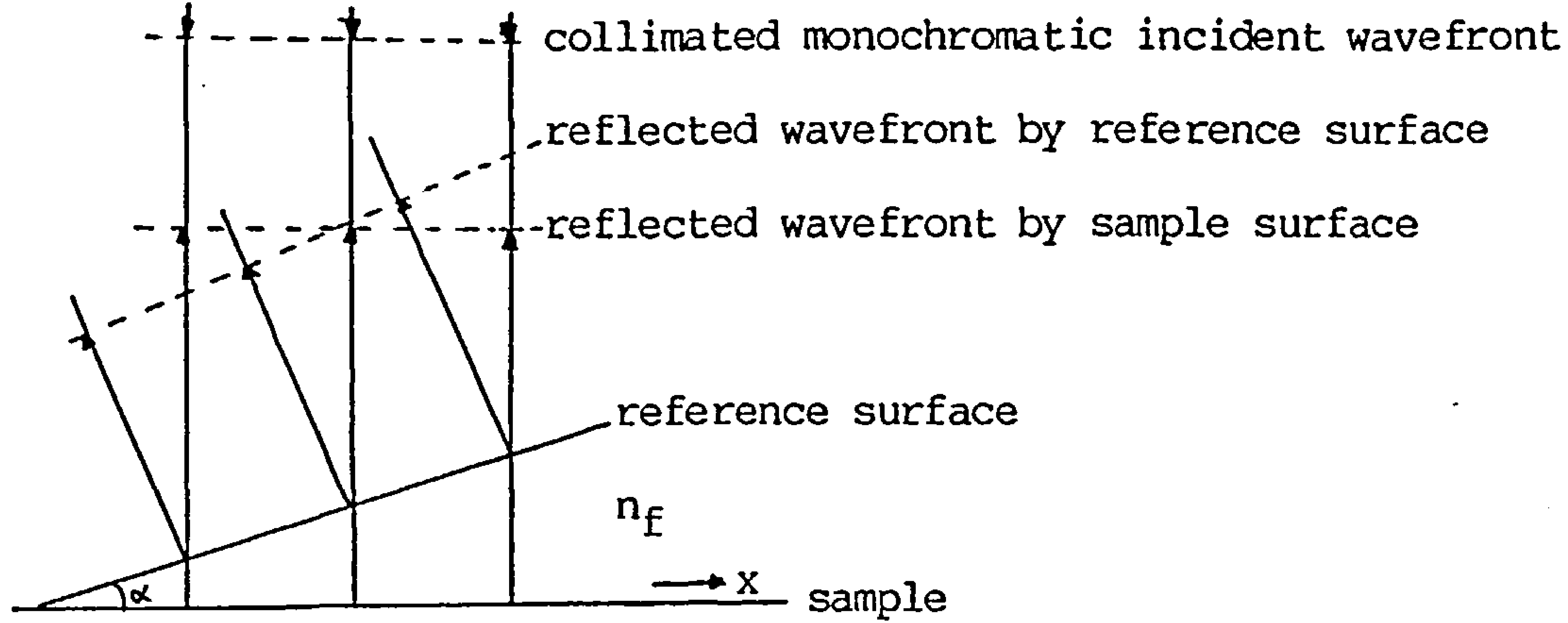


Figure 6.5. Fizeau interferometer

$L=2n_f x\alpha$. For small angles of incidence the condition for an interference maximum becomes

$$(m+1/2)\lambda = 2n_f x\alpha \tag{6.4}$$

The surface roughness will cause the fringes to deviate from straight lines. These deviations can be transformed into surface profiles, which subsequently can be processed for the extraction of various surface parameters.

Along these lines Eastman [1980] has developed a scanning Fizeau interferometer, working in transmission. It is capable of producing surface profile measurements of a 1 mm strip in less than three minutes. The lateral resolution is approximately 4 um. The horizontal resolution is .5 nm. Work is under progress to improve the lateral resolution to about the diffraction limit. Mention is made that the instrument is very sensitive to vibrations, but this is a general problem with interferometers.

FECO interferometry.

A FECO (Fringes of Equal Chromatic Order) scanning interferometer has been developed by Bennet [1976]. Essentially, the instrument is a combination of a Fabry-Perot interferometer and a spectrograph to analyse the light reflected by the interferometer.

Initial techniques, procedures and results were reported in 1953 by W.F. Koehler [1953]. The FECO fringes are formed when a collimated beam of white light undergoes multiple reflections between two partially silvered parallel surfaces, one of which is the surface whose profile is being measured and the other is a supersmooth reference surface. When the light is dispersed by a prism or grating, the FECO fringes occur as wiggly black lines in a bright continuum. The wiggles are the contours in the wavelength regime of microirregularities on the pair of surfaces comprising the interferometer.

A schematic of a Fabry-Perot interferometer is shown in figure 6.6.

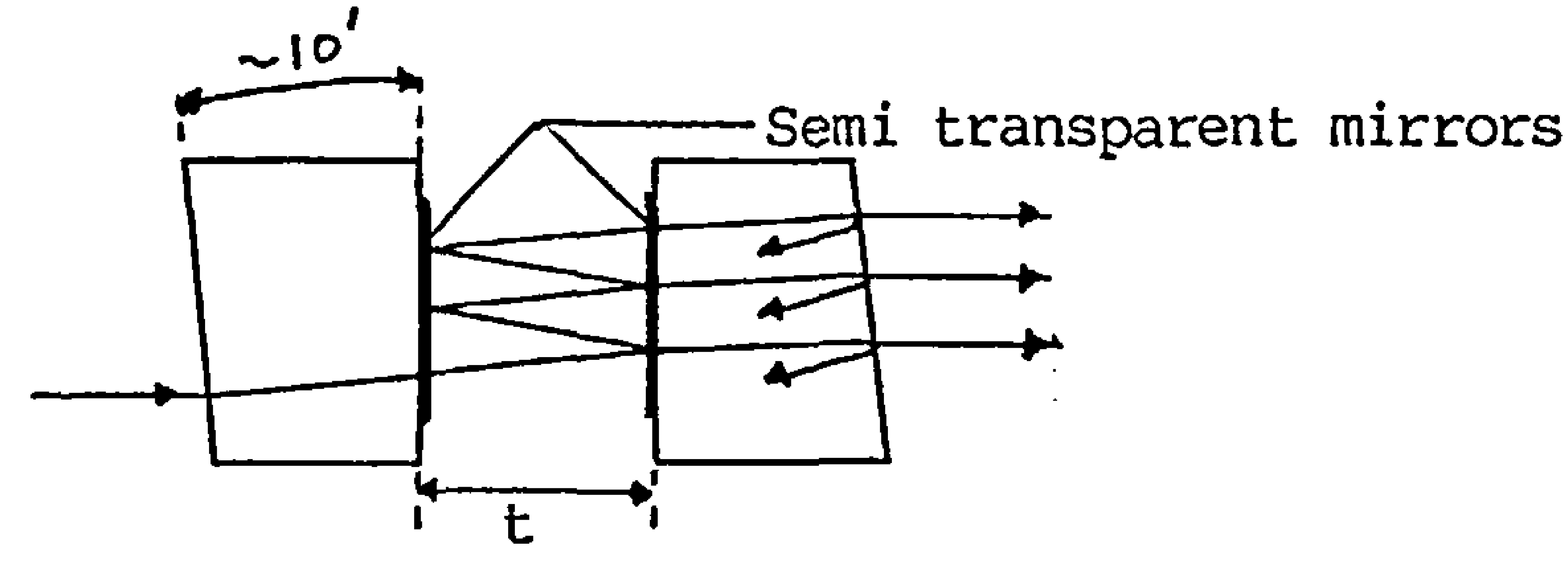


Figure 6.6. Fabry-Perot interferometer.

The Fabry-Perot interferometer is a plane-parallel optical

resonator, which transmits light at those wavelengths for which, under normal incidence, $n\lambda=2t$, where n is the order of interference and t the thickness of the gap between the two semi-transparent mirrors. The reflected light will be missing the wavelength which have been transmitted, and is subsequently analysed by a spectrograph, which produces dark fringes in the wavelength regime. A small change t of the airgap, due to height variations of the sample surface, will produce a fringe shift according to $\Delta t=n\Delta\lambda/2$. So the fringe shift is a measure for the height variation of the surface.

The camera scan of the image of one single interference fringe is fed to a signal averager and stored in a minicomputer. In spite of this automated data handling, the FECO interferometer is essentially a profiling instrument and hence acquires data slowly. A height sensitivity of a few angstroms and a lateral resolution of about 2 um, the resolution limit of the optical system, can be achieved.

The FECO interferometry has some advantages over the more common Fizeau interferometry:

i. Since no wedge angle is necesssary, there is no beam walkoff. All the multiply reflected beams can be collected, and extremely high finesse can be obtained when both surfaces are coated with silver films.

ii. The order of interference can be obtained unambiguously so there are never any errors of integral numbers of halve wavelengths.

iii. Since no wedge angle is required, very low orders of interference are possible. The combination i,ii and iii

makes possible the determination of surface roughnesses as small as 8 angstroms rms.

iv. The roughness of any portion of the sample surface may be observed merely by translating the entire interferometer without having to readjust a wedge angle; the latter is a difficult, delicate operation .

v. The measured irregularities on the fringe are independent of the figure of the surface. In Fizeau interferometry figure errors change the fringe spacing and can affect roughness measurements.

Heterodyne interferometry.

An heterodyne interferometer which produces a profile has been developed by Sommargren [1981]. When a surface is illuminated by two focussed beams of sligtly different frequency and the reflected beams interfere, the phase of the sinusoidal intensity modulation is related to the height difference between the illuminated points on the surface. If one of the beams remains focussed on a fixed point while the other beam is moved along the surface, height variations along the scanned line are measured. The basic component of the instrument is the quasi common path heterodyne interferometer shown in fig. 6.7 . The interferometer is comprised of a Wollaston prism, microscope objective and the surface being measured. The orthogonally linearly polarized colinear beams, with a frequency difference of 2

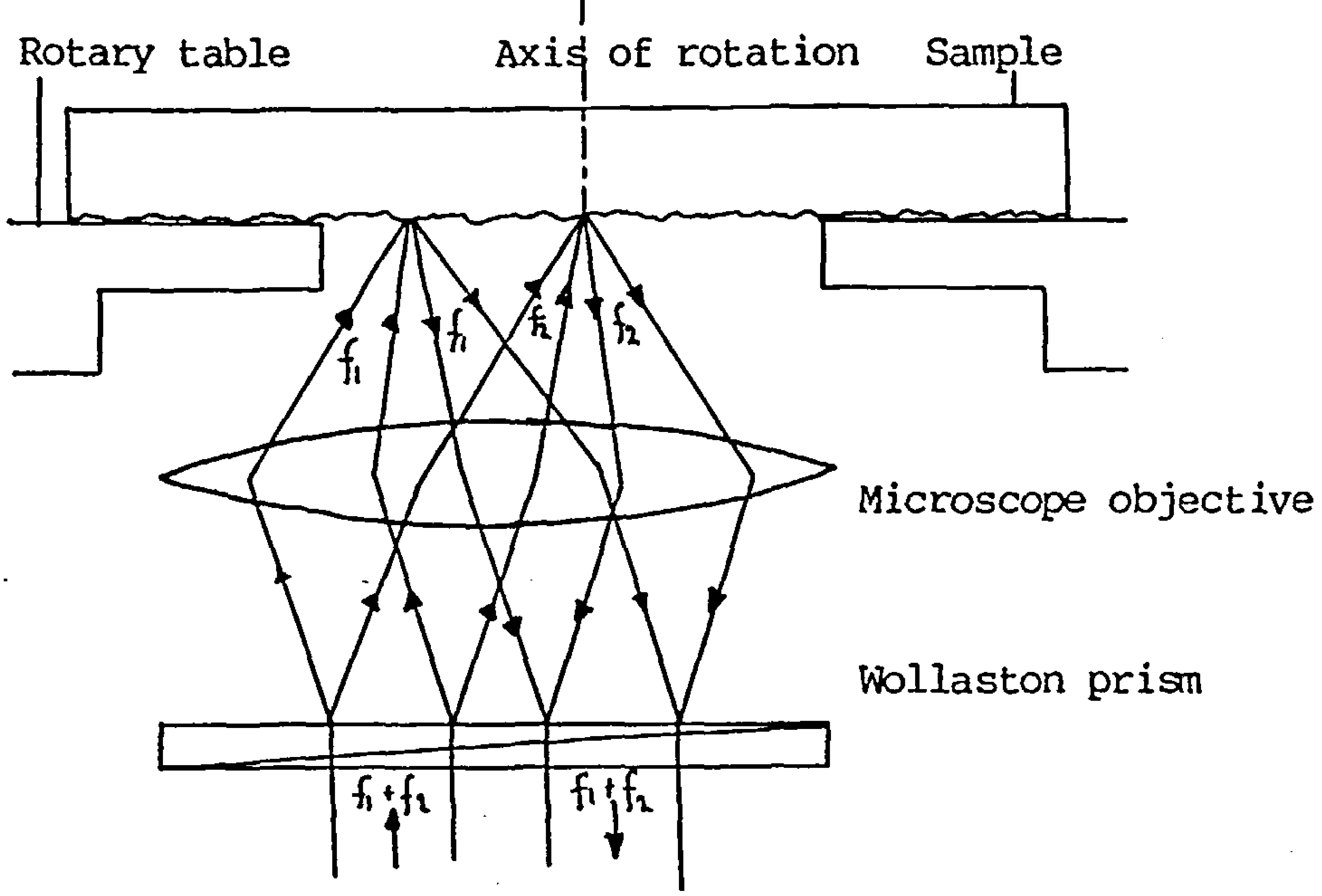


Figure 6.7. Detail of the ac-interferometer.

MHz,incident on the Wollaston prism, are refracted slightly, one to each side of the normal. The microscope objective brings them to focus at two distinct points on the surface, one on either side of the optical axis with one on the axis of rotation of the rotary table. The points are 2 um in diameter and separated by 100um. The reflected beams are recombined by the Wollaston prism and leave on a path parallel to but displaced from the incident beams.

As the table rotates the surface, one of the measurement beams is scanned in a circle along the surface. Changes in surface height, Δy, introduces an optical path change, Δz.
The measured phase change, $\Delta\phi$, is related to the optical path change by:

$$\Delta\phi = \frac{2\pi}{\lambda}\Delta Z \tag{6.5}$$

where λ is the wavelength.

Thus, the surface profile can be determined from a length standard λ and a measured quantity $\Delta\phi$. It has a height sensitivity of .1nm.

A similar instrument has been developed by Makosch & Drollinger [1984]. Theirs, however has a linear scan, and a single frequency. The light beams sensing the the sample surface are two colinear, othogonally polarized light bundles split from a He-Ne laser. The diameter of the laser spots on the surface is variable between between 1 and 10um. The optical phase difference of the two beams reflected from the test object is proportional to the height difference h between the illuminated points on the surface. Using an electrooptical phase modulation technique combined with a numerical calculation of the phase difference, this instrument is capable of measuring height differences with a resolution of 1 nm.

In the latter two systems, the specimen surface is referenced against itself, which makes the measurement almost insensitive to vibrations and imperfections of the sample scanning mechanism.

6.6.2. Geometrical optics.

Optical stylus.

The common factor of optical styli is, that they are based upon focussing a small spot (1-2 um) onto a surface while the spot is scanned across the surface. When there is relative, parallel motion

between the surface and the instrument, then the spot will remain focussed provided the surface is perfectly flat. However, when the surface has a certain amount of roughness, the spot will not remain perfectly focussed, i.e. a defect-of-focus occurs. This defect-of-focus can be sensed by means of suitable optics and electronics. The methods of producing and processing the defect-of-focus signal has lead to the development of two different categories of instruments.

i Direct conversion of the defocus signal into height information. This technique was originally reported by Dupuy [1967/68]. With the advent of more sophisticated electronics and laser diodes variants based upon this idea have been developed by Sawatari & Zipin [1979], Mignot & Gorecki [1983] Lou et al. [1984]. In general, the output signal is linear with respect to displacement within a range of 2 um about focus.

ii The other technique is based upon dynamic focussing. In this system, the defect-of-focus signal is employed to move the focussing lens such that the light spot is refocussed onto the surface. The position of the focussing lens is measured, which is directly related to the surface profile. A higher vertical range is possible compared with the previous technique. It is interesting to note that this type of instrument is a straight development from the optical styli originally designed for reading optical

compact discs. So a similar situation exists as before with mechanical styli for record players. The various variants of this technique are described by amongst others: Ertl [978], Arecchi et al. [1979], Fainman et al. [1982]

Light sectioning.

Uchida et al. [1979] constructed instruments based on a light sectioning microscope [Schmaltz 1936, Abbott and Goldschmidt 1937 Schlesinger 1942,Kayser 1943,Howes 1974]. In the light sectioning microscope the image of a slit is thrown onto the surface at an incident angle of 45 and viewed by a microscope objective at a reflected angle of 45 . The image of the slit will appear as a straight line if the surface is smooth, and as an undulating line if the surface is rough. The relative vertical magnification of the profile is the cosecant of the angle of incidence, in this case 1.4. Resolution is about .5um, and it is quite easy to measure peak-to-valley roughness.

The image need not be of a slit; a straight-edge, such as a razor blade, will suffice as object [Kayser 1943]. Shaw and Peklinek [1963] used a variation of this technique to measure the roughness of a razor blade edge by projecting its image onto an inclined screen. They reported discrepancies in the apparent magnification which they attributed to the finite thickness of the razor blade.

Taper sectioning

In taper sectioning, as its name implies, a section is cut through the surface to be examined at a shallow angle α, thus effectively magnifying height variations by a factor $\cot\alpha$, and subsequently examined by optical microscopy. Practical details are discussed at length by Nelson [1969] and by Rabinowicz [1950].

The main advantage claimed for taper sectioning is its accuracy; indeed Shaw and Peklinek [1963] have described it as "probably the most accurate method that has ever been devised for studying the profile of a surface". However, the same authors claim a vertical resolution of only .25 um.

Tarasov [1945] claims its ability of showing up deep scratches which a stylus will not penetrate. The only measurement which can conveniently be made from a micrograph of a taper section is the peak-to valley height, and Tarasov compares this measurement with rms roughness found by a stylus instrument for a number of surfaces. According to other results quoted by Shaw and Peklinek, a comparison of taper section profiles with those of a stylus instrument revealed larger peak-to-valley roughness in every case, up to a maximum of 100%. This technique is tedious and destructive.

Optical Fourier technique.

Whitefield [1975] has developed a scanning profilometer which is based on optical Fourier transforms. In this system, the surface is

illuminated by a divergent beam. After reflection from the sample surface one beam follows the diffraction analysis path that yields the signal representing the slope average over a localised area of the sample. The sample is located at the focal distance in front of the lens used as a Fourier transform lens. Integration of the output signal yields the surface profile. The author reports only results obtained from perfectly flat surfaces with one single scratch. The system is rather complicated and does not lend itself to be transformed into a rugged instrument.

Stereoscopic SEM.

Hillmann et al. [1984] describe the use of stereoscopic pairs of microphotographs obtained from a scanning electron microscope. The stereoscopic pair are produced by viewing the test surface under two different angles. The height information gained by this technique is very good. Fig. 6.8. shows the comparison of the surface profiles

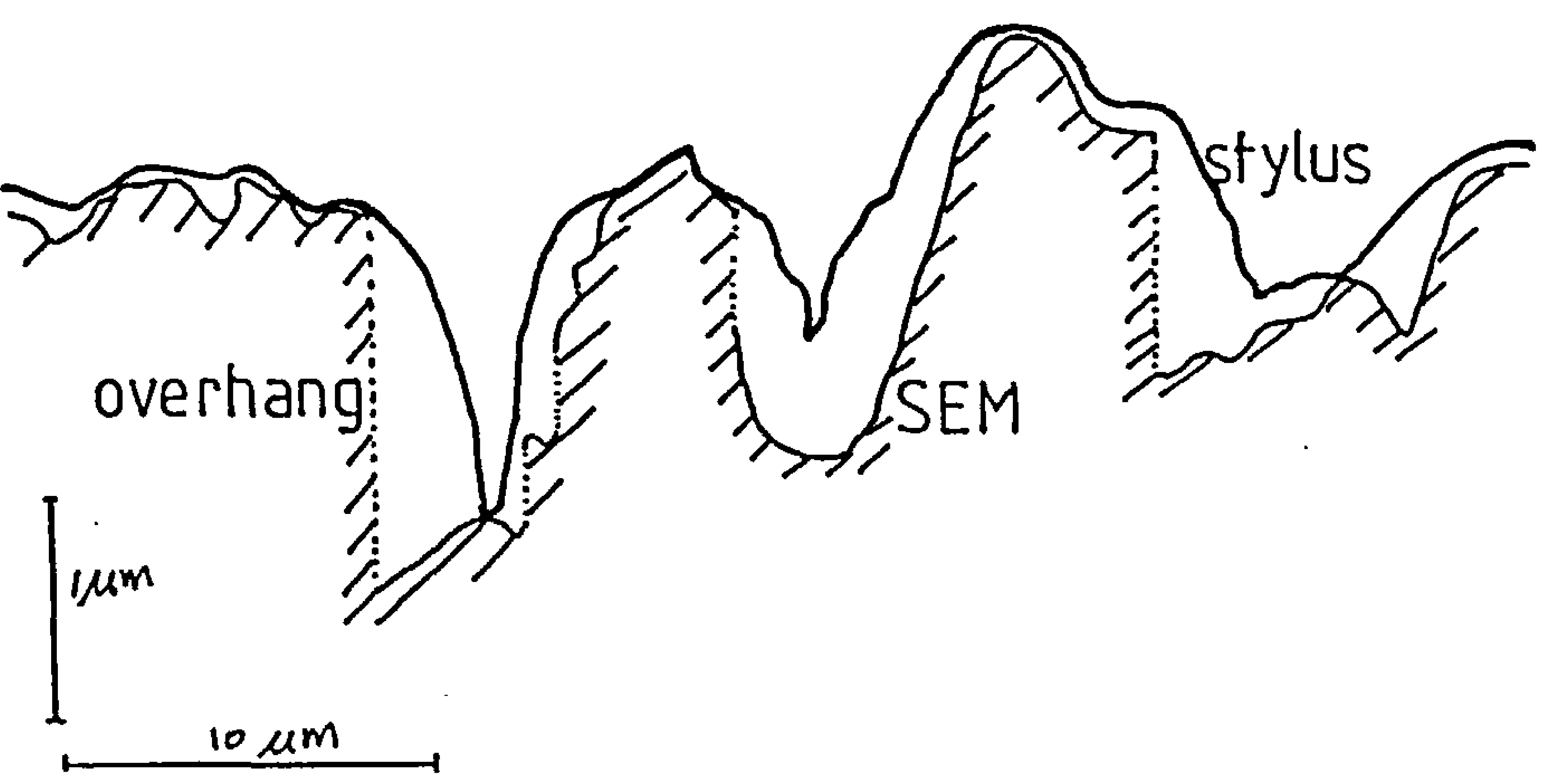


Figure 6.8. Profile obtained by SEM and contact stylus method.

obtained by a conventional stylus instrument and the stereoscopic technique. It shows an improved lateral resolution.

For rapid and routine assessment of surface finish, however, this technique has its disadvantages, since making stereoscopic SEM photographs is time consuming and the equipment is expensive and needs trained operators and also the samples must be small enough to fit into an electron microscope.

Triangulation.

Arecchi et al. [1979] image a 40 um pinhole onto the surface and then re-image it onto a photodiode array. The height of the surface is measured by triangulation. The apparatus can sample the surface every 10 um. The height sensitivity is .7 um and the measuring range equals 150 um.

Holography.

Surface profiles have also been obtained by Vienot & al. [1977] with a new technique known as temporal holography, which combines the methods of interferometry, spectroscopy, and holography to produce surface roughness profiles. This work is still in the experimental stage and more work is required before an adequate evaluation can be made.

Other methods.

Mitsui and Sato [1976] have developed an apparatus whose basics can be described as follows. Fig. 6.9. depicts the schematic lay-out.

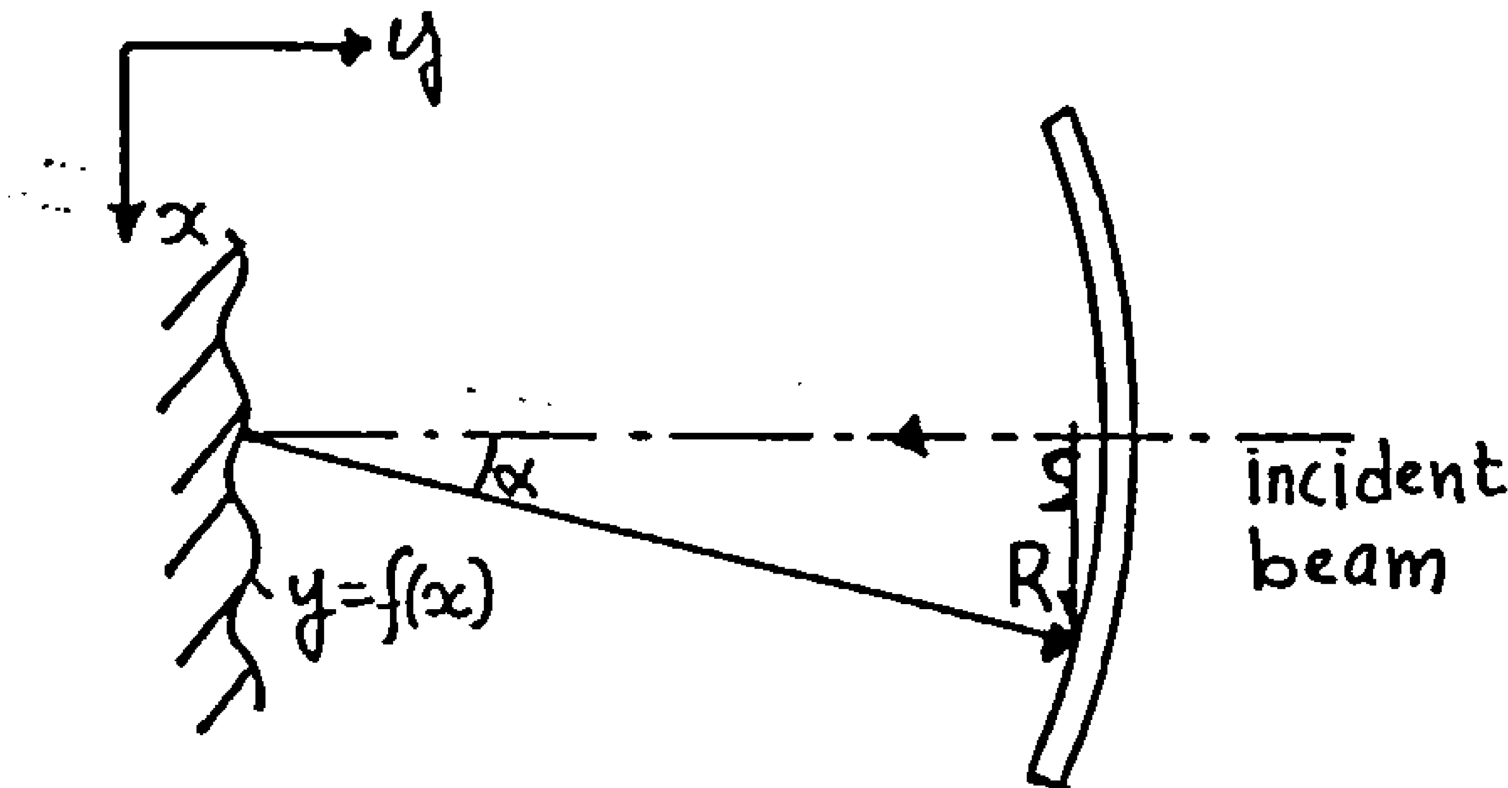


Figure 6.9. Detection of the surface slope.

The laser beam is directed horizontally onto the object surface so that the focus point may be obtained just on it. Photo diodes are lined up on a curved screen having a radius R with the reflecting point as its centre. It is assumed that the slope in the direction perpendicular to the plane of the paper is negligible. If the surface profile is expressed by y=f(x) using the coordinates shown in the figure, and the reflected light has an angle α from the incident beam, then

$$\tan\alpha = (-2f'(x))/(1-2f'(x)^2) \tag{6.6}$$

Assuming $f'(x)^2 \ll 1$ and $\alpha=\tan\alpha$, then $\alpha=-2f'(x)$. let S be the distance on the curved screen from the incident beam axis, then $S=-2Rf'(x)$, and

$$y = -\int (S/R)dx = -v\int (S/R)dt \tag{6.7}$$

for x=vt. This means that the shape of the surface profile can be obtained by detecting the position of the reflected light beam and by integrating it with respect to time. The measuring speed is given by the frequency characteristics of the photo diodes and the integration electronic circuit. The actual experimental system uses a beam with an angle of incidence of 45°.

Shiraisi [1983] has developed a system in which two beams of 10 um diameter are directed normally onto a turned surface. These beams are spaced so that one beam produces a spot on the peak of the turning feed mark, and the other beam produces a spot in the trough. An optical system is then used together with photo diodes in order to measure the radial position of these spots. The unit is attached to the toolpost of the lathe and by this means the spots are fixed relative to the peak and trough of the feed mark. During machining each spot scans along peaks and troughs individually. A resolution .5 um is claimed. No mention is made of the fact by which means the spots are positioned. Furthermore it seems to be only applicable to turning processes.

6.7. Optical Parametric techniques.

Most optical parametric techniques are based upon the angular scattering, when a collimated beam of laser light is reflected by a rough surface. This scattering can be caused by:

i scratches,digs and other surface features whose dimensions are large compared to the wavelength of light.

ii microirregularities, whose heights are much smaller than the wavelength of light, and covering the whole surface.

The scattering distribution from the first type of surface defects can, in principle, be calculated using the laws of geometrical optics. In practice, however, this is quite difficult because of the lack of a good model for the shape and size of the facets making up the surface features. The scattering caused by the second type can be treated by the application of physical optics. The light intensity distribution pattern of the scattered radiation depends on the roughness heights, the spatial surface wavelengths, and the wavelength of the illumination. In general, small spatial wavelength components diffract the light into large angles relative to the specular direction and large spatial wavelength diffract the light into small angles. For random surfaces, the power spectrum has a broad continuous spectrum of spatial components, and the light is therefore diffracted over a continuous range of angles centred about the specular direction [Bennett et al. 1961,1963] . Periodic surfaces, however, produce a discrete distributed optical diffraction pattern [Church et al. 1975,1977, Thwaite 1980,1982a,b and Hingle & Rakels 1983,1984]. Using a scalar scattering theory based upon the Kirchhoff diffraction integral, Davies [1954] and Bennett & Porteus [1961] have derived expressions for the relation between the angular light distribution and certain surface parameters. In their derivation the surface is represented by a statistical model having

the following properties:

i The rms roughness R_q is small compared with the optical wavelength.

ii The distribution of heights of the surface irregularities is Gaussian about the mean.

iii The autocorrelation function of the surface irregularities is also Gaussian with standard deviation a.

iv The surface has the statistical properties of stationarity and ergodicity with respect to the position along the surface.

Under these conditions, the following expression for the specular reflected component R_s is obtained, for normal incident illumination:

$$R_s = R_o \, e^{\left(-(4\pi R_q/\lambda)^2\right)} \tag{6.8}$$

where R_o is the specular reflectance of a perfectly smooth surface of the same material. The angular distribution of the diffusely scattered light $R_d(\theta)$ is expressed by:

$$R_d(\theta)d\theta = R_o \, 2\pi^4 (a/\lambda)^2 (R_q/\lambda)^2 (\cos\theta + 1)^4 \sin\theta \, e^{\left(-(\pi a \sin\theta/\lambda)^2\right)} d\theta \tag{6.9}$$

where $R_d(\theta)d\theta$ refers to the fraction of the reflected light which is scattered into an angle between θ and $\theta+d\theta$ at an angle θ from the normal to the surface.

From these expressions it follows that for very smooth surfaces ($R_q \ll \lambda$), most of the reflected light propagates in the specular direction. As R_q increases, the intensity of the specular beam decreases, while the scattered radiation increases in intensity (conservation of energy).

In addition, the angular distribution of diffuse radiation consists of a fine grainy stucture called speckle, which exhibits itself as an intensity contrast between neighbouring points.

Finally, the light wave may undergo a change in polarization upon reflection from the surface.

All of these phenomena, the relative intensity or reflectance in the specular direction, the total intensity of the scattered light, the diffuseness of the angular scattering pattern, the speckle contrast and the change in polarization, depend on the surface roughness, and all of them have served as the bases for potential surface roughness measuring instruments, albeit that some of them should be classified as comparators while other ones can produce absolute surface roughness values.

Generally, the methods based upon the measurement of the angular light distribution are limited by available theories to studies of surfaces, whose rms roughness R_q is less than $\lambda/30$. There are a few studies, mostly empirical, which have pushed beyond this limit [Tanner & Fahoum 1976]. With the ubiquitous HeNe laser as the illumination source, the above constraint dictates that these techniques can be applied mainly on optical quality surfaces, where R_q<20nm. This might appear to be a severe limitation, but in defence, these fine surfaces are difficult to measure with contacting

techniques without causing any possible damage, visibly or mechanically. Within this regime, they can provide high speed quantitative measurements of the rms roughness of both isotropic surfaces and those with a pronounced lay. The ultimate vertical resolution is 1 nm or better. For surfaces where $R_q/\lambda>.1$ this technique may be usefull as a comparator for monitoring both the amplitude and surface wavelengths properties as described by Brodmann et al. [1983a,b, 1984,1986a,b], Gast and Thurn [1984] and Rau et al. [1985a]. The information obtained is the slope distribution and the waviness of surfaces. The latter is obtained by measuring the positional change of the centre of gravity of the scattered light on the photo diode array in the sensor. They also report that this measurement method can be aplied to obtain 2-D surface lay information [Rau et al. 1985b].

There are no particular limitations on the part geometries which may be studied by these methods. Although most applications have been confined to external surfaces, internal ones can be accommodated.

6.7.1. Total Integrated Scatter.

This technique is only applicable to random surfaces with a Gaussian height distribution. The problem with surfaces consisting of a periodic structure is, that the specular component does not decrease exponentially with increasing roughness, but follows more or less the squared value of the Bessel function of the first kind of order zero $J_0^2(4\pi R_q/\lambda)$, with $4\pi R_q/\lambda$ as argument. From this it is obvious

that the intensity of the zeroth order is not a single valued function of the roughness R_q.

The simplest approach is to measure the specular reflectance of a rough surface [Bennett et al. 1961, Tanner et al. 1976a,b, Depew et al. 1971 and Chandley 1976]. Much of the initial research has been carried out by Davies and the group of Bennett. Commercial instruments following this general approach are often called glossmeters, although some of them measure in the non-specular direction as well and could be seen to measure the diffuseness of the scattered light. Several different arrangements for these types of measurement are described by Westberg [1967/68]. Peters [1965] outlines an in-process sensor employing white light scattering. His instrument measures the specular and a diffuse component, and was intended for use in grinding processes.

Total Integrated Scatter (TIS) refers to the integrated sum of light scattered into all directions within the scattering hemisphere. It is not necessary to be aware of the angular distribution of the scattered light, only of the fraction of the reflected light which is scattered out of the specular beam. This scattering arises from correlated irregularities which have heights which are small relative to the illumination wavelength.

The fraction of the total reflected light, i.e. specular plus scattered, scattered outside the specular direction by the surface roughness can be described by scalar scattering theory [Davies 1954] based upon the Kirchhoff diffraction integral. The dependence of TIS on the illumination wavelength and the surface roughness R_q is given by [Bennett & Porteus 1961]:

$$TIS = 1 - R_s/R_t = 1 - e^{(-(4\pi R_q \cos\theta_0/\lambda)^2)} \simeq (4\pi R_q \cos\theta_0/\lambda)^2 \qquad (6.10)$$

where R_t is the total reflected light.

R_s is the fraction, which is reflected in the specular direction θ_0 , the angle of incidence.

R_q is the rms roughness of the surface.

λ is the wavelength of the illumination.

The above formula expresses the fact that TIS is only a function of the height but not of the lateral separation of the irregularities, so long as they fall within an appropriate range where the scalar theory is valid. In most cases, the light scattering obeys the equation very well in the ultraviolet, visible and near infrared region of the electromagnetic spectrum. Fig. 6.10. shows a

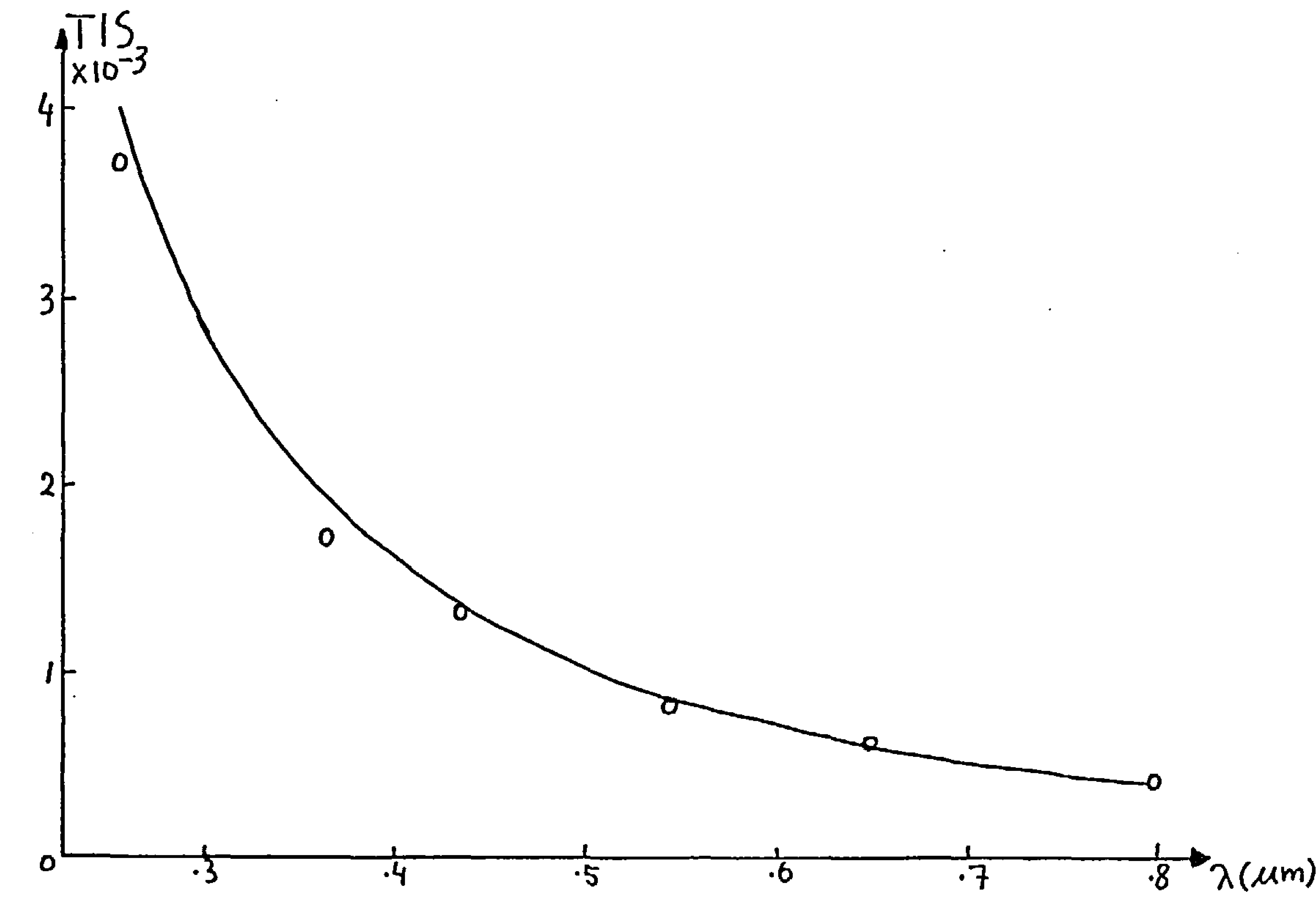


Figure 6.10. Wavelength dependence of TIS. o measured, __ calculated.

typical plot of the total scattered light as a function of wavelength in this region.

This technique relies heavily on the ability to detect the specular reflected component. This component effectively disappears when the surface roughness R_q exceeds $\lambda/10$.

Detrio and Miner [1985] have designed a standardized test for TIS measurements. This work was undertaken for the American Society of Materials and Testing ,ASTM. Their method has been applied by 8 laboratories on 4 test surfaces. The precision of the test was found to be better than 15% by this interlaboratory test program.

6.7.2. Angular Resolved Scatter.

Although TIS measurements are relatively simple to perform and generally satisfying for assessing optical surfaces, some applications require a more detailed knowledge of scattering into a particular angle or range of angles. For example, in a laser gyro system, where light is incident on the mirrors at 30 or 45 degrees, the retroscattered light, i.e. the light that is backscattered along the incident beam direction, is of crucial importance since it can cause the gyro to malfunction because of unwanted optical feedback in the laser resonator cavity. Another important disadvantage of TIS is that it cannot be applied to periodic surfaces.

Angular scattering from correlated surface microirregularities is most tractable to handle theoretically, when the heights are much smaller than the wavelength of light. There is, however, no simple

relation describing the angular distribution of scattered light as a function of the surface roughness, similar to the TIS relation. The angular distribution depends not only on the rms roughness as well as on the lateral spacing of the irregularities. A convenient measure of the random spacing of features is the autocorrelation function (ACF), and for periodic surfaces the spatial wavelength of the various harmonics which make up the surface structure.

The scalar scattering theory can provide information about the angular scattering distribution close to the specular direction, otherwise polarization effects will occur. From this theory it has been shown [Bennett & Porteus 1961], that for a surface with a Gaussian height distribution and a Gaussian ACF, the scattering into an angle measured from the specular direction is:

$$R(\theta) = R_o \cdot 16 \cdot \pi^4 R_q^2 a^2 \theta^2 / \lambda^4 \tag{6.11}$$

where R_o is the reflectance of a perfectly smooth surface of the same material.

R_q is the rms roughness value.

λ is the illumination wavelength.

a is the correlation length, the length at which the ACF has decayed to 1/e of the maximum value

This equation describes the scattering into a cone of half angle θ. This expression holds only for small scattering angles since it is the first term of an expansion [Porteus 1963]. This small angle approximation makes this formula independent of the functional form

assumed for the autocorrelation function, so it should also be applicable to surfaces which do not possess a Gaussian ACF.

Another approach of analyzing the angular distribution caused by light-scattering from surfaces with a random height distribution, is to consider the surface to be a superposition of sinusoidal gratings. These constituent gratings have random amplitudes, spatial wavelengths and relative phases, furthermore a one-dimensional lay is considered [Church et al. 1977,Stover 1976]. This surface profile y(x) can then be expressed by:

$$y(x) = \sum_{n=1}^{\infty} A_n \sin(2\pi x/d_n + \phi_n) \qquad (6.12)$$

where A_n is the amplitude of the n-th grating.

d_n is the spatial wavelength of the n-th grating.

ϕ_n is the relative phase the n-th grating.

The rms roughness R_q of this profile is:

$$R_q = \sqrt{\frac{1}{2}\sum_{n=1}^{\infty} A_n^2} \qquad (6.13)$$

It is assumed that each individual grating creates independently its own diffraction pattern. The angular light distribution of the diffracted orders m, of each grating, is defined by the familiar grating equation, assuming normal incidence:

$$\sin\theta_m = m\lambda/d_n \qquad (6.14)$$

where m is the diffraction order (m is integer)

θ_m is the diffraction angle

λ is the illumination wavelength

d_n is the spatial wavelength of the n-th grating

The amplitudes a_m of the diffraction orders, produced by the n-th grating, can be expressed by Bessel functions of the first kind of integral order m, $J_m()$:

$$a_{m,n} = R_o J_m^2(4\pi A_n/\lambda) \tag{6.15}$$

where R_o is the reflectance of a perfectly smooth surface of the same material.

Now assuming that the second and higher diffraction orders of each individual grating can be neglected, then the diffraction pattern can be treated as a modified power spectrum of the profile y(x). The amplitudes in the power spectrum are modified by the Bessel function. When $R_q < \lambda/30$, then equation (6.15) can be approximated by:

$$a_{m,1} = a_n = R_o \cdot (0.44 A_n/\lambda)^2 \tag{6.16}$$

Which means that the angular light distribution is a relative power spectrum of y(x). To obtain from this optical spectrum the powerspectrum, the value for R_o has to be evaluated, i.e. the total amount of light or energy reflected by the surface. This, however, is a simple integration of the complete diffraction pattern with the angle as the integration variable. After this it is a straightforward

matter of applying the formulae (6.13 and 6.16) to obtain R_q.

It should be noted that most of the literature is concerned with random surfaces.

In this thesis, however, formulae will be derived for periodic surfaces whose surface roughness can extend as far as $R_q/\lambda < 0.100$, thereby increasing the useful range of this technique considerably.

6.7.3. Ellipsometry.

Ellipsometry measures the change in the polarization state of a beam of light when it is reflected from a surface. Originally ellipsometry was mainly employed for the determination of the refractive index and thickness of materials. Fenstermaker and McCrackin [1969], however, investigated the influence of surface roughness on the errors in the measurement of the refractive index. The surface finishes ranged from 0 to 50 nm. From their experiments they came to the conclusion that even at R_q=5 nm, large errors can occur. In traditional null ellipsometry the important quantities are the angle of incidence of the illuminating wavefront and the angular positions of the polarizing and analysing elements (P and A in fig 6. 11) in the light path that produce a null output in the detector. These measurements enable the determination of R_p/R_s, the ratio of the complex reflection coefficients for the p- and s-components of the electromagnetic field. The electric field component of the p-wave

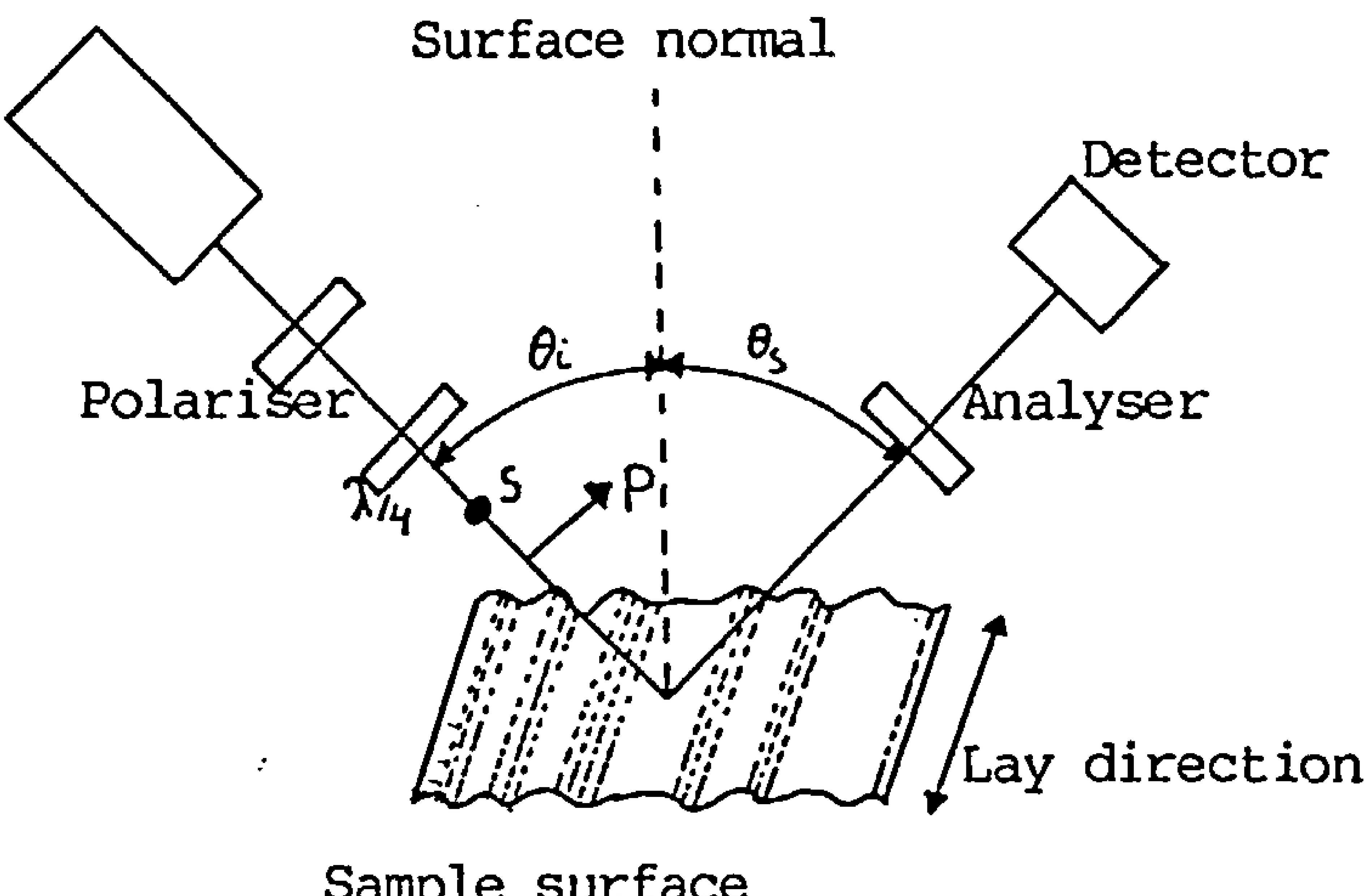


Figure 6.11 Schematic diagram of an ellipsometer.

is parallel to the plane of incidence, whereas the electric field component of the s-wave is perpendicular to the plane of incidence. The ratio R_p/R_s is customarily related to the following ellipsometry parameters, ψ and ϕ, by the definition:

$$R_p/R_s = \tan\psi \, e^{i\phi} \qquad (6.17)$$

where $\tan\psi$ is equal to the ratio of the magnitudes of the p and s reflection coefficients and ϕ is equal to the phase shift between them.

Ellipsometry measurements are sensitive to a number of surface properties including composition; surface structure such as damage, defects, or surface crystal faces; temperature; strain state; and surface roughness.

Since the roughening of surfaces can significantly change the results from the ellipsometer, thereby obscuring other surface changes to be observed, the question arises whether ellipsometry can be used to measure the surface roughness of engineering surfaces directly. The investigations sofar have been largely empirical and they are contradictory.

Lonardo [1974,1978, Chiesorin & Lonardo 1977] has studied ellipsometry as a potential tool for in-process detection of surface roughness. They experimented with ground and polished steel surfaces with R_a values ranging between 0.01 to 1.1 um and found that both ψ and ϕ decreased with increasing R_a. They defined and measured a parameter which combined the effects of both in order to increase the sensitivity of ellipsometry to R_a. Their parameter α, given by:

$$\tan(2\alpha) = \tan(2\psi) \cdot \cos(\phi - 180^\circ) \tag{6.18}$$

was found to vary almost linearly with R_a.

However, results obtained by Vorburger and Ludema [1980] show that the ellipsometric parameters cannot be simply related to surface parameters like R_q or S_q, the rms roughness and slope. They measured the ellipsometric parameters for six lapped and ground surfaces of a comparison standard typical of those used in the machine-tool industry, and they compared the results with surface parameters obtained by a stylus instrument.

To isolate the effects due to surface topography from other surface effects, the data was taken under four sets of experimental conditions with varying angles of incidence, optical wavelength, and

surface compositions. Since the trends for the four different sets of experimental conditions are quite similar, Vorburger and Ludema concluded that the experiment had isolated topographical effects in the ellipsometry results. The behaviour was very complicated, however. The results did not vary monotonically with R_a, nor did they vary monotonically with R_q, S_a or S_q. Similar behaviour was observed for ψ and for Lonardo's parameter on these surfaces. In general, it is not at all clear how the parameters of the surface geometry affect the ellipsometry results.

From these contradicting results it can be concluded that the use of ellipsometry as a tool to measure surface roughness would be problematic. This point is substantiated by results reported by Smith [1976].

Ohlidal et al. [1974] have experimentally shown the possibility of determining the ratio of the surface roughness and correlation length , R_q/T, very accurately, i.e. the rms slope of the surface finish. This is not a surprising result, considering that the Fresnel reflection coefficients for polarised light are slope dependent. They suggest to employ ellipsometry for obtaining the rms slope and TIS measurements for the determination of R_q.

6.7.4. Speckle.

When a rough surface is illuminated with partially coherent light, the reflected beam consists in part of random patterns of bright and dark regions, called speckle. These patterns can be

interpreted in terms of Huygen's principle whereby the intensity at a field point is caused by interference of wavelets, scattered from different points within the illuminated area, with their phases randomized by height variations of the surface. The spatial pattern and contrast of the speckle depend on the optical system used for observation, the coherence condition of the illumination, and the surface roughness of the scatterer.

The spatial intensity variations classified as speckle should be carefully distinguished from those variations termed the angular distribution as discussed in previous paragraphs. Speckle is the local intensity variation between neighbouring point or noise superimposed on the angular distribution. This speckle noise can be averaged out to obtain a clear angular distribution pattern.

Speckle may be observed either in the simple "Fresnel" (fig. 6.12a) or "Fraunhofer" (fig. 6.12b) configuration, the designations

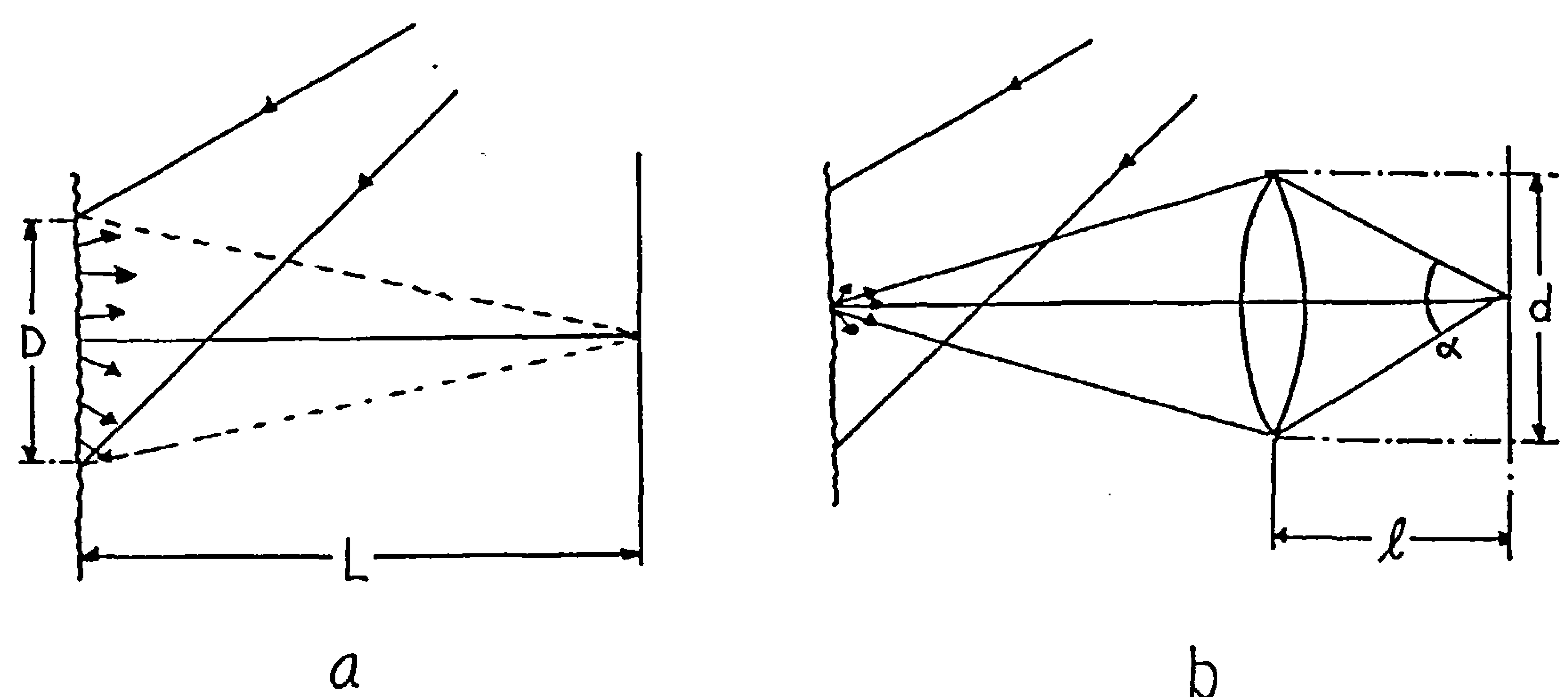


Figure 6.12. Illumination and recording configurations for speckles.

being given by analogy with regular diffraction phenomena. In figure

6.12a, the speckle pattern is found in the whole space in front of the optically rough surface and the screen AB has a section of the three-dimensional interference pattern projected onto it. For a circular patch of diameter D, illuminated by light of mean wavelength λ, the average size $\langle\delta\rangle$ of the speckles is given by the approximate formula:

$$\langle\delta\rangle = 1.2\,\lambda\,L/D \qquad (6.19)$$

where L is the distance from the illuminated surface to the screen.

When the interference pattern is imaged with an optical system such as depicted in fig b, the size $\langle\delta\rangle$ of the speckle is dependent on the angular aperture subtended by the exit pupil of the imaging system at the image plane, and $\langle\delta\rangle=1.2\lambda/\alpha=1.2\ l/d$, where l and d are as defined in fig. 6.12b. Thus for systems with apertures, the area of the object plane contributing to the speckle formation is not the entire illuminated area but just the area of the point spread function of the optical system.

The number of scatterers which contribute to the interference producing the speckle pattern is an important factor that determines the properties of the pattern. As pointed out by Dainty [1976], the minimum possible scatterer size is λ^2. For an aberration-free lens of numerical aperture, na, the area of the point spread function is λ^2/na^2. Thus the maximum number of scatterers contributing to a particular image point is $1/na^2$. For an illuminated area of diameter D, the maximium number of scatterers making a contribution is approximately dD^2/λ^2. In a majority of the applications discussed

below, it is assumed that a large number of scatterers contribute to the pattern. This condition is probably satisfied for most practical situations but will not hold for high resolution optical systems of highly focussed beams. If 100 scatterers is chosen as the lower bound for sufficiently large number to obtain acceptable statistics, then the properties of speckle patterns should be highly dependent on surface statistics for na greater than about 0.1 and illuminating diameters less than about 10λ. In these regimes, relationships between surface characteristics and speckle are not well understood but are active areas of investigation [Dainty 1975,1976 and Erf 1978]

The effect of surface scatterer motion on the properties of speckle patterns is particularly important for understanding the potential of speckle as an on-line measurement technique. The basic concept governing these effects is that a three-dimensional interference pattern fixed relative to the scattering surface exists in the space of the scattered illumination. If a surface is illuminated with a parallel beam of light at an arbitrary angle of incidence, then the speckle pattern follows any translation motion of the surface in its mean plane. The mean plane refers to the ideal flat surface about which the peaks and valleys of the surface topography are distributed and which give rise to the specularly reflected light beam. The bright and dark regions of the speckle patterns extend over three-dimensional cigar-shaped volumes of lengths $4\lambda/\alpha^2$. Object-image plane displacements over distances comparable to this dimension should, therefore, cause no first order changes in the speckle pattern.

On these basics, three broad classes of measurement methods are

based, for determining the roughness properties of a surface from speckle patterns. These are speckle contrast, speckle pattern decorrelation and speckle elongation. In all these methods, the roughness properties are obtained from the speckle patterns by empirically relating either the contrast or the degree of pattern correlation to the roughness of the surface under study. The empirical relationship is then interpreted with first order theories of speckle formation, which generally assume that only single scattering of the electromagnetic waves takes place and that the scattering surfaces can be characterized by a Gaussian distribution of heights with a correlation length much less than the dimension of the scattering region.

Speckle contrast.

In speckle contrast measurements the intensity variations are quantified in terms of an average contrast V defined as the normalised standard deviation of intensity variations at the observation plane. Let I(r) represent the intensity at the observation point r, then V is defined as:

$$V = \left[\langle I^2(r) \rangle - \langle I(r) \rangle^2 \right]^{\frac{1}{2}} / \langle I(r) \rangle \tag{6.20}$$

where the brackets <> indicate an average over a large number of intensity measurements. The intensity variations are determined by either moving the specimen thereby the speckle pattern past a fixed

detector with an aperture smaller than the speckle size or by moving the detector through the speckle field. High-contrast speckle patterns are produced when all the interfering wavelets have sufficient phase difference ($\phi > 2\pi$) to give complete destructive interference at some points in the pattern and the illumination has a high degree of spatial and temporal coherence. Spatial coherence means that the phases of the electromagnetic field at two points spaced across the propagating wavefront are highly correlated and temporal coherence that the phases of the field at two points spaced along the direction of light propagation are highly correlated.

Speckle contrast is unity for fully coherent monochromatic light illuminating a surface whose R_q is much larger than λ so that the wavelet phases are uniformly distributed over the interval from 0 to 2π. Correspondingly, for coherent monochromatic illumination, as the reflecting surface becomes smoother and less complete destructive interference occurs, the contrast decreases towards zero.

For any roughness value a decrease in speckle contrast occurs when the degree of either spatial or temporal coherence of the illumination is reduced. For temporal coherence the contrast decreases when the coherence length L_c of the illuminating wave field begins to have dimensions comparable to the roughness heights. In this regime as the roughness becomes larger than the dimensions of the region of coherence, an increasingly smaller number of scatterers give coherent contributions to the scattered field. If $\Delta\nu$ represents the frequency bandwidth of the broadband illumination, then $L_c = c/\Delta\nu$ where c is the speed of light; equivalently $L_c = (\lambda^2/\Delta\lambda)$ where λ is the mean wavelength of illumination and is the effective width of the

wavelength spectrum.

The effects due to decreasing spatial coherence are more difficult to conceptualize. Dainty [1976] showed that the speckle pattern produced in an image plane by partially spatially coherent illumination is simply the weighted sum of the intensities of speckle patterns produced by different angles of incidences of coherent illumination. The decrease in contrast for illumination with partial spatial coherence may then be thought of as arising from washing-out of the interference from this averaging over different angles of effective incidence. Figure 6.13. summarizes the combined effects of different degrees of coherence and scatterer surface roughness on speckle contrast.

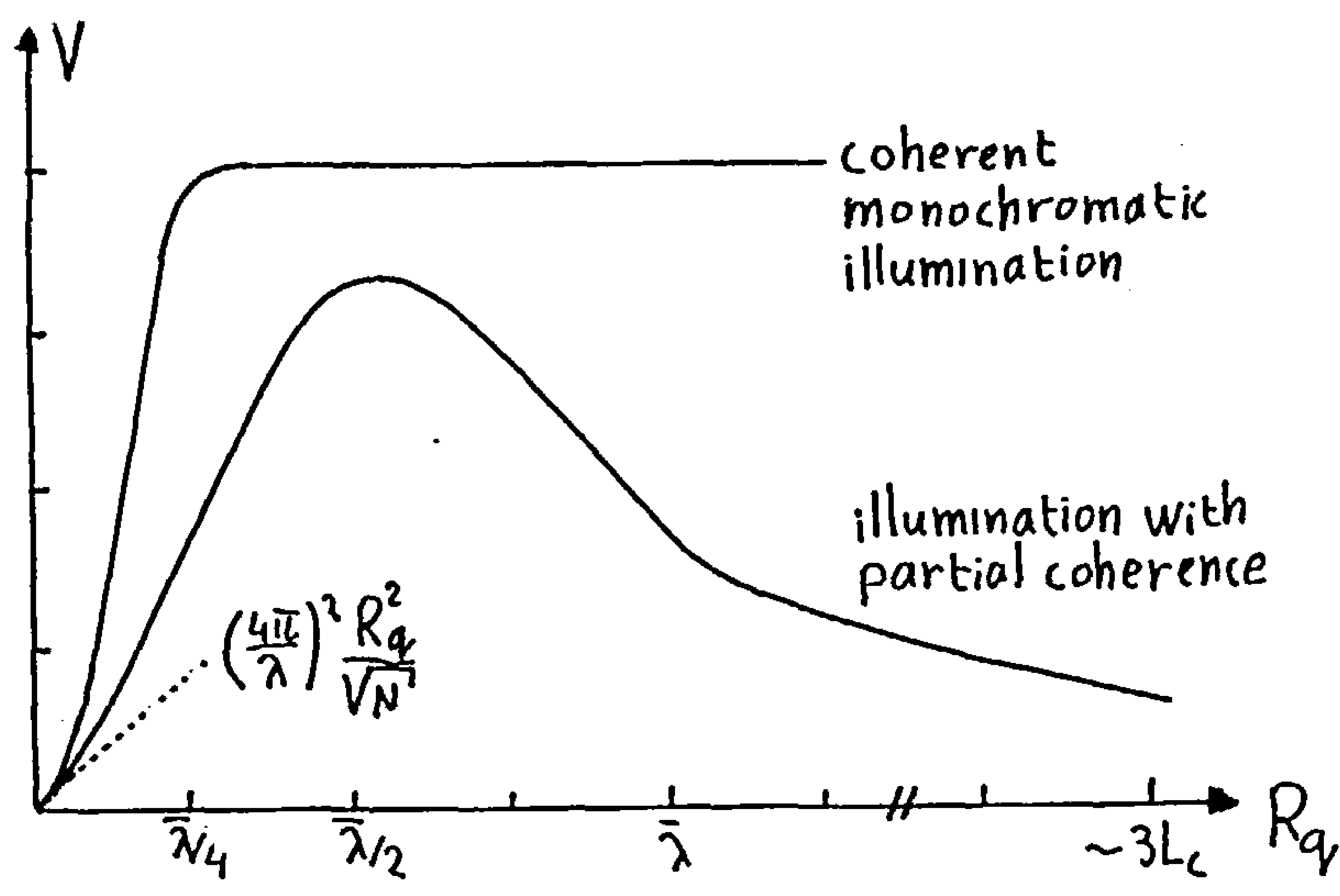


Figure 6.13. Variation of speckle contrast .

The first measurements to relate polychromatic speckle contrast and surface roughness were performed by Sprague [1972]. The light source in his experiment in combination with the spectral response of his photodetector produced a mean wavelength of about 0.5 um and a

coherence length of about 1.5um. From intensity variations as a function of specimen position Sprague calculated the arithmetic average of absolute deviations from a mean value rather than the rms V. Results of his measurements show a good monotonic relation between polychromatic speckle contrast and surface roughness as determined with a stylus instrument for roughnesses from about $\lambda/2$ up to about $2L_c$, i.e. 0.2um to 3.2 um. In contrast to the specular reflectance and diffuseness measurements, the correlation of contrast to roughness seems to be relatively independent of the type of finishing process used to form the surface.

A modified version of Sprague's experiment has been reported by Leonhardt and Tiziani [1982]. Their claim is that the white light speckle contrast can be improved by superposition of an incoherent light intensity onto the speckle pattern. Ahlers [1985] describes a prototype of an instrument based upon this superposition of white light onto the speckle pattern. The results are not very convincing, since it is difficult to decide how much light has to be superimposed, however it should be useful as a comparator.

Experiments to relate surface roughness to the contrast of speckle patterns produced by coherent monochromatic illumination were first performed Fujii [1978] and Asakura [1978]. Results of their contrast measurements show that a strong, linear correlation exists between V and R_a, determined by stylus profilometry, for R_a values up to 0.13um.

Speckle pattern decorrelation.

The second class of techniques for relating surface roughness and speckle is speckle pattern decorrelation measurement. Here two speckle patterns are obtained from the test surface by illumination with different angles of incidence or different wavelengths of light. Correlation properties of the speckle patterns are then studied by recording the patterns on the same photographic plate by double exposures [Mendez et al. 1975, Tribillon 1974 and Leger et al. 1975], or by photoelectric detection of the two patterns [Fujii et al. 1978, Giglio et al. 1979, Wykes 1977 and Leger et al. 1976, Ruffing 1986]. The primary attribute of this type of speckle measurement is that R_q values as large as 30 to 50 um can be measured.

Theoretical work by Parry [1974] and Goodman [1976] form the basis for decorrelation measurements. The basic idea for intensity correlation measurements can be obtained from an equation derived by Parry for the intensity cross-correlation function, γ, which is a measure of the similarity of the two speckle patterns, is defined by:

$$\gamma = e^{\left[-R_q^2 (k_1-k_2)^2\right]} \; e^{\left[-\beta |k_1 r_1 - k_2 r_2|^2\right]} \tag{6.21}$$

where $k_i=(2\cos\theta_i/\lambda_i)$, β is a constant proportional to the laser spot size on the scatterer, and r_i describes the observations points in a far field plane. θ_i is the angle of incidence of the light of wavelength λ_i. Thus, the value of γ is determined by the two radial scales, k_1 and k_2, of the speckle patterns. If $k_1>k_2$, the speckle pattern at k_1 is a radially contracted replica of the pattern for k_2.

The degree of replication depends on the first exponential, which is controlled by the surface roughness of the test specimen. If $R_q^2 (k_1-k_2)^2 << 1$, a high degree of correlation is found and a marked radial structure appears. If $R_q^2 (k_1-k_2)^2 >> 1$, the two patterns are totally uncorrelated and the radial structure disappears. As regards the second factor, experimental conditions are generally such that $k_1 r_1 = k_2 r_2$ is maximised. The large range of R_q for the technique is clearly determined by the experimental capabilities for producing small values of k_1-k_2 . Different values of k are obtained by either fixing θ and varying λ or vice versa.

From a theoretical point of view, this seems a more reliable application of speckles to the measurement of surface finish than the speckle contrast method.

Speckle elongation.

Stansberg [1979] has developed a method of analysing the shape of the speckles, which are recorded on photographic film. The speckle pattern has generally a radial structure with the optical axis of the system as its centre. The speckle shape depends on its radial position and the surface roughness of the test sample. This speckle elongation is assessed by means of Fourier optics. The experimental and theoretical results compare favourably for $10 < R_q < 20$ um. Below 10 um the experimental and theoretical results start to deviate. Above 20 um, the speckle elongation seems to be constant.

6.7.5. Interferometric Technique.

Motycka [1969] and Ribbens [1972,1974] have proposed variations of the interferometric technique which rely on the fringe visibility caused by the interference of holograms and the reflection of surfaces, to determine the surface roughness. The system proposed by Ribbens [1974] uses two wavelengths in order to increase the useful surface finish range. This system uses an effective optical wavelength equal to $(1/\lambda_1 - 1/\lambda_2)$, λ_1 and λ_2 are the two wavelength used. The roughness parameter is related to the interference fringe contrast ratio. This technique is still in experimental stages and the method is rather cumbersome.

6.8. Comparison of the optical techniques and their suitability for in-process application.

As with mechanical techniques, there is a trade off between accuracy and speed of measurement. The parametric are generally faster than the profiling (interferometry and optical stylus) methods. However, the operation of the parametric techniques is based upon various assumptions about their height, slope distribution and or the correlation length. These instruments will still yield a result, even if the operating conditions are not met, but these results are then very suspect. This is very much the case with ellipsometric measurements, since not only is the measurement by the surface finish parameters but by other material properties as well.

This problem does not occur with profiling techniques, which indeed will produce very accurate profiles, from which the various parameters can be calculated.

Table 6.2. shows a comparison of the various techniques discussed before. The heading "ease of use" takes into account the sample preparation which is often necessary for the interferometric measurements. The speed of operation is mostly governed by the number of detectors which have to be sampled and whose data needs processing. ARS stands for Angular Resolved Scatter.

Few of the optical techniques are suitable for in-process measurement. All the profiling techniques need careful setting-up of the instrumentation in a vibration free environment, which excludes them from in-process measurement. From the parametric techniques only the light scattering method is suitable for in-process measurement, since the scattering distribution is independent from surface movement as long as the surface characteristics are the same all over the surface. They can also cope with a fair amount of vibration. In speckle methods again only stationary measurements can be made and some need careful alignment of the optical system.

Table 6.2. Comparison of optical techniques.

	horizontal		vertical		vibration	ease of	assumptions	speed
	resolution	range	resolution	range	problems	use		
Interferometry	1 nm	1 um	2 um	2 mm	2	3	none	1000
Optical stylus	1 nm	50 um	2 um	10 mm	1	2	none	1000
TIS	1 nm	0.1um	20 um	5 mm	0	1	few	2
ARS	1 nm	0.1um	20 um	5 mm	0	1	few	100
Speckle contrast	10 nm	2 um	10 um	1 mm	0	2	few	100
Speckle correlation	100 nm	20 um	10 um	1 mm	1	2	few	10
Ellipsometry	?	?	?	?	0	1	many	2

7. Comparison of the various scattering theories.

As can be concluded from chapter 6, the most promising techniques for the determination of surface roughness parameters, which can be applied in-process, are those based on light scattering, namely Total Integrated Scatter (TIS) and Angular Resolved Scatter (ARS). The TIS technique can be considered a simplified ARS method.

In general the ARS technique is preferable, since it can produce more surface information than TIS. TIS can only produce the R_q roughness parameter of surfaces with a random height distribution, whereas ARS can produce slope, Δ_q, and surface wavelength, λ_q, parameters as well as the R_q value for both periodic and random surface profiles. However, TIS measurements are generally faster to make and therefore more suitable for quality control applications [Bennett,1978], since detoriation of a machining process will in general produce a change in surface finish.

To predict the angular scattering from rough surfaces, various approaches have been employed. These include geometrical optics, scalar and vector diffraction theories, and methods based on numerical methods. These various approaches will be briefly discussed and then their relative merits for the determination of surface finish will be compared. In chapter 8 a more detailed discussion will follow on the scalar theory on which the operation of the surface finish sensor is based.

7.1 Various scattering theories.

7.1.1. Geometrical optics.

Geometrical or ray optics can be applied to light scattering problems when the dimensions of the surface roughness are large compared to the wavelength of light. A typical geometrical optics approach assumes that the scattering surface consists of a system of connected plane facets, having lateral dimensions that are large relative to the illumination wavelength and reflecting like plane mirrors. Furthermore it is assumed that the diffraction effects at the edges of the facets are negligible. Each facet has a surface normal vector and a value of surface area, and there will be a statistical distribution of these two quantities. Thus, scattering from a geometrical optics type of rough surfaces consists of adding the contributions from each tilted facet making up the surface. In order to actually treat a surface using this approach, one needs to have a characterisation of the surface: either a model of the shapes and their distribution of the surface facets or an equivalent statistical characterisation. A statistical model of a facet surface generated by a Markov chain has been given by Beckmann and Spizzichino [1963].

Geometrical optics fail to describe the effects of diffraction. Polarisation effects can be taken into account by introducing the Fresnel reflection coefficients. The geometrical optics approach can be considered as an extension of the scalar diffraction theory from surfaces whose roughness is much larger than the wavelength of light. In both cases the scattered light contains information about the surface slope distribution and not of the surface heights [Lynch 1970].

7.1.2. Scalar diffraction theory.

The Helmholtz-Kirchhoff diffraction integral is the starting point of the scalar theory treating scattering from rough surfaces. This integral and the resulting diffraction formula expresses the value of the wave equation at some observation point in space, surrounded by an arbitrary closed surface (a mathematical construction), in terms of the values of both the electromagnetic field and its normal derivative at the rough surface. So, when applied to scattering from rough surfaces, strictly speaking, the field must be known at the rough surface itself. However, generally this field is not exactly known and to overcome this problem certain approximations, known as the Kirchhoff boundary conditions are made.

Davies [1954] has derived a relation predicting the angular dependence of the scattered light from rough surfaces at angles near the specular direction. He assumed that the surface was perfectly conducting, i.e. that the specular reflectance of a perfectly smooth surface or the total reflectance of a rough surface of the same material was 100%. The surfaces were also assumed to have Gaussian height distributions functions and Gaussian autocovariance functions. Bennett and Porteus [1961] modified Davies' relation to include the actual reflectance of the material and then experimentally verified the validity of their relations for prediction of the specular reflectance of rough surfaces with $R_q \ll \lambda$. Porteus [1963] expanded the relations to include arbitrary height and autocovariance functions, however, they had to be random surfaces. The restriction that the roughness had to be much smaller than the wavelength was

also removed.

Beckmann [Beckmann & Spizzichino 1963] has also derived a scattering formula based upon scalar diffraction. His formula has the advantage that various autocovariance functions of the surface profile can be introduced to predict the scattered light distribution as well as the specular reflectance. Welford and Chandley [1975] have reformulated his results into conventional optical terms. Like Davies, Beckmann was interested in radar wave scattering. The Beckmann theory is based on geometrical optics for the determination of the field at the scattering surface, subsequently this field is used as the input for the Kirchhoff scalar diffraction integral. Chandley [1976a,b] has successfully verified the Beckmann theory for ratios of R_q/λ up to 0.6.

Theoretically the validity of the scalar theory is limited to scattering angles near the specular direction. However, experimental results have shown that in practice reasonably large scattering angles are no problem as has been shown by Welford [1977]. Polarisation properties of light scattering are not included.

Scalar theories have been applied to surfaces whose roughness is much smaller than the illumination wavelength. According to Church [1979] this limit is $R_q/\lambda < 0.01$. However, this is a pessimistic estimate, based upon first order scattering. In this thesis algorithms are devised for periodic surfaces, which are valid for a surface roughness to wavelength ratio ,R_q/λ , up to 0.100. These algorithms have been tested on actual scattering data and have been proven to be very accurate [Rakels,1986c]. For random surfaces the useful surface roughness range is up to $R_q/\lambda < 0.2$ [Welford, 1980].

7.1.3 Vector diffraction theory.

To include the vector nature of electromagnetic fields, a somewhat different technique has to be employed than with the scalar theory. There are two main methods. The first is a variational or perturbation method, of which there are many variations. The other is analogous to the Helmholtz-Kirchhhoff scalar integral and is its vector equivalent, known as the Stratton-Chu-Silver (SCS) integral [Stratton 1941, Silver 1949]. Leader [1971] has rewritten the SCS integral for the scattered field $\vec{E}_S(P)$ at a point P in the $\vec{n}_2$ direction as:

$$\vec{E}_S(P) = K\vec{n}_2 \times \iint_S [\vec{n}\times\vec{E} - \eta\vec{n}_2 \times (\vec{n}\times\vec{H})] e^{ik\vec{r}\cdot\vec{n}_2} dS \tag{7.1}$$

where $\vec{r}$ is the position vector from the origin to the surface-area element dS in question, k is the wavenumber, $K=ike^{(-ikR/4\pi R)}$, with R the distance from the origin to the field and $\vec{E}=\vec{a}E_0e^{(-i\vec{k}_1\cdot\vec{r})}$ with $\vec{a}$ as the polariztion vector. η is the free-space intrinsic impedance and $\vec{n}$ the local normal vector. Furthermore, the terms $\vec{n}\times\vec{E}$ and $\vec{n}\times\vec{H}$ can be expressed by:

$$\vec{n}\times\vec{E} = [(1+R_1)(\vec{a}\cdot\vec{t})(\vec{n}\times\vec{t}) - (\vec{n}\cdot\vec{n}_1)(1-R_2)(\vec{a}\cdot\vec{d})\vec{t}]\, E \tag{7.2}$$

and

$$\vec{n}\times\vec{H} = -[(1+R_2)(\vec{a}\cdot\vec{d})(\vec{n}\times\vec{t}) + (\vec{n}\cdot\vec{n}_1)(1-R_1)(\vec{a}\cdot\vec{t})\vec{t}]\, E/\eta \tag{7.3}$$

where R_1 and R_2 are the Fresnel reflection coefficients for the polarization vector $\vec{a}$ lying normal to, and in the plane of incidence,

respectively. $\vec{n}, \vec{t}$ and $\vec{d}$ form a local orthonormal coordinate system given by:

$$\vec{t} = (\vec{n}_1 \times \vec{n}) / [1 - (\vec{n}_1 \cdot \vec{n})^2]^{\frac{1}{2}} \quad , \vec{d} = \vec{n} \times \vec{t} \tag{7.4}$$

where $\vec{n}_1 = \vec{k}_1 / |\vec{k}_1|$ with $\vec{k}_1$ as the propagation vector of the incident wave.

The SCS integral was originally applied to micro-wave diffraction problems. It is based on the calculation of the radiation caused by a distribution of currents over a surface [Kalhor and Neureuther 1971]. The integral fully contains the vector nature of the currents (based upon the electromagnetic field at the surface), and yields the radiation from this distribution of surface currents. In principle, the integral can yield the scattering from any magnitude of surface roughness or shape, but in practice the results are often limited to roughnesses which are small compared to the wavelength of the illumination. One major reason for this limitation is that the actual field at the rough surface is not known (as is necessary to properly evaluate the integral) and, consequently, approximations have to be used.

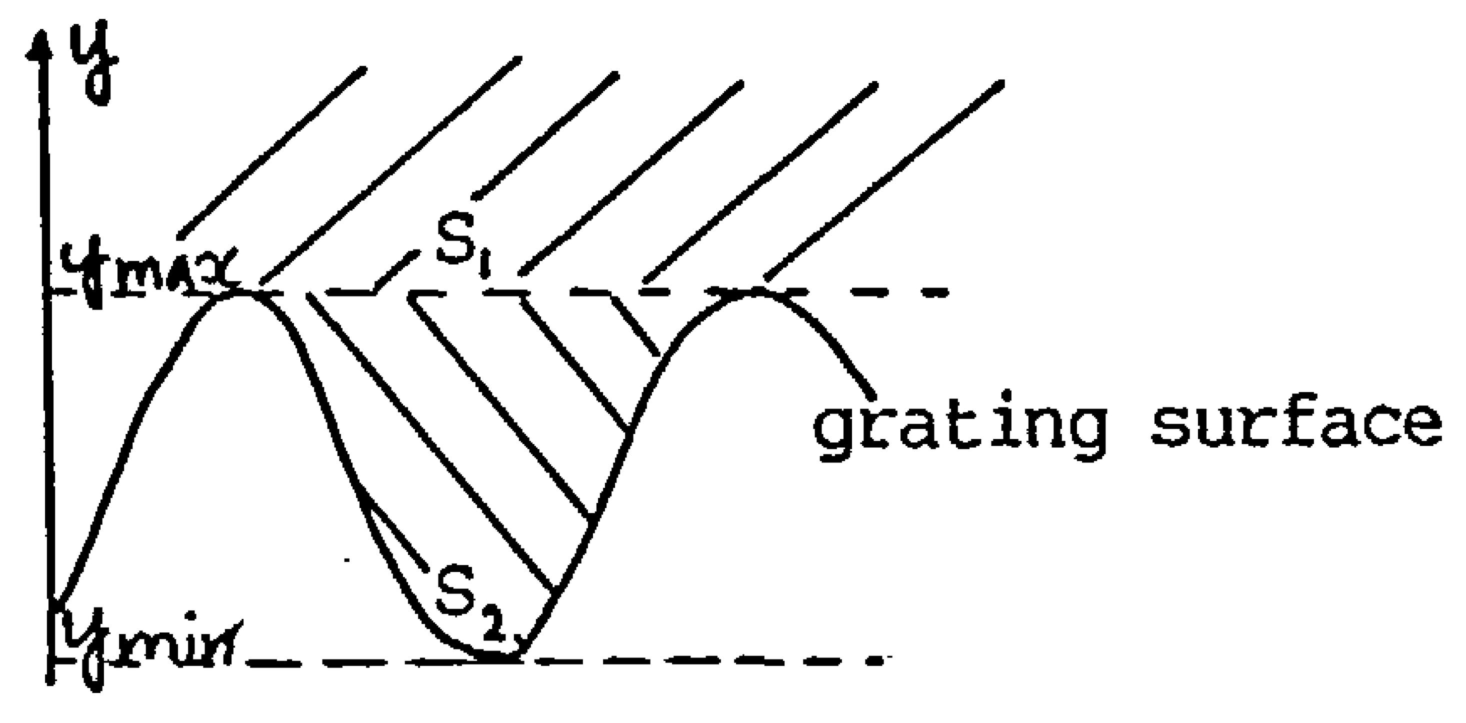
The variational technique [Meecham 1956, Stroke 1960], which was first applied by Rayleigh [1945] for acoustic scattering, depends again on the surface having a small roughness value compared to the wavelength of light, so that the incident field at the surface can be approximated by the fields which would be present at a smooth surface. This method is suitable for periodic surfaces with a one directional lay [Fortuin 1969]. In this method a linear combination of known solutions to the wave equation is chosen to represent the

reflected field. The coefficients will be chosen here so that they minimize the square of the error in the boundary conditions. These boundary conditions depend on the polarisation of the field. For an incident plane wave with its propagation vector lying in the plane formed by the surface normal and the direction across the lay and which is polarised so that the electric vector is perpendicular to that plane, the boundary condition is that the total (incident plus reflected) electric field is zero at the surface boundary. For incident radiation polarised so that the electric field is parallel to the one-directional lay, the boundary condition is that the derivatives of the total electric field, taken normal to the surface, are zero. The directions of these fields, which make up the total reflected field, are determined by the grating formula:

$$\sin\theta_m = \sin\theta \pm m\,\lambda/D \tag{7.5}$$

where θ_m is the diffraction angle, θ the incident angle, m is the order of diffraction and D is the period length of the grating.

This variational method has caused quite some controversy, concerned with the so-called Rayleigh hypothesis. This hypothesis states that the reflected waves between the corrugations of a grating (figure 7.1) and in the region above the grating can be expressed by



S_1 denotes the domain $y_{max}<y<\infty$

S_2 denotes the domain $y_{min}<y<y_{max}$

Fig. 7.1. Grating configuration.

the same series of plane waves having directions as determined by the grating formula. The validity was first argued by Lippmann [1953] and later by Petit [1966]. Following Petit and Cadilhac [1966], Millar [1968,1970a,b,1973] investigated the conditions under which the Rayleigh hypothesis can be applied for a sinusoidal grating. Their conclusion is that when the grating is expressed by $y(x)=h\cos(2\pi x/D)$, the hypothesis is valid for $h/D<0.071$. The validity of the hypothesis for more general type of gratings have been reported by Neviere and Cadilhac [1970] and van den Berg & Fokkema [1978]. In physical terms, the condition under which the Rayleigh hypothesis is valid is governed by the period length, D, of the grating and its amplitude,h. The period length determines under which angles the reflected waves will leave the surface and the amplitude determines the magnitude of these waves. Now when h is large and D is small relative to wavelength of light, then the reflective waves existing in the corrugations will hit the grating surface thereby causing secondary diffraction effects modifying the finally reflected waves and thereby invalidating Rayleighs hypothesis.

Wirgin [1980a,b,1982] has shown that for h/D ratios as high as 0.35, provided that D/λ is inferior to about two, the Rayleigh method will yield numerical solutions that are close to the rigorous integral method.

7.1.4 Numerical methods.

All of the surface scattering theories discussed previously can

yield results in closed form. To obtain these results it is necessary to make various approximations and to limit the ratio of the rms roughness to the wavelength of light. With the advent of high speed computers, numerical solutions have become a practical alternative. The usual method involves a solution of the wave equation for the reflected wave as an infinite series expansion, similar to the variational method. The coefficients are determined numerically by solving equations for the appropriate boundary conditions. The majority of the solutions that have been obtained by numerical methods are for scattering from gratings. This technique is especially useful when applied to high efficiency gratings where the grating amplitude is not small compared to the wavelength of light. These are exactly those cases which are difficult to cover by the classical theories. It is usually possible to truncate the numerical algorithm after the required accuracy has been achieved. Some authors have calculated the diffraction caused by a colinear array of line currents [Yakovlev 1964, Zaki & Neureuther 1971]. Several different methods (Rayleigh, integral and differential) are outlined by Petit [1980]. Most of the initial techniques are actually based on the Rayleigh method [Jiracek 1973, Ikuno & Yasuura 1973, Kalhor 1976 and Whitman and Schwering 1977]. Because of the problem of specifying the exact surface profile, numerical methods have not been often applied to the scattering from random surfaces. One of few treatments of random surfaces can be found in contributions by Axline & Fung [1978] and Maystre [1983,1984]. In general the numerical solution leads to large matrix equations whose solutions do not give much insight into the features of the scattering process. However their accuracy can be

very high, depending on number of terms taken into account in the series expansion of the reflected wave.

7.1.5. Comparison of the scattering theories.

No one single method can be singled out as the ultimate for the calculation of the angular distribution caused by scattering from rough surfaces. Indeed, when the surface geometry is known, then the various theories can predict the light scattering, provided that the assumptions made for their derivation are not violated. However, an instrument for the determination of surface finish, has to be able to extract the surface finish parameters from the scattered light distribution, which is in effect an inverse scattering problem. So, in this case, the criterion on which to base the preference of any of the theories is: "Can this theory be used to extract the surface finish parameters from the scattered light?"

This requirement eliminates two of the discussed theories, namely the geometrical and numerical methods.

The geometrical method is based on the size and slope distributions of facets constituting the surface. Here the direction of scattering is governed by the facet slopes and directional intensity by the facet sizes. The information about the sequence in which these facets occur is not available. Because of this the surface roughness parameters R_a and R_q cannot be determined. This can be demonstrated by the simple surface models of figure 7.2.a,b. Here the two surfaces consist of 4 equal size facets of unity length, of

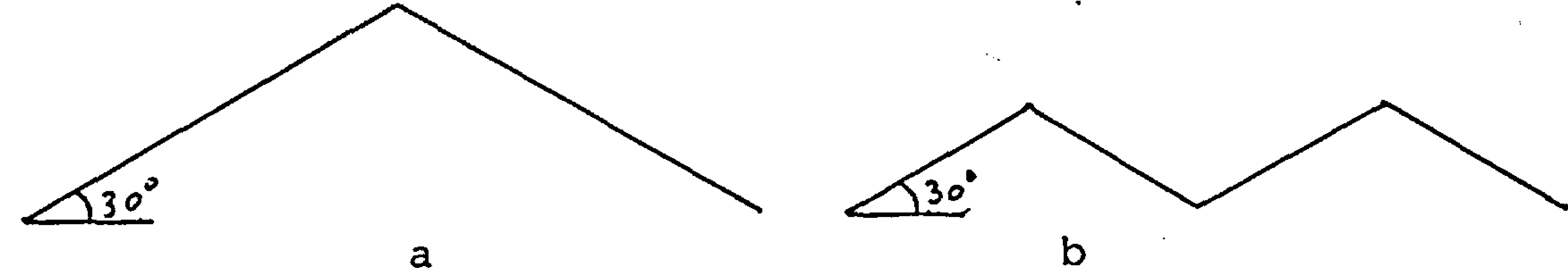


Fig. 7.2. Two surfaces with the same slope distribution but differing in R_a and R_q.

which two have a surface slope of 30° and the other two of -30°. The angular light distribution for both surfaces will be the same, since the areas orientated in a certain direction are equal for both cases. Applying the formulae for evaluating the R_a and R_q parameters leads to the following results:

$$\text{surface a: } R_a=1/4 \quad , R_q= 3/6$$
$$\text{surface b: } R_a=1/8 \quad , R_q= 3/12$$

Another disadvantage of geometrical optics is that its application is only valid when the sizes of the facets are much larger than the wavelength of light. This condition is generally not met by real surfaces.

The numerical approach has to be abandoned because it lacks a solution in closed form. Indeed, the scattered light intensities can be predicted very accurately, but has to be done for individual cases, for which the geometry is exactly known. A statistical description of the surface profile will not suffice, and one can expect that the inverse problem is difficult to solve for general

cases. However, it must be mentioned that Mata Mendez et al. [1983] have published results of a numerical method for retrieving a surface profile from a diffraction pattern. But their method has been applied only to computer generated diffraction patterns caused by simple surface geometries of short lengths with only a few undulations. Results from actual diffraction patterns have not yet been presented. Furthermore, it is an iterative method which is rather time consuming.

This leaves only two candidates as a basis for a surface finish sensor based on light scattering techniques: (a) scalar diffraction and (b) vector diffraction theory. Although the vector theory is more complete than the scalar diffraction theory, the latter is, under certain conditions, not inferior, as can be concluded from various publications, whose contents will be discussed.

So now the question "Under what conditions is the Kirchhoff scalar Approach, KA, valid and are these conditions compatible with the characteristics of diamond turned surfaces?" has to be answered. There exists no unanimously quantitative response to this question.

Sommerfeld [1967] writes that the KA is applicable "to large grating constants $D \gg \lambda$, to grating elements which are not too deeply cut, and to moderate angles of incidence and diffraction."

Beckmann and Spizzichino [1963] consider the KA to be valid for $\min(2kr_c\cos\theta) \gg 1$, with r_c the radius of curvature at a point P on the surface and θ the local incident angle of a plane wavefront at P. Then they write that this condition implies that the KA "will break down completely if the roughness includes sharp edges or sharp points," but they thereupon add, in connection with a comparison of

the KA with rigorous results obtained for a grating with rectangular grooves (four corners per period): "This comparison shows that the Kirchhoff method may still be a good approximation when edges are present and even when the basic condition $D \gg \lambda$ is not adequately satisfied." Lynch's studies [1970] suggest that the KA may be a good approximation when $\min(2kr_c\cos^3\theta) \gg 1$.

It will be appreciated that most of the aforementioned conditions for the validity of the KA involve "much larger" or "much smaller" symbols. Whereas this may be satisfying for the mathematician, it is quite annoying for solving the problem of deciding whether it is justifiable to use the KA. The studies which will be discussed now have produced more quantitative conditions for the validity of the KA.

Stover [1975] and Church and Zavada [1975] studied the scattering from micro-rough sinusoidal gratings (R_q=20nm). These surfaces produce only two diffraction peaks at angles given by the grating equation (7.5.) for m=1. For these surfaces the ratio of the power diffracted in the plus one (P_1) and zero (P_0) order direction are:

$$P_{sca} = P_1/P_0 = (k/h)^2\left[(\cos\theta_s + \cos\theta_i)^2/4\cos\theta_i\right]^2 \quad (7.6a)$$

for scalar diffraction, and for vector diffraction in the case of s- and p- polarisation, respectively:

$$P_{v,s} = P_1/P_0 = (k/h)^2 \cos\theta_i \cos\theta_s \quad (7.6b)$$

$$P_{V,p} = P_i/P_0 = (k/h)^2 [(\cos\theta_i - \lambda \tan\theta_i/d)^2] \cos\theta_i \sec\theta_s \quad (7.6c)$$

where $k=2\pi/\lambda$ and h is the amplitude of a sinusoidal grating with a period length D. For normal incidence ($\theta_i=0$), these formulae reduce to:

$$P_{sca} = (k/h)^2 [(1+\cos\theta_s)^2/4\cos\theta_s]^2 \quad (7.7a)$$

$$P_{V,s} = (k/h)^2 \cos\theta_s \quad (7.7b)$$

$$P_{V,p} = (k/h)^2/\cos\theta_s \quad (7.7c)$$

These formulae differ only in the factor depending on the angular scattering angle θ_s. This angle θ_s can be calculated by the grating equation, which is in this case $\sin\theta_s = \lambda/D$. Figure 7.3. shows the

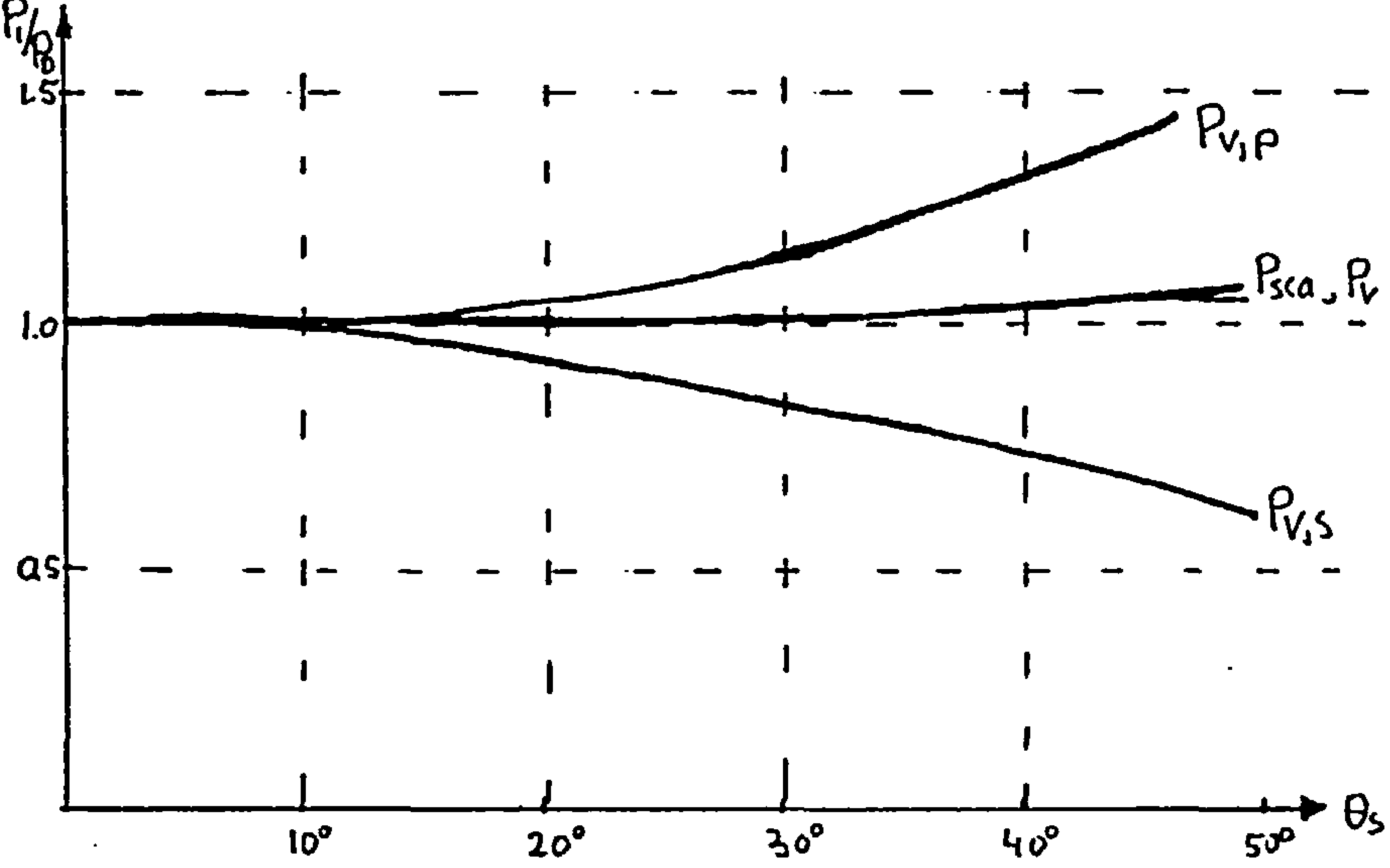


Fig. 7.3. Angular dependence of scalar and vector scattering.

angular dependence of the three formalae (7.7), together with the

arithmetical mean value of the two vector expressions P_V. From this figure it can be concluded that:

i the result for the scalar theory does not differ from the result obtained by the vector theory for a scattering angle up to 10°.

ii the result for the scalar theory and the result for the average value from the vector theory are very close up to a scattering angle of 45°.

This last point suggests that the scalar theory can be applied to rather large scattering angles when the light source is randomly polarised. The angles of 10° and 45° correspond to surface period lengths of resp. 7λ and 1.4λ.

Experimental data produced by Hensler [1972], using TIS measurements, substantiate the validity of the scalar theory under the conditions mentioned previously. Here the surface roughnesses of randomly rough surfaces were calculated from scattering data by applying Beckmann's formula, which is based on the Kirchhoff scalar diffraction theory. These results were compared with data obtained by a stylus instrument. These values compare very well for roughness values up to $\lambda/8$. The correlation lengths of these surfaces were in the order of 3.5λ.

A very comprehensive study on the validity range of the KA has been conducted by Wirgin [1983]. His study is based upon the scattering from sinusoidal gratings. He compares the solutions obtained by the KA with those from rigorous theories. For normal

incident illumination, he concludes that the KA is valid under the conditions that $h/D<0.011(D/\lambda)$ and $h/D<0.13$. The latter is the limit where multiple scattering effects become non-negligible.

Figure 7.4 shows the validity range as obtained from the studies by Wirgin and Stover. The relative scales are the same as in figure 5.6. Comparison of these two figures show that the KA covers quite a few machining processes. From this it can be concluded that the scalar diffraction theory is a valid base for the development of a surface finish sensor.

Vorburger et al. [1986] have checked the validity of the Kirchhoff approach by comparing experimental diffraction patterns from sinusoidal gratings with those predicted by the scalar theory. Here the surface finish range is $1.3\lambda<R_q<4\lambda$, and the period length of the gratings ranged from 50λ to 1000λ. These results are also entered in figure 7.4.

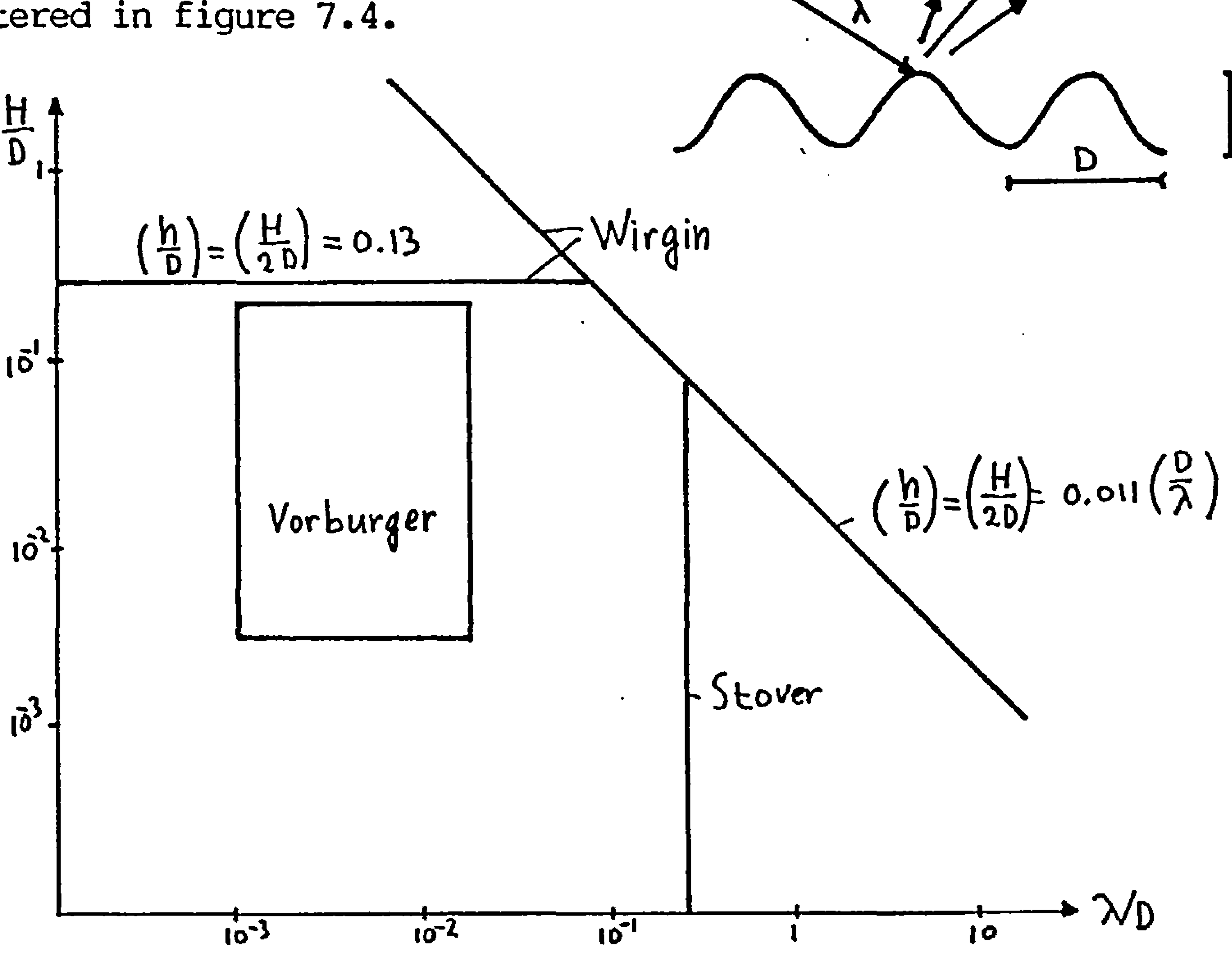


Fig. 7.4. Validity range for the scalar diffraction theory.

8. The Kirhhoff diffraction integral.

Since the operation of the surface finish sensor is based upon the Kirchhoff scalar diffraction theory, it seems appropriate to show the derivation of the Kirchhhoff diffraction integral.

Although diffraction can be applied as an advanced, modern surface metrology tool, it should be remembered that diffraction is an old phenomenon as can be concluded from the following brief account of the history of diffraction.

8.1. History of diffraction.

What is now called Fresnel diffraction, has been studied in detail by Grimaldi (1618-1663). He noticed the deviation of light from rectilinear propagation, something he called "diffractio". This effect is a general characteristic of wave phenomena, occurring whenever a portion of a wavefront, be it sound or light, is obstructed in some way. Grimaldi had observed bands of light within the shadow of a rod illuminated by a small source. Hooke (1635-1703), later also observed diffraction effects. He was the first to study coloured interference patterns, which brought him to propose a wave theory of light.

The wave theory of light was not favoured very much by Newton (1642-1727). Perhaps his main reason for rejecting this wave theory as it stood then was the blatant problem of explaining rectilinear propagation in terms of waves which spread out in all directions, and he became more committed to the emission (corpuscular) theory.

At the same time that Newton was emphasising the emission theory in England, Huygens (1629-1695), on the continent was greatly extending the wave theory.

However, the great weight of Newton's opinion stopped the further evolution of the wave theory. Young (1773-1829) revived the wave theory, but did not meet much appreciation because of the foregoing.

Unaware of Young's efforts, Fresnel (1788-1827) began the revival of the wave theory in France. Fresnel synthesised the concepts of Huygens' wave description and the interference principle. The corresponding Huygens-Fresnel principle states that every unobstructed point of a wavefront, at a given instant in time, serves as a source of spherical secondary wavelets. The amplitude of the optical field at any point beyond is the superposition of all these wavelets (considering their amplitudes and relative phases.)

Under the criticism from Laplace (1749-1827) and Biot (1774-1862), who advocated the emission theory, Fresnel's theory took on a more mathematical emphasis. He was able to calculate the diffraction patterns from various obstacles and accounted for the rectilinear propagation in homogeneous isotropic media, thus dispelling Newton's main objection to the wave theory.

Kirchhoff (1824-1887) developed a more rigorous theory based directly on the solution of the differential wave equation. His refined analysis lent credence to the assumptions of Fresnel and led to an even more precise formulation of Huygens' principle as an exact consequence of the wave equation. Whilst the integral theorem of Kirchhoff embodies the basic idea of the Huygens-Fresnel principle,

the laws governing the contributions from different elements of the obstacle are more complicated than Fresnel assumed. Kirchhoff showed, however, that in many cases the theory may be reduced to an approximate but much simpler form, which is essentially equivalent to the formulation of Fresnel, but which in addition gives an explicit formula for the inclination factor that remained undetermined in Fresnel's theory. Even so, the Kirchhoff theory itself is an approximation which is valid for sufficient small wavelength, i.e. when the diffracting objects have dimensions which are large in comparison to the illuminating wavelength. Kirchhoff's theory, however, works fairly well even though it only deals with scalar waves and is insensitive to the fact that light is a transverse vector field.

8.2. Derivation of the Kirchhoff scalar diffraction formula.

Light is a form of electromagnetic energy covering a small portion of the electromagnetic spectrum. The visible light region corresponds to frequencies ranging from $0.3 \cdot 10^{15}$ to 10^{15}Hz. Therefore the fundamental properties of light can be described by Maxwell's equations, which relate the time varying vector fields $\vec{E}$ and $\vec{B}$, where $\vec{E}$ is the electric field strength and $\vec{B}$ the magnetic induction. A summary of the used vector expressions and vector operations can be found in appendix A1. In free space the Maxwell equations in integral form are [Hecht and Zajac 1974]:

$$\oint_C \vec{E}\cdot d\vec{l} = -\iint_A \frac{\partial \vec{B}}{\partial t}\cdot d\vec{S} \quad (8.1a)$$

$$\oint_C \vec{B}\cdot d\vec{l} = \mu_0 \epsilon_0 \iint \frac{\partial \vec{E}}{\partial t}\cdot d\vec{S} \quad (8.1b)$$

$$\iint_A \vec{B}\cdot d\vec{S} = 0 \quad (8.1c)$$

$$\iint_A \vec{E}\cdot d\vec{S} = 0 \quad (8.1d)$$

where in (8.1a) and (8.1b) A is an open surface bounded by the closed contour C and in (8.1c) and (8.1d) A is a closed surface covering the volume V. In free space the magnetic permeability μ equals μ_0 and the electric permeability ϵ equals ϵ_0.

These integral equations can be rewritten in differential form which are more useful for deriving the wave aspects of the electromagnetic field. This transition can be readily accomplished by making use of two theorems from vector calculus, namely Gauss's divergence theorem:

$$\iint_A \vec{F}\cdot d\vec{S} = \iiint_V \nabla\cdot F \, dV \quad (8.2)$$

and Stokes' theorem:

$$\oint_C \vec{F}\cdot d\vec{l} = \iint_A \nabla\times \vec{F}\cdot d\vec{S} \quad (8.3)$$

where $\vec{F}$ is a vector quantity.

Applying Stokes' theorem to the electric field strength yields:

$$\oint_C \vec{E} \cdot d\vec{l} = \iint_A \nabla \times \vec{E} \cdot d\vec{S} \tag{8.4}$$

Upon comparing this with equation (8.1a) it follows that:

$$\iint_A \nabla \times \vec{E} \cdot d\vec{S} = -\iint_A \frac{\partial \vec{B}}{\partial t} \cdot d\vec{S} \tag{8.5}$$

This result must be true for all surfaces bounded by the contour C. This can then only be the case if the integrands are themselves equal, i.e. if:

$$\nabla \times \vec{E} = -\frac{\partial B}{\partial t} \tag{8.6a}$$

A similar application of Stokes' theorem to $\vec{B}$, using equation (8.1b) results in:

$$\nabla \times \vec{B} = \mu_0 \varepsilon_0 \frac{\partial \vec{E}}{\partial t} \tag{8.6b}$$

Gauss's divergence theorem applied to the electric field strength $\vec{E}$ yields:

$$\oiint_A \vec{E} \cdot d\vec{S} = \iiint_V \nabla \cdot \vec{E} \, dV = 0 \tag{8.7}$$

and since this is to be true for any volume (i.e. for an arbitrary closed domain) the integrand must equal zero. Consequently, at any point (x,y,z,t) in space-time:

$$\nabla \cdot \vec{E} = 0 \tag{8.8a}$$

In the same fashion Gauss's divergence theorem applied to the $\vec{B}$ field and combined with equation (8.1d) yields:

$$\nabla \cdot \vec{B} = 0 \tag{8.8b}$$

The equations (8.6a,b) and (8.8a,b) are Maxwell's differential equations in free space.

Taking the curl of equation (8.6b) and assuming $\vec{E}$ is a well behaved function so that the space and time derivatives can be interchanged, yields:

$$\nabla \times (\nabla \times \vec{B}) = \mu_0 \varepsilon_0 \frac{\partial}{\partial t}(\nabla \times \vec{E}) = -\mu_0 \varepsilon_0 \frac{\partial^2}{\partial t^2} \vec{B} \tag{8.9}$$

The vector triple product can be simplified by using the operator identity:

$$\nabla \times (\nabla \times \vec{F}) = \nabla (\nabla \cdot \vec{F}) - \nabla^2 \vec{F} \tag{8.10}$$

where ∇^2 is called the Laplacian. It operates on each component of the vector field $\vec{F}$:

$$\nabla^2 \vec{F} = \nabla^2 F_x \vec{e}_x + \nabla^2 F_y \vec{e}_y + \nabla^2 F_z \vec{e}_z \tag{8.11}$$

where F_x, F_y and F_z are the vectorial components of $\vec{F}$ in the directions $\vec{e}_x, \vec{e}_y$ and $\vec{e}_z$ respectively, using Cartesian coordinates, and:

$$\nabla^2 F_n = \frac{\partial^2}{\partial x^2} F_n + \frac{\partial^2}{\partial y^2} F_n + \frac{\partial^2}{\partial z^2} F_n \tag{8.12}$$

where n can be x,y or z. Application of (8.10) on equation (8.9) leads to the result:

$$\nabla^2 \vec{B} - \mu_0 \epsilon_0 \frac{\partial^2}{\partial t^2} \vec{B} = \nabla^2 \vec{B} - \frac{1}{c^2} \frac{\partial^2}{\partial t^2} \vec{B} \tag{8.13a}$$

with $c=1/\sqrt{\mu_0 \epsilon_0}$, the speed of light in free space. A similar expression can be derived for the electric field strength $\vec{E}$.

$$\nabla^2 \vec{E} - \frac{1}{c^2} \frac{\partial^2}{\partial t^2} \vec{E} = 0 \tag{8.13b}$$

These two vector equations actually represent a total of six equations. One of these equations is:

$$\frac{\partial^2}{\partial x^2} E_x + \frac{\partial^2}{\partial y^2} E_x + \frac{\partial^2}{\partial z^2} E_x = \frac{1}{c^2} \frac{\partial^2}{\partial t^2} E_x \tag{8.14}$$

with precisely the same form for E_y, E_z, B_x, B_y and B_z. Equations of this kind, which relate to time and space variations of some physical quantity, describe wave phenomena. Each and every component of the electromagnetic field therefore obeys the scalar wave equation:

$$\nabla^2 \psi = \frac{1}{c^2} \frac{\partial^2}{\partial t^2} \psi \tag{8.15}$$

Without specifying the precise spatial nature of the wave, ψ can be written as:

$$\psi = U e^{-ikct} \tag{8.16}$$

where $k=2\pi/\lambda$, with λ the wavelength of the monochromatic wave ψ. Here U represents the complex space part of the scalar ψ. Substitution into the wave equation yields:

$$\nabla^2 U + k^2 U = 0 \tag{8.17}$$

This expression is known as the Helmholtz equation. In the remainder of the text, the time dependent part expression e^{-ikct} will be suppressed unless stated otherwise. Equation (8.17) can be solved with the aid of Green's theorem as follows. Suppose that there are two scalar functions U_1 and U_2, then according to Green's theorem

$$\iiint_V (U_1 \nabla^2 U_2 - U_2 \nabla^2 U_1)\, dV = \oint\!\!\oint_A (U_1 \nabla U_2 - U_2 \nabla U_1) \cdot d\vec{S} \tag{8.18}$$

It is obvious that if U_1 and U_2 are solutions of the Helmholtz equation, i.e. if:

$$\nabla^2 U_1 + k^2 U_1 = 0 \tag{8.19a}$$

and

$$\nabla^2 U_2 + k^2 U_2 = 0 \tag{8.19b}$$

then

$$\oint\!\!\oint_{A_2} (U_1 \nabla U_2 - U_2 \nabla U_1) \cdot d\vec{S} = 0 \tag{8.20}$$

Let $U_1=U$, the space portion of an unspecified scalar optical field and let

$$U_2 = \frac{e^{ikr}}{r} \tag{8.21}$$

where r is measured from a point P. Both of these choices U_1 and U_2 satisfy the Helmholtz equation. There is a singularity for U_2 at point P, where r=0. By surrounding P by a small sphere A' in order to exclude P from the region enclosed by the surface A, see figure 8.1,

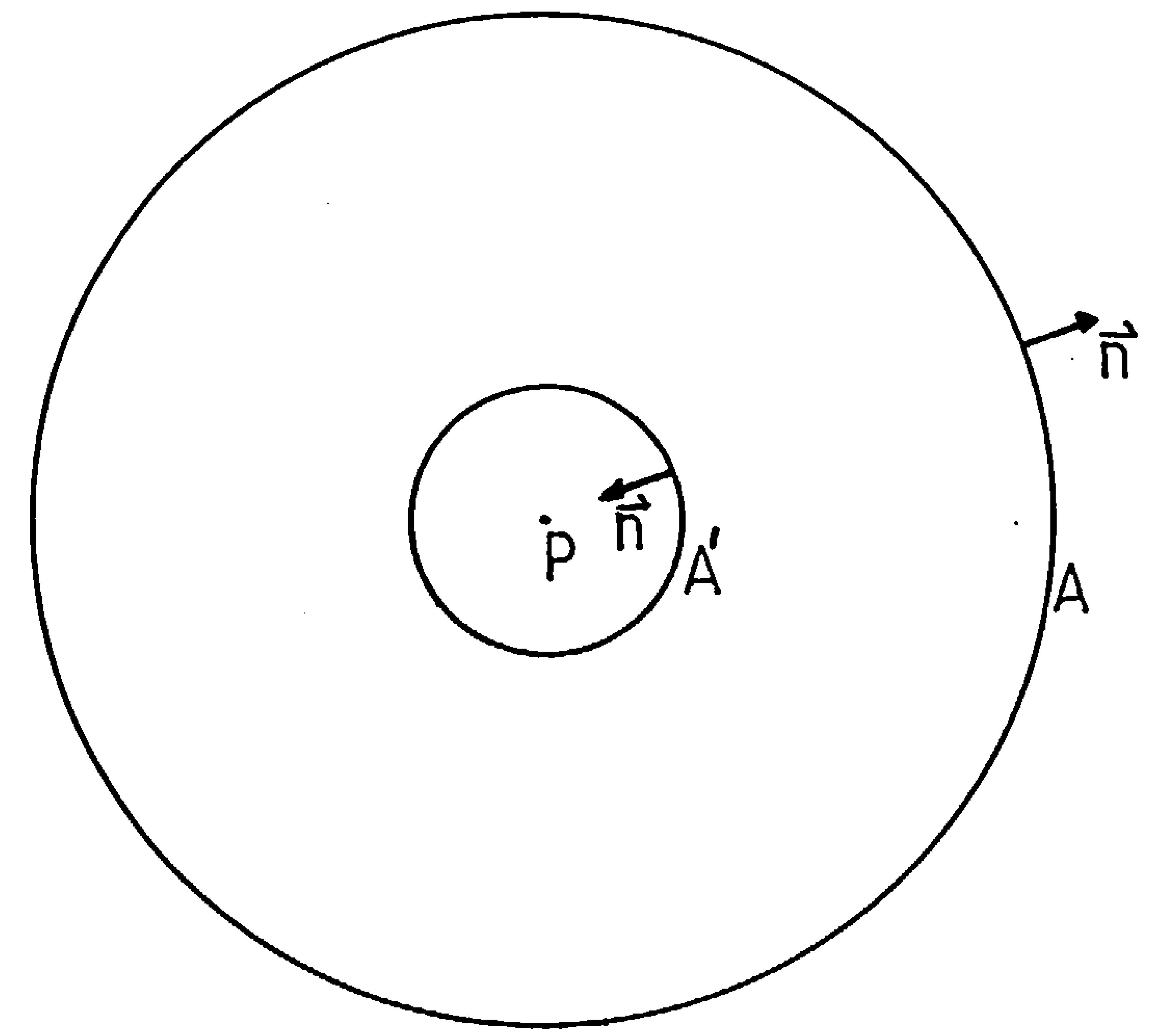


Figure 8.1. Geometry involved for the evaluation of integral 8.22.

equation (8.20) now becomes:

$$\oint\!\!\!\oint_A \left[U \nabla\left(\frac{e^{ikr}}{r}\right) - \frac{e^{ikr}}{r} \nabla U \right] \cdot d\vec{S} + \oint\!\!\!\oint_{A'} \left[U \nabla\left(\frac{e^{ikr}}{r}\right) - \frac{e^{ikr}}{r} \nabla U \right] d\vec{S} = 0 \quad (8.22)$$

On the small sphere, the unit normal $\vec{n}$ points towards the origin at P and

$$\nabla\left(\frac{e^{ikr}}{r}\right) = \left(\frac{1}{r^2} - \frac{ik}{r}\right) e^{ikr} \vec{n} \quad (8.23)$$

since the gradient is directed radially outward. In terms of the solid angle ($dS=r^2 d\Omega$) measured at P, the integral over A' becomes:

$$\oiint_{A'} \left(U - ikrU + r\frac{\partial U}{\partial r} \right) e^{ikr}\, d\Omega \tag{8.24}$$

where $\nabla U \cdot d\vec{S} = -(dU/dr)r^2 d\Omega$. As the sphere surrounding P shrinks, $r \to 0$ on A' and $e^{ikr} \to 1$. Because of the continuity of U, its value at any point on A' approaches its value at P, i.e. U(P). The last two terms in equation (8.24) go to zero and the integral becomes $4\pi U(P)$. Finally equation (8.22) becomes:

$$U(P) = \frac{1}{4\pi} \oiint_{A} \left[\frac{e^{ikr}}{r} \nabla U - U \nabla \left(\frac{e^{ikr}}{r} \right) \right] \cdot d\vec{S} \tag{8.25}$$

Which is known by names of the Kirchhoff integral theorem and the Helmholtz integral. This integral expresses the optical field U(P) at point P in terms of the field and gradiant evaluated on an arbitrary closed surface A, enclosing P.

The Helmholtz integral is easily applied to surfaces that are not closed; e.g. in the application of calculating the field scattered from a rough surface. Here the rough surface part A can be regarded as part of any closed surface A', of which only A has a non-vanishing reflection coefficient; the remaining integral over A-A' vanishes so that there is no need to introduce the surface A'. Now the integral (8.25) can be further developed as follows. Let E be a component of the electric field strength due to a point source at some point Q, so that

$$E = E_0\, e^{i(k\rho - \omega t)} / \rho \tag{8.26}$$

in which case $U=E_0 e^{ik\rho}/\rho$. Substitution into equation (8.25) leads to

$$U(P) = \frac{1}{4\pi} \iint_A \left[\frac{e^{ikr}}{r} \frac{\partial}{\partial \rho}\left(\frac{E_0}{\rho} e^{ik\rho}\right) \cos(\vec{n},\vec{\rho}) - \frac{E_0}{\rho} e^{ik\rho} \frac{\partial}{\partial r}\left(\frac{e^{ikr}}{r}\right) \cos(\vec{n},\vec{r}) \right] dS \tag{8.27}$$

where $d\vec{S}=\vec{n}dS$, $\vec{n}$, $\vec{r}$ and $\vec{\rho}$ are unit vectors.

$$\nabla \left(\frac{e^{ikr}}{r}\right) = \vec{r} \frac{\partial}{\partial r}\left(\frac{e^{ikr}}{r}\right) \tag{8.28a}$$

and

$$\nabla U = \vec{\rho} \frac{\partial U}{\partial \rho} \tag{8.28b}$$

The differentiations under the integral sign are

$$\frac{\partial}{\partial \rho}\left(\frac{e^{ik\rho}}{\rho}\right) = e^{ik\rho}\left(\frac{ik}{\rho} - \frac{1}{\rho^2}\right) \tag{8.29a}$$

and

$$\frac{\partial}{\partial r}\left(\frac{e^{ikr}}{r}\right) = e^{ikr}\left(\frac{ik}{r} - \frac{1}{r^2}\right) \tag{8.29b}$$

when $\rho >> r$ and $r >> \lambda$ (so called Kirchhoff conditions) the $1/\rho^2$ and $1/r^2$ terms can be neglected. Proceeding, one obtains:

$$U(P) = \frac{E_0 ik}{2\pi} \iint_A \frac{e^{ik(r+\rho)}}{r\rho} \left[\cos(\vec{n},\vec{\rho}) - \cos(\vec{n},\vec{r})\right] dS \tag{8.30}$$

This formula applies to the configuration in figure 8.2a. However,

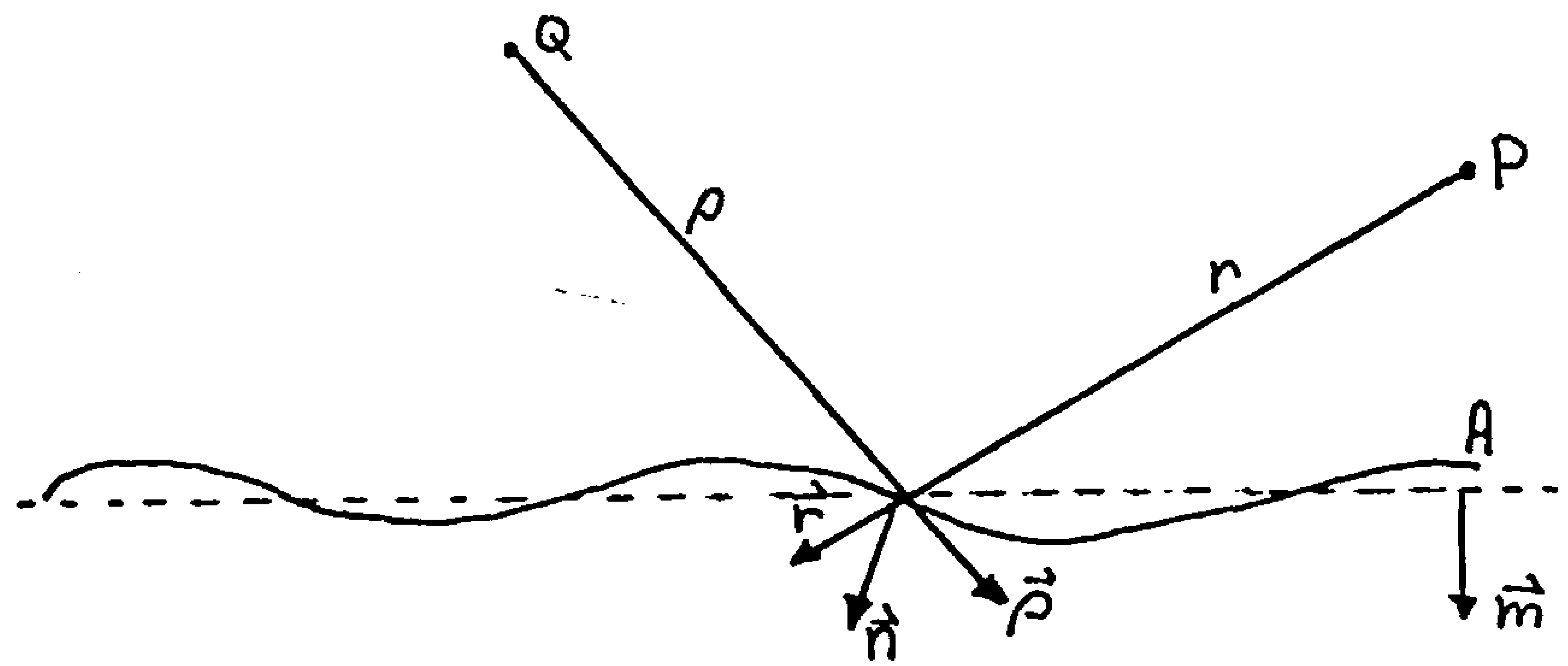


Figure 8.2a Geometry for solving integral (8.30).

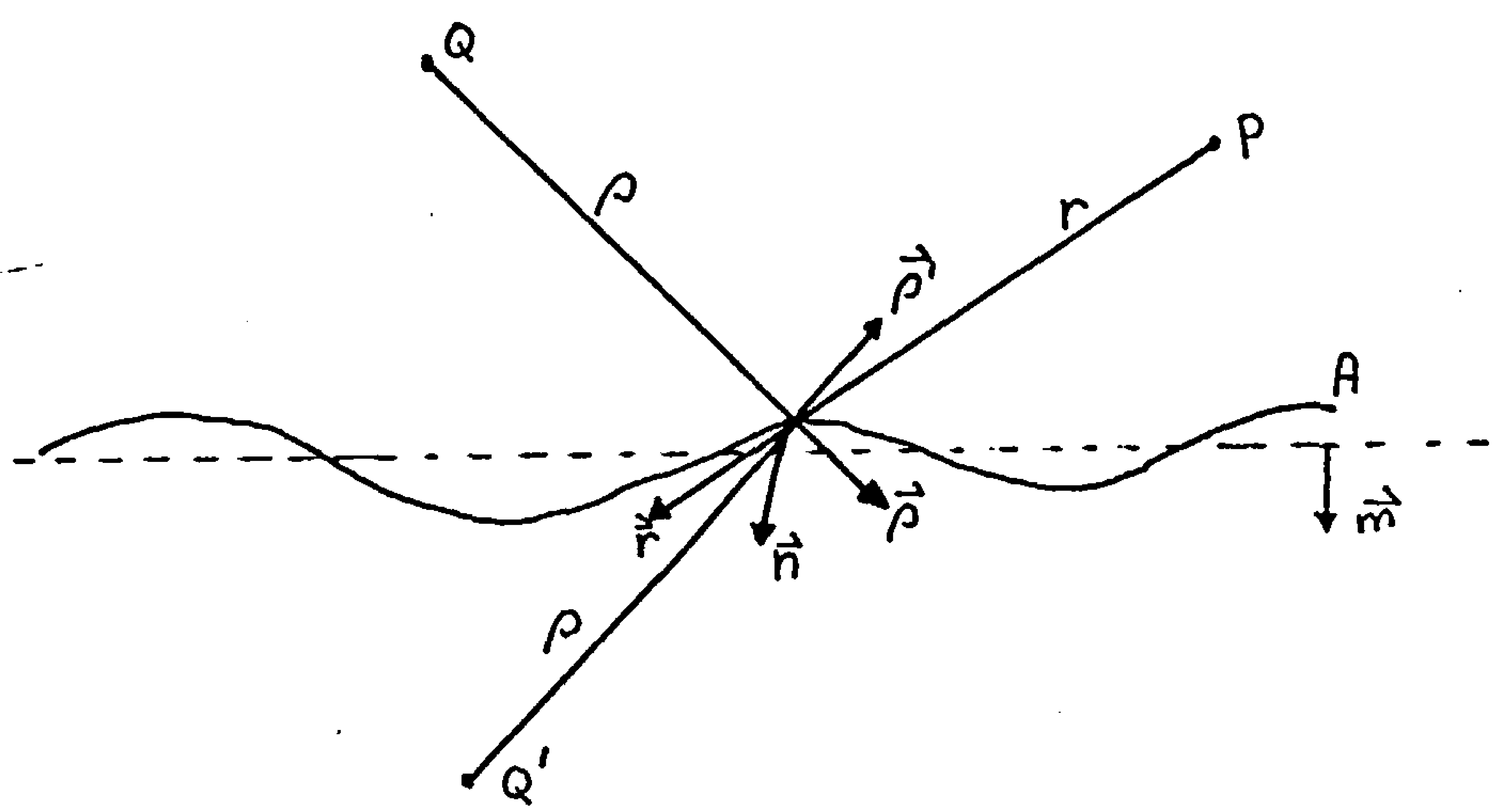


Figure 8.2b Geometry for solving integral (8.31).

when the surface is a perfect conductor, then the field at the surface A is a summation of the original field and the field produced by the eddy currents in the surface (see appendix A2). This induced

field can be represented by a source present at the mirror image position of the original optical source. Here the surface A acts as the mirror and the strength of the image is proportional to the original source. This results in the situation as sketched in figure 8.2b., and the expression (8.30) has to be rewritten as:

$$U(P) = \frac{E_0 ik}{2\pi} \oint\!\!\int_A \frac{e^{ik(r+\rho)}}{r\rho} \left[\cos(\vec{n},\vec{\rho}) - \cos(\vec{n},\vec{r})\right] dS$$
$$+ \frac{C_i E_0 ik}{2\pi} \oint\!\!\int_A \frac{e^{ik(r+\rho)}}{r\rho} \left[\cos(\vec{n},\vec{\rho}\,') - \cos(\vec{n},\vec{r})\right] dS \qquad (8.31)$$

Under the so-called Kirchhoff assumptions that:

- i the actual surface normal $\vec{n}$ does not differ much from the normal $\vec{m}$ of the mean surface,
- ii the light is under normal incidence,
- iii the normal distance between the observation point P and the surface A is much larger than the dimensions of A, and
- iv the lateral position of P is in the order of magnitude of the dimensions of A,

then the cosine terms in (8.31) take on the values 0 and -2 resp., so that (8.31) becomes:

$$U(P) = C E_0 \oint\!\!\int_A \frac{e^{ik(r+\rho)}}{r\rho} dS \qquad (8.32)$$

It has to be remembered that the following assumptions have been made for the derivation of the above formula:

i the surface is a perfect conductor.

ii the surface slopes are very small.

iii near normal incident illumination.

iv near normal incident observation.

v no multiple scattering occurs.

vi no shadowing by the surface occurs.

This integral forms the basis of the development of the surface sensor. In chapter 9, this integral is further developed in order to enter expressions for various surface profiles into it.

9. Application of the Kirchhoff diffraction integral to the inverse scattering problem for periodic surfaces.

The Kirchhoff diffraction integral (8.32) is evaluated for several shapes of periodic surface profiles. From these results it appears that their scattered light distributions are almost identical for a surface roughness range $0<R_q/\lambda<0.10$. This property is employed to obtain certain surface finish parameters from the scattered light distribution (the inverse scattering problem). A so-called phase modulation model is derived to enable the insertion of the various surface profiles into integral (8.32). With the aid of this model, the integrand of the diffraction integral can be decomposed in two parts, one part containing the surface profile and the other one a Fourier transform expression.

9.1. Derivation of the phase modulation model.

For simplicity, the scattering problem can be made two-dimensional (i.e. all functions depend only on x and y). This simplification is justified for turned surfaces. This geometry is described by the following, as depicted in figure 9.1a.

i The surface S is fully described by $y=y(x)$, and is independent of z.

ii The source point Q of the illumination is located at infinity in $Q(0,Y_1)$, where $Y_1 \rightarrow \infty$. This ensures normal

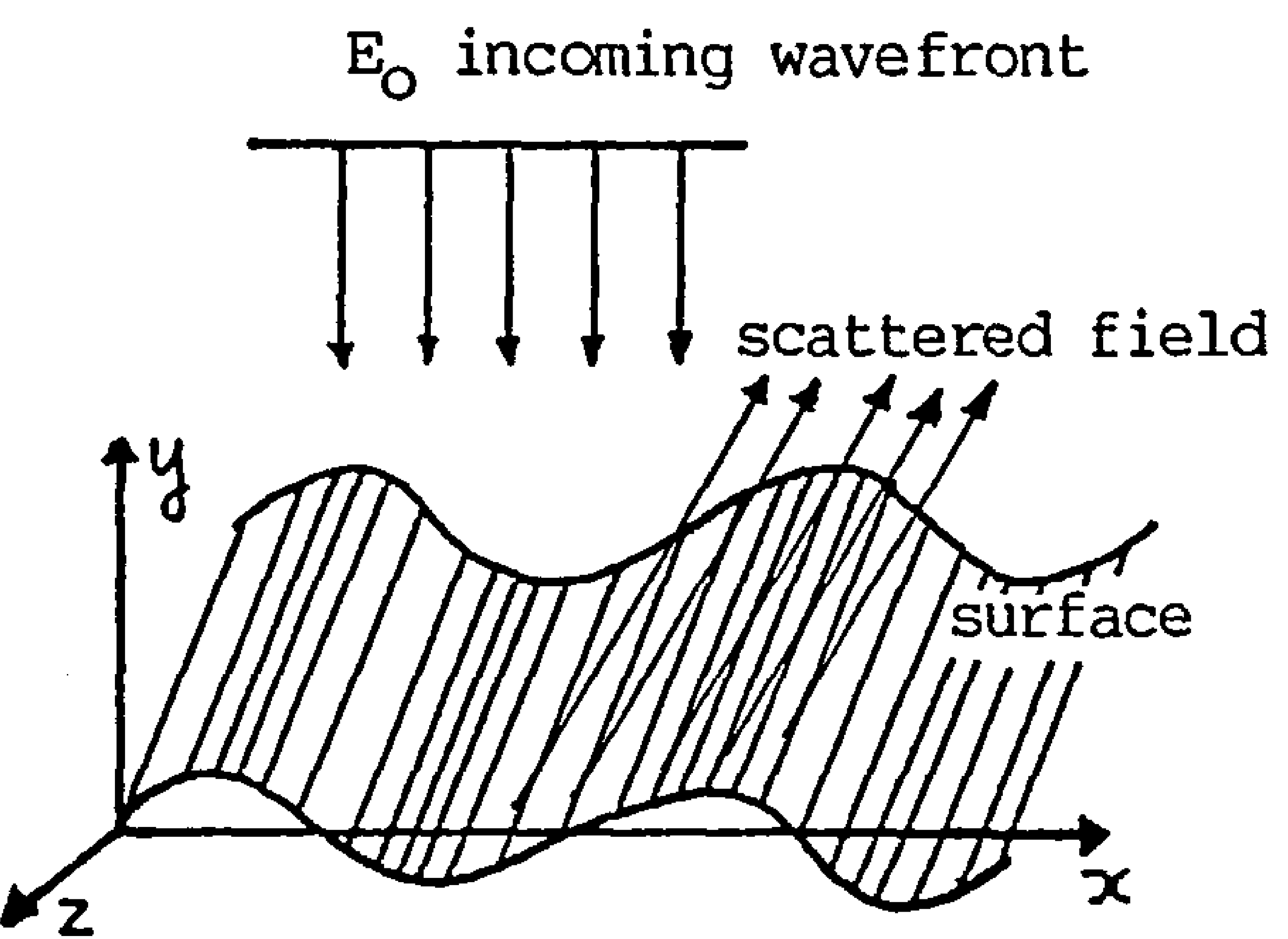


Figure 9.1a Three-dimensional scattering geometry.

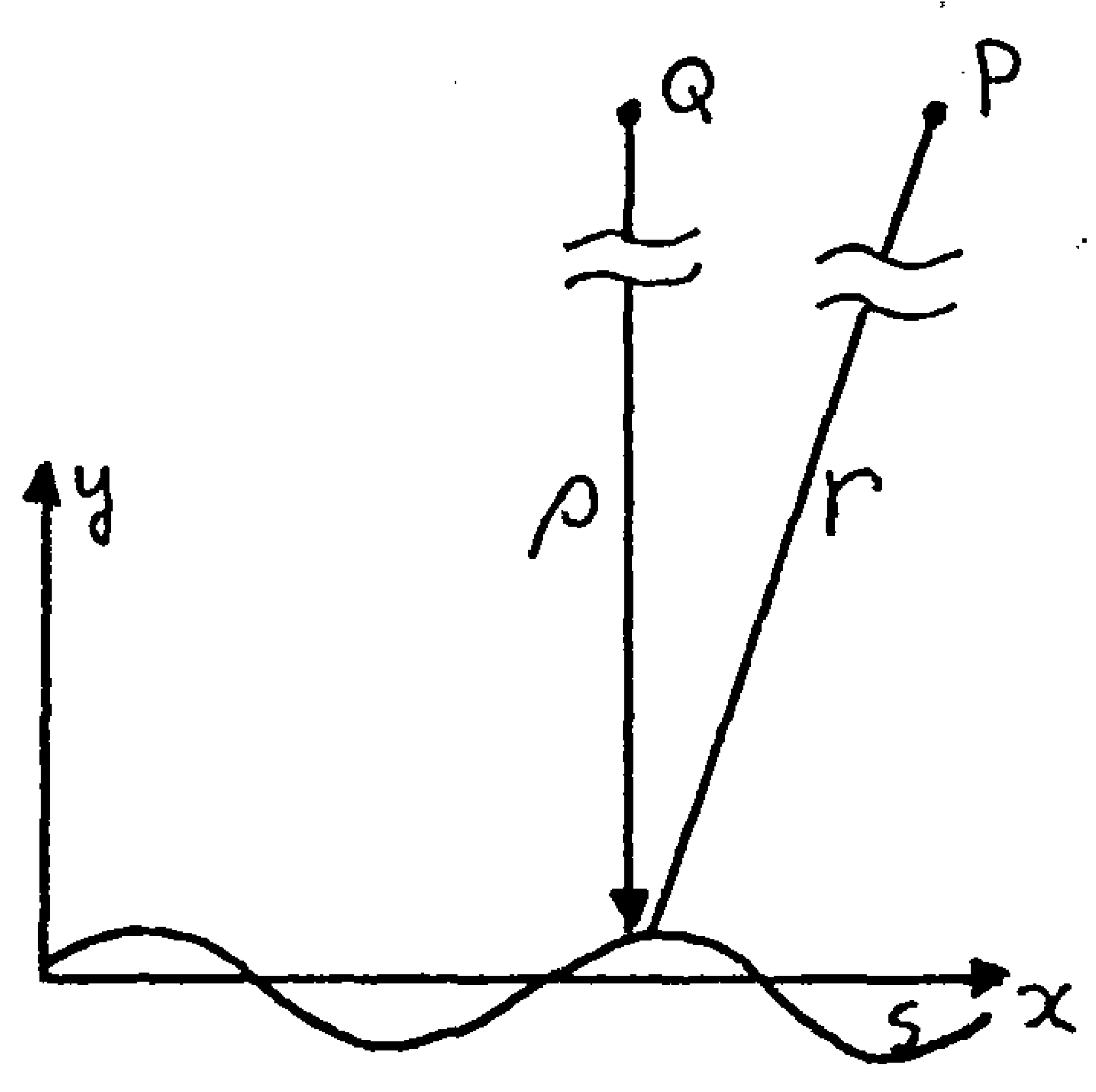


Figure 9.1b Two-dimensional scattering geometry.

incident illumination with a plane wavefront.

The three-dimensional geometry from figure 9.1a can then be reduced to the two-dimensional scattering model shown in figure 9.1b. The integral (8.32) can then be rewritten as:

$$U(X) \propto E_0 \int_S \frac{e^{ik(r+\rho)}}{r\rho} \, dx \tag{9.1}$$

Assuming that the source point Q is at $(0,Y_1)$ and that the observation point P is at (X,Y_2), where $Y_2 \to \infty$ and $X \ll Y_2$, then the distances r and ρ from a point T on the surface y=y(x) to the points P and Q can be expressed by:

$$r = \left[(X-x)^2 + (Y_2 - y(x))^2 \right]^{1/2} \tag{9.2a}$$

and

$$\rho = \left[x^2 + (Y_1 - y(x))^2 \right]^{1/2} \tag{9.2b}$$

or

$$r = Y_2\left[1 - \frac{2(Xx + Y_2 y(x))}{Y_2^2} + \frac{X^2 + x^2 + y^2(x)}{Y_2^2}\right]^{1/2} \tag{9.3a}$$

and

$$\rho = Y_1\left[1 - \frac{2Y_1 y(x)}{Y_1^2} + \frac{x^2 + y^2(x)}{Y_1^2}\right]^{1/2} \tag{9.3b}$$

thus

$$r = Y_2 - \frac{Xx + Y_2 y(x)}{Y_2} + \frac{X^2 + x^2 + y^2(x)}{2Y_2} + \cdots \tag{9.4a}$$

and

$$\rho = Y_1 - \frac{Y_1 y(x)}{Y_1} + \frac{x^2 + y^2(x)}{2Y_1} + \cdots \tag{9.4b}$$

on expanding (9.3a,b) as far as the first term by the binomial theorem. Small changes in the exponential in (9.1) are important since it can cause the integrand to go through a complete cycle. Thus small changes in r and ρ are important in the exponentional. On the other hand it is assumed that Y_1 and Y_2 tend to infinity, so small changes in the $1/r\rho$ factor can be ignored, and can be replaced by a constant. Since only small angles of diffraction are considered $X \ll Y_2$, $x \ll Y_2$ and $y(x) \ll Y_2$, only the first two terms of the expressions (9.4a,b) need considering. Equation (9.1) becomes:

$$U(X) \propto \int_S e^{ik\left(Y_1 + Y_2 - \frac{Xx}{Y_2} - 2y(x)\right)} dx \tag{9.5}$$

ignoring the constant phase term $e^{ik(Y_1+Y_2)}$ leads to:

$$U(X) \propto \int_S e^{-i2ky(x)} \cdot e^{ik\frac{Xx}{Y_2}} dx \tag{9.6}$$

From this expression it can be recognised that the optical wave at X, U(X), is proportional to a scaled Fourier transform of the phase

function p(x), where

$$p(x) = e^{-ik2y(x)} \tag{9.7}$$

In this derivation it is assumed that the observation point X is at infinity. This is not a practical proposition for an instrument which has to contain the observation point X. However, infinity can be optically achieved with the aid of a convex lens. By definition, parallel rays, entering a convex lens with focal length F, will cross in the back-focal plane of this lens. This is exactly the definition of infinity, that is the point where all parallel lines cross. This can be introduced in equation (9.6) by replacing the quotient X/Y_2 with X/F. Here both qoutients represent the scattering angle under which the light leaves the surface. Equation (9.6) now becomes:

$$U(X) = \int_S e^{-i2ky(x)} e^{-ik\frac{Xx}{F}} dx \tag{9.8}$$

Equation (9.6) can also be derived by assuming that on reflection of the plane, normal incident wavefront, the reflected wavefront, very close to the surface, has become phase modulated in a geometrical way. Figure 9.2. illustrates this process for a simple

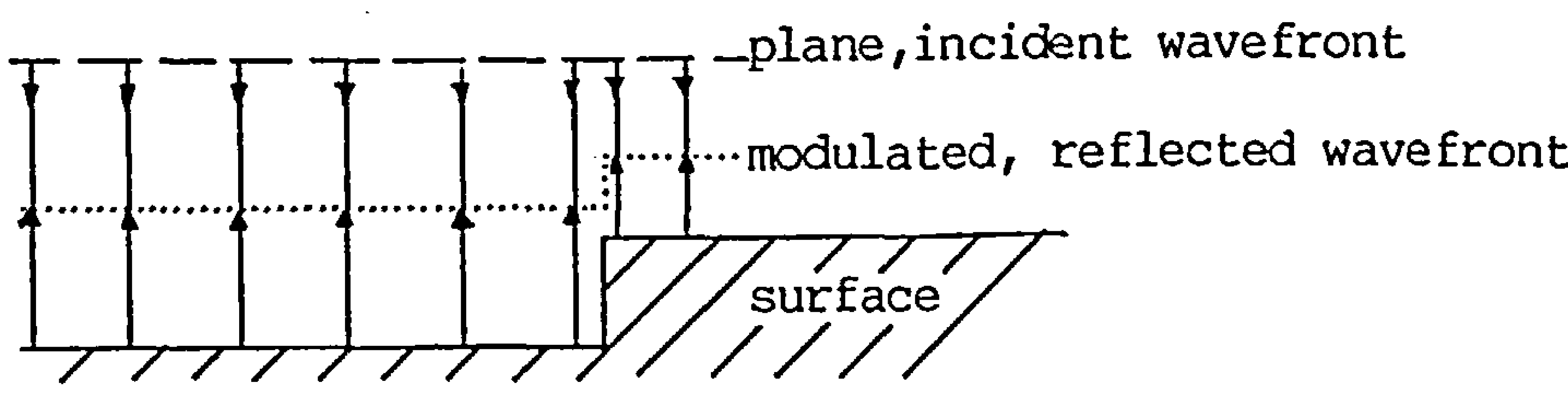


Figure 9.2 Phase modulation model.

surface geometry. In general the phase modulation, or phase advance, on reflection can be expressed by:

$$\varphi(x) = \frac{-2\pi}{\lambda} \cdot 2y(x) \tag{9.9}$$

where y(x) is the surface profile. The phase function $p_1(x)$ in a plane parallel to the mean plane of the surface and close to this surface, can be expressed by:

$$P_1(x) = e^{-i\frac{2\pi}{\lambda}2y(x)} \tag{9.10}$$

and assuming that the amplitude of the incoming wavefront is independent of x, then the reflected field can be expressed by:

$$E(x) = E_0\, e^{-ik2y(x)} \tag{9.11}$$

and the Kirchhoff integral for this situation is:

$$U(X) \propto E_0 \int_S e^{-ik2y(x)}\, e^{ikr} \tag{9.12}$$

where the integration has to be performed over a plane surface $y=y_0$, so

$$r = \left[(X-x)^2 + (y_2-y_0)^2\right]^{1/2} = (y_2-y_0)\left[1 - \frac{2Xx}{(y_2-y_0)^2} + \frac{X^2+x^2}{(y_2-y_0)^2}\right]^{1/2} \tag{9.13}$$

which on expansion as far as the first terms becomes:

$$r = (y_1 - y_0) - \frac{Xx}{(y_1-y_0)} + \frac{X^2 + x^2}{2(y_1-y_0)} + \cdots \tag{9.14}$$

and for the same reasons as before in (9.4a), the term in the exponential can be approximated by:

$$r = (y_1 - y_0) - \frac{Xx}{y_1 - y_0} \tag{9.15}$$

and the integral (9.12) can be rewritten as:

$$U(X) \propto E_o\, e^{ik(y_1-y_0)} \int_S e^{-ik2y(x)}\, e^{\frac{-ikXx}{y_1-y_0}}\, dx \tag{9.16}$$

which with equating y_o to zero becomes:

$$U(X) \propto \int_S e^{-ik2y(x)}\, e^{-\frac{ikXx}{y_1}}\, dx \tag{9.17}$$

So usage of the geometrical reflection model near the reflecting surface, leads to the same result as the more strict application of the Kirchhoff integral (9.1).

9.2 Relation between the light intensity and the optical field.

Because of the extremely high frequencies of visible light waves ($f=3.10^{14}$ Hz for λ=1um), direct observations of the optical (electrical field are normally not possible. The only measurable quantity is the intensity I(X), which is the time average of the amount of energy which, in unit time, crosses a unit area normal to

the direction of the energy flow. This is for optical fields proportional to the time average of the square of the electric field,

$$I(X) = \lim_{T \to \infty} \frac{1}{2T} \int_{-T}^{T} E^2(X,t)\, dt \tag{9.18}$$

where E(X,t) is the real part of $U(X)e^{i\omega t}$ with $e^{i\omega t}$ the time dependent part of the optical field which is normally suppressed. E(x,t) can be expressed by:

$$E(X,t) = \left[U(X)\, e^{i\omega t} + U^*(X)\, e^{-i\omega t} \right] / 2 \tag{9.19}$$

where U(X)* is the complex conjugate of U(X). Insertion into (9.18) yields:

$$I(X) = \lim_{T \to \infty} \frac{1}{2T} \left[\frac{1}{4} \int_{-T}^{T} \left(U^2(X)\, e^{i2\omega t} + U^{*2}(X)\, e^{-i2\omega t} + 2U(X)U^*(X) \right) dt \right] \tag{9.20}$$

The time average of the first two time dependent terms equals zero, and the integral becomes:

$$I(X) = \frac{1}{2}\, U(X)U^*(X) = \frac{1}{2} \left| U(X) \right|^2 \tag{9.21}$$

So the intensity I(X) in the observation point X is proportional to the squared value of the optical field U(X) at X.

9.3. <u>Relationship between the light intensity and the PSDF of surface profiles with small amplitudes.</u>

For general surface profile shapes with small amplitudes and a zero mean, it is possible to show that light intensity I(X) is proportional with the PSDF (see section 4.1.2) of this surface profile y(x). The autocorrelation theorem or Wiener-Khintchine theorem has to be applied to obtain this result. This theorem states that the Fourier transform of the autocorrelation of a function equals the squared of the Fourier transform of the function itself taken over all limits.

$$I(X) \propto |U(X)|^2 = \int_{-\infty}^{\infty} \left[\int_{-\infty}^{\infty} p(x+s)\, p^*(x)\, dx \right] e^{-ikXs/F}\, ds \qquad (9.22)$$

where p(x) is the exponential phase function as defined by (9.7). p(x) can be approximated by $(1+i\varphi(x)-\varphi^2(x)/2)$, under the condition that $|\varphi_{max}| = |2ky_{max}(x)| < 0.5$ or $R_q/\lambda < 0.03$. Insertion in equation (9.22) yields:

$$I(X) = \int_{-\infty}^{\infty} \left[\int_{-\infty}^{\infty} \big(1 - \varphi(x) + \varphi(x+s) - \varphi^2(x)/2 - \varphi^2(x+s)/2 + \varphi(x)\varphi(x+s) + i\varphi^2(x+s)\varphi(x)/2 - i\varphi(x+s)\varphi^2(x)/2 + \varphi^2(x+s)\varphi^2(x)/4 \big) dx \right] e^{-ikXs/F}\, ds \qquad (9.23)$$

The integral over the first order terms equal zero because of the zero mean of the $\varphi(x)$. The third and fourth order terms can be neglected since $|\varphi_{max}| < 0.5$. Furthermore $\int_{-\infty}^{\infty} \varphi^2(x)\, dx = \int_{-\infty}^{\infty} \varphi^2(x+s)\, dx$. This leads to:

$$I(X)=\int_{-\infty}^{\infty}\left[\int_{-\infty}^{\infty}\left(1-\left(\frac{4\pi}{\lambda}\right)^2 y^2(x)+\left(\frac{4\pi}{\lambda}\right)^2 y(x)y(x+s)\right)dx\right]e^{-ikXs/F}\,ds \quad (9.24)$$

With the aid of the definitions for R_q (see page 16) and R(s) (4.1), this expression can be rewritten as:

$$I(X)\propto\int_{-\infty}^{\infty}\left[\left(\frac{\lambda}{4\pi}\right)^2-R_q^2\right]e^{-ikXs/F}\,ds+\int_{-\infty}^{\infty}R(s)e^{-ikXs/F}\,ds \quad (9.25)$$

Here the first integral represents the light intensity reflected into the zero diffraction order or specular direction, and the second integral represents the light intensity scattered outside the zero order. Comparison of this second integral with expression (4.4a) shows that the scattered light distribution is proportional with the power spectral density function (PSDF) of the surface profile y(x).

From this follows that, when $R_q/\lambda < 0.03$, then various statistics of the surface profile can be extracted from zeroth, second and fourth moment of the scattered light distribution in the same way as outlined in section 4.1.3.

9.4. Calculation of the intensity for various diffraction patterns produced by periodic surfaces.

Formula (9.17) is employed to calculate U(X) for various surface profiles y(x). These profiles y(x) are expressed by a finite Fourier series:

$$y(x) = \sum_{n=1}^{7} a_n \cos(2\pi n x/D) \qquad (9.26)$$

where D is the length of one period of the profile. The following table lists the coefficients a_n for the profiles. These profiles approximate the named curves. The maximum rms difference between these approximations and the ideal curves is 0.037%. They resemble very closely surface profiles obtained in diamond turning.

Table 9.1.

coefficient	a_1	a_2	a_3	a_4	a_5	a_6	a_7
cosine	1	0	0	0	0	0	0
triangle	1	0	0.111	0	0.04	0	0.02
|sine|	1	-0.2	0.086	-0.047	0.03	-0.021	0.015
parabola	1	-0.25	0.111	-0.0625	0.04	-0.028	0.02

These profiles are depicted in the figures 9.3a-d.

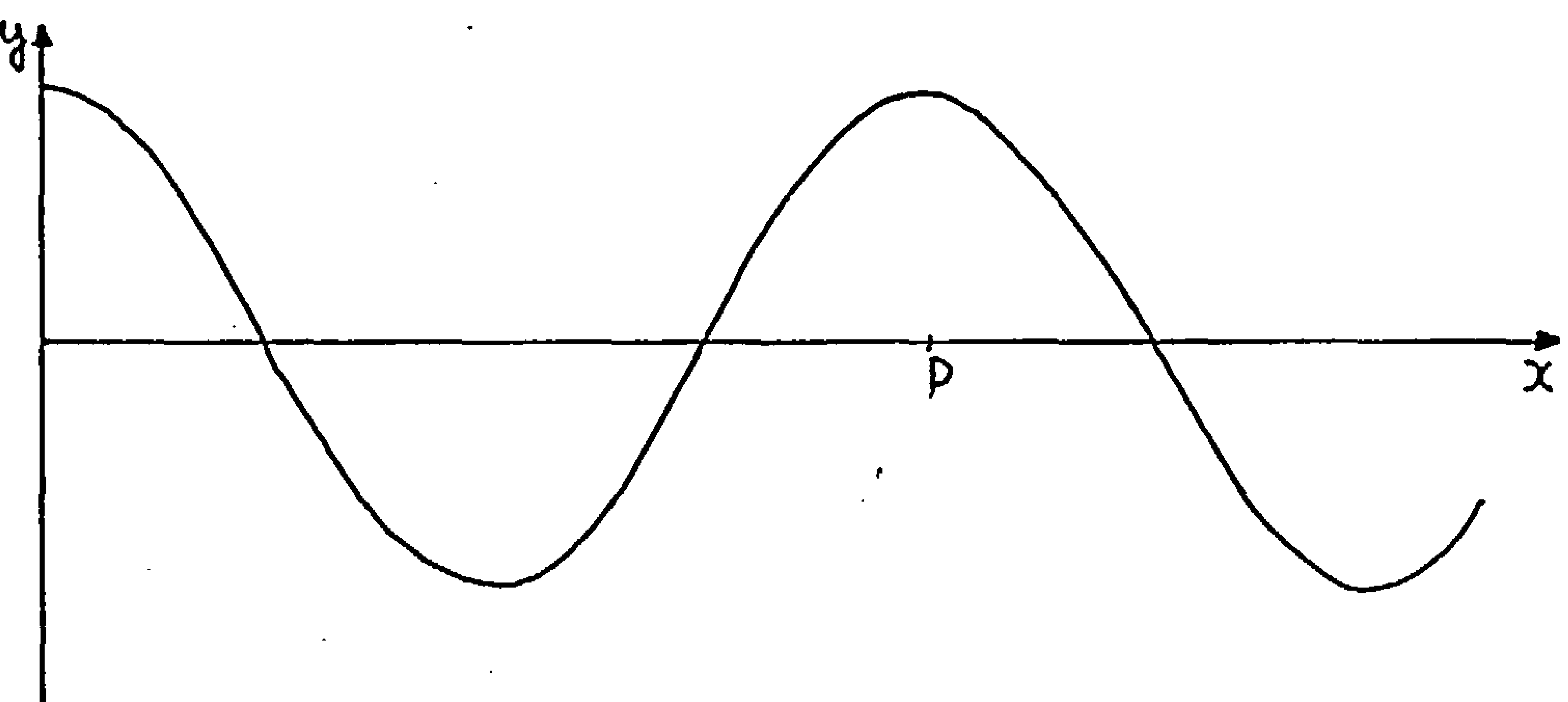


Figure 9.3a Sinusoidal surface profile.

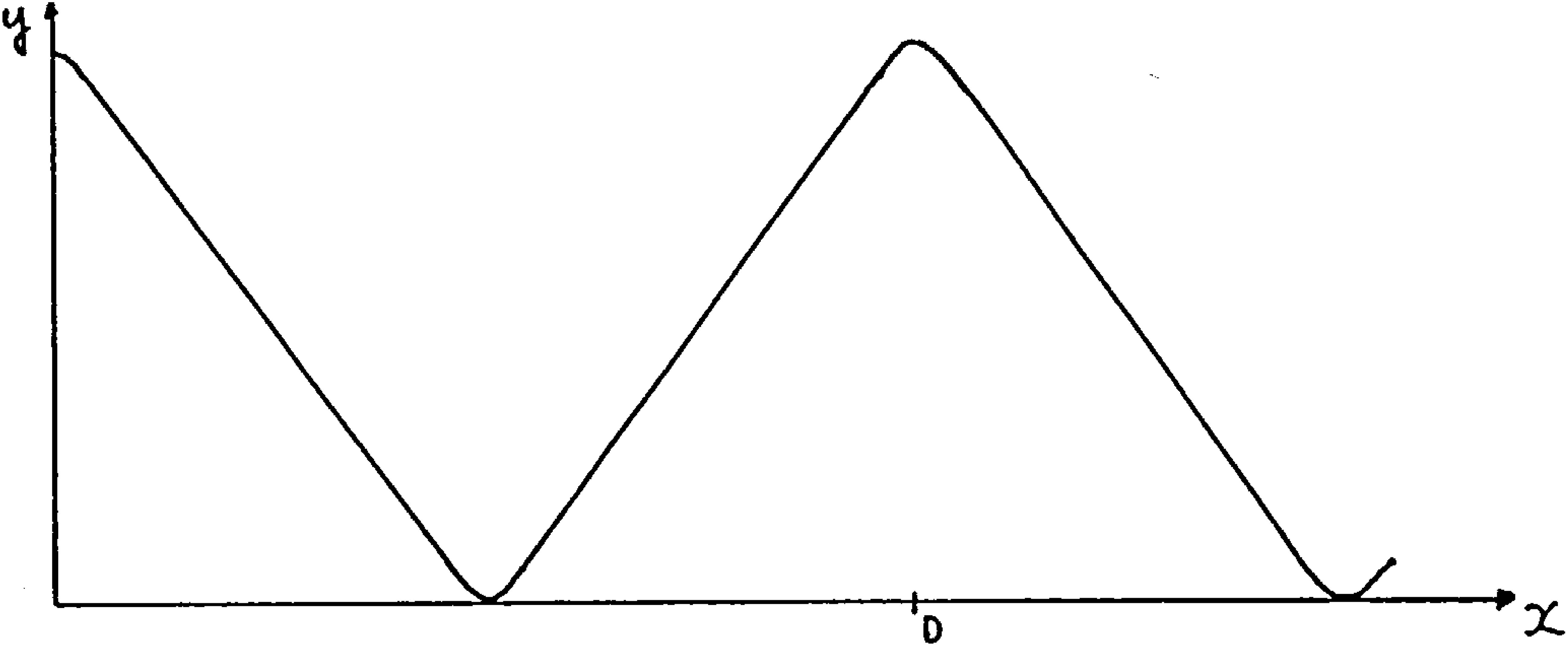


Figure 9.3b Approximated triangular surface profile.

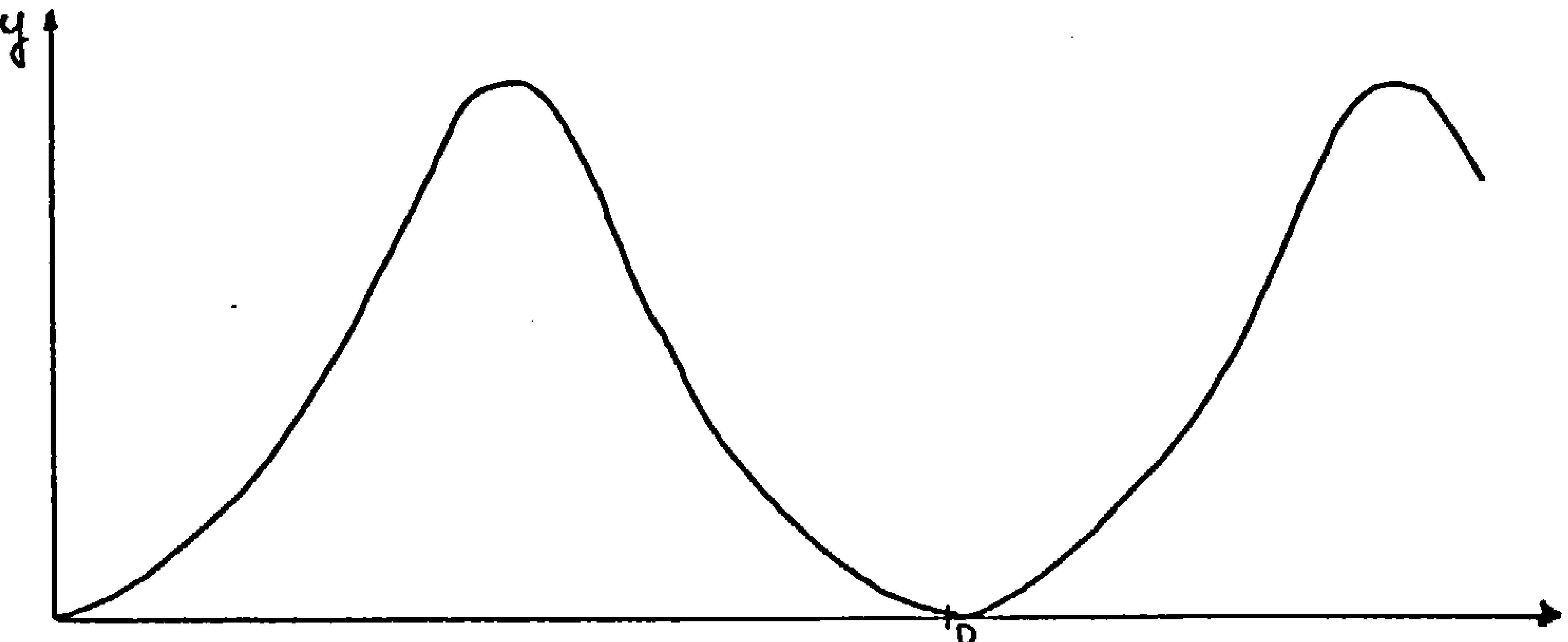


Figure 9.3c Approximated rectified sinusoidal surface profile.

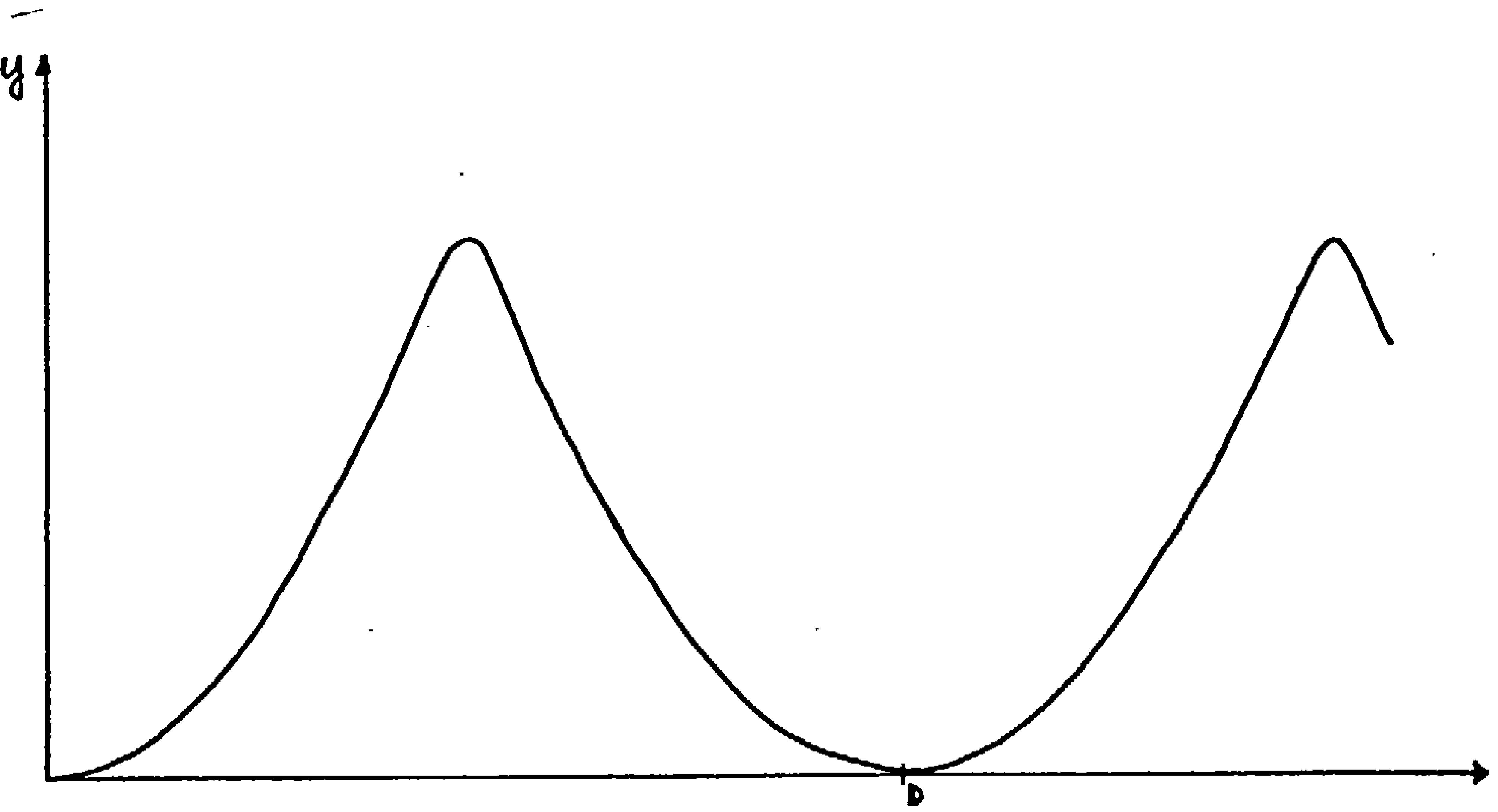


Figure 9.3d Approximated parabolic surface profile.

Expression (9.26) can be rewritten as:

$$y(x) = \left(R_q \sum_{n=1}^{7} a_n \cos(2\pi n x/D)\right) \Big/ \sqrt{\tfrac{1}{2}\sum_{n=1}^{7} a_n^2} \tag{9.27}$$

insertion in formula (9.9) gives the next expression for the phase advance $\varphi(x)$

$$\varphi(x) = 4\pi\sqrt{2}\,\frac{R_q}{\lambda}\left(\sum_{n=1}^{7} a_n \cos(2\pi n x/D)\right) \Big/ \sqrt{\sum_{n=1}^{7} a_n^2} \tag{9.28}$$

The resulting exponential function p(x),(9.7), has been expanded up to 12 terms, being:

$$p(x) = e^{-i\varphi(x)} = \sum_{l=0}^{12} \left(i\varphi(x)\right)^{l}/l! \tag{9.29}$$

This summation results in a series which can be represented by:

$$p(x) = \sum_{m=0}^{84} i^m A_m \cos(2\pi m x/D) \tag{9.30}$$

The profiles are even functions of x, from which it follows that p(x) is an even function of x, therefore the Kirchhoff integral can be replaced by:

$$U(X) \propto \int_{-\infty}^{\infty} p(x) \cos\left(\frac{kXx}{F}\right) dx \tag{9.31}$$

Insertion of (9.30) into expression (9.31), and evaluating this integral leads to a solution of the form:

$$U(X) = \sum_{m=-84}^{84} B_m \, \delta\left(X = \frac{2\pi m}{D} \frac{F}{k}\right) \tag{9.32a}$$

or

$$U(X) = \sum_{m=-84}^{84} B_m \, \delta\left(X = \frac{m\lambda F}{D}\right) \tag{9.32b}$$

where $\delta\left(X = \frac{m\lambda F}{D}\right)$ is a Dirac δ function and the resulting expression for the intensity I(X) is, according to formula (9.21):

$$I(X) = \sum_{m=-84}^{84} |B_m|^2 \, \delta\left(X = \frac{m\lambda F}{D}\right) \tag{9.33}$$

Here $|B_m|^2$ is the light intensity of the mth diffraction order in the diffraction plane at the position $X = m\lambda F/D$, as depicted in figure 9.4.

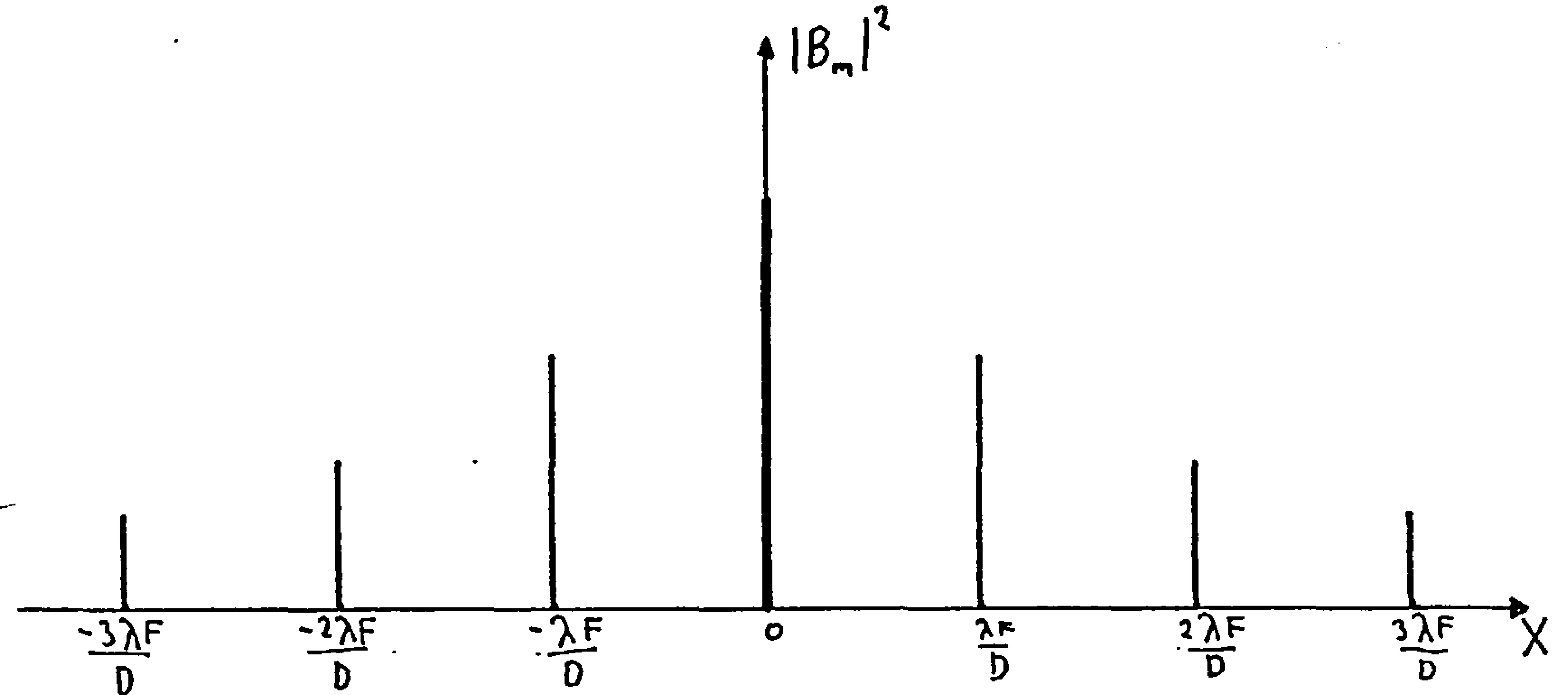


Figure 9.4. Idealised light intensity pattern in the diffraction plane.

It must be remembered that the values $|B_m|^2$ are relative intensity values, which are normalised with respect to the total reflected light.

The values for B_m have been calculated for values of $R_q/\lambda < 0.20$ and for $0 \leqslant m \leqslant 7$. Appendix A3 shows the listing of the program for the evaluation of $|B_m|^2$. The results are plotted in the figures 9.5a-d. For

each of the results the following summation $S(R_q/\lambda)$ has been calculated:

$$S(R_q/\lambda) = \sum_{m=-84}^{84} |B_m|^2 \tag{9.34}$$

This summation represents the total reflected light intensity in the diffraction plane, which must be constant because of conservation of energy. Over the mentioned surface roughness range the calculated value of $S(R_q/\lambda)$ equals unity. From this it can be concluded that the expansion of p(x) up to 12 terms is sufficient.

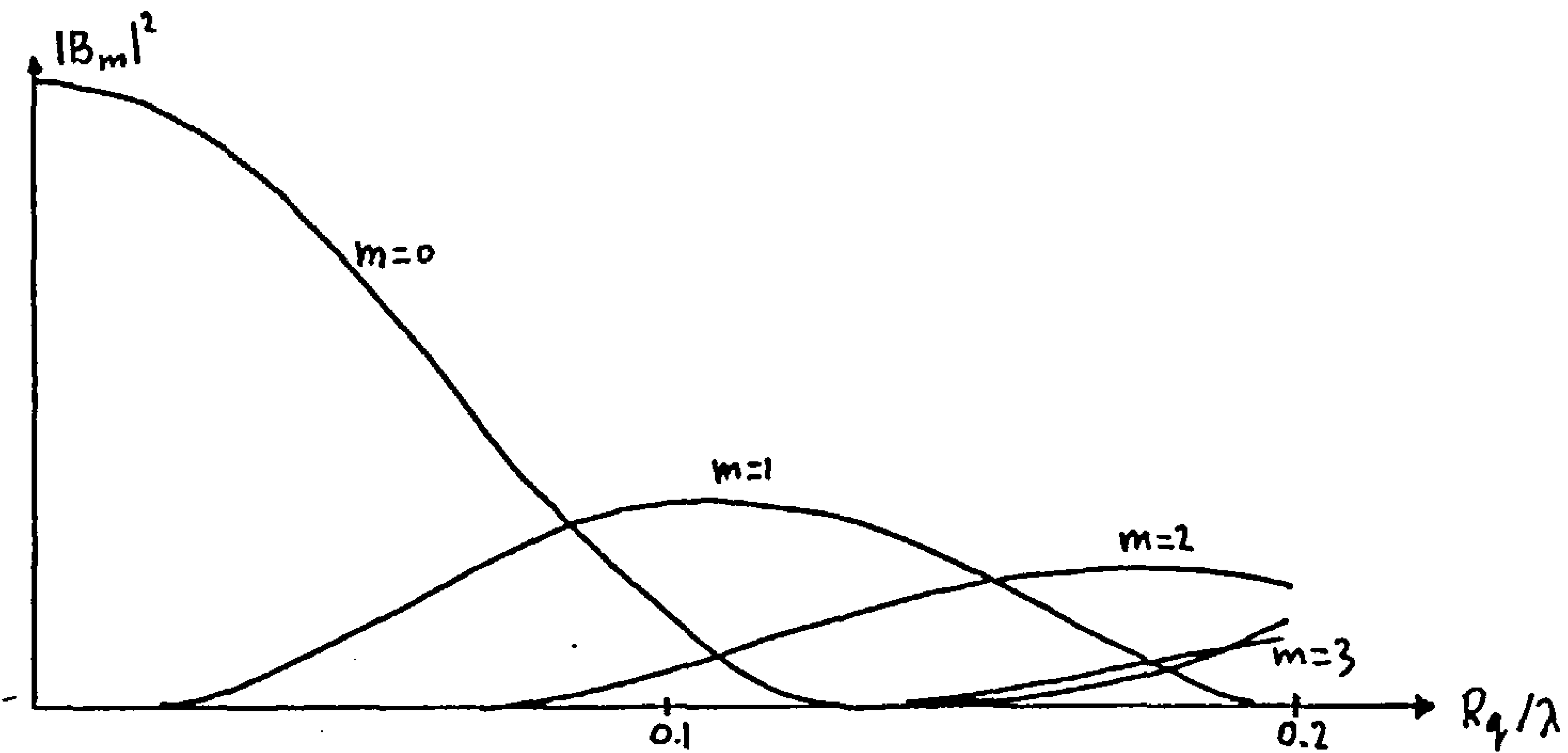


Figure 9.5a Intensity amplitudes for a sinusoidal surface profile.

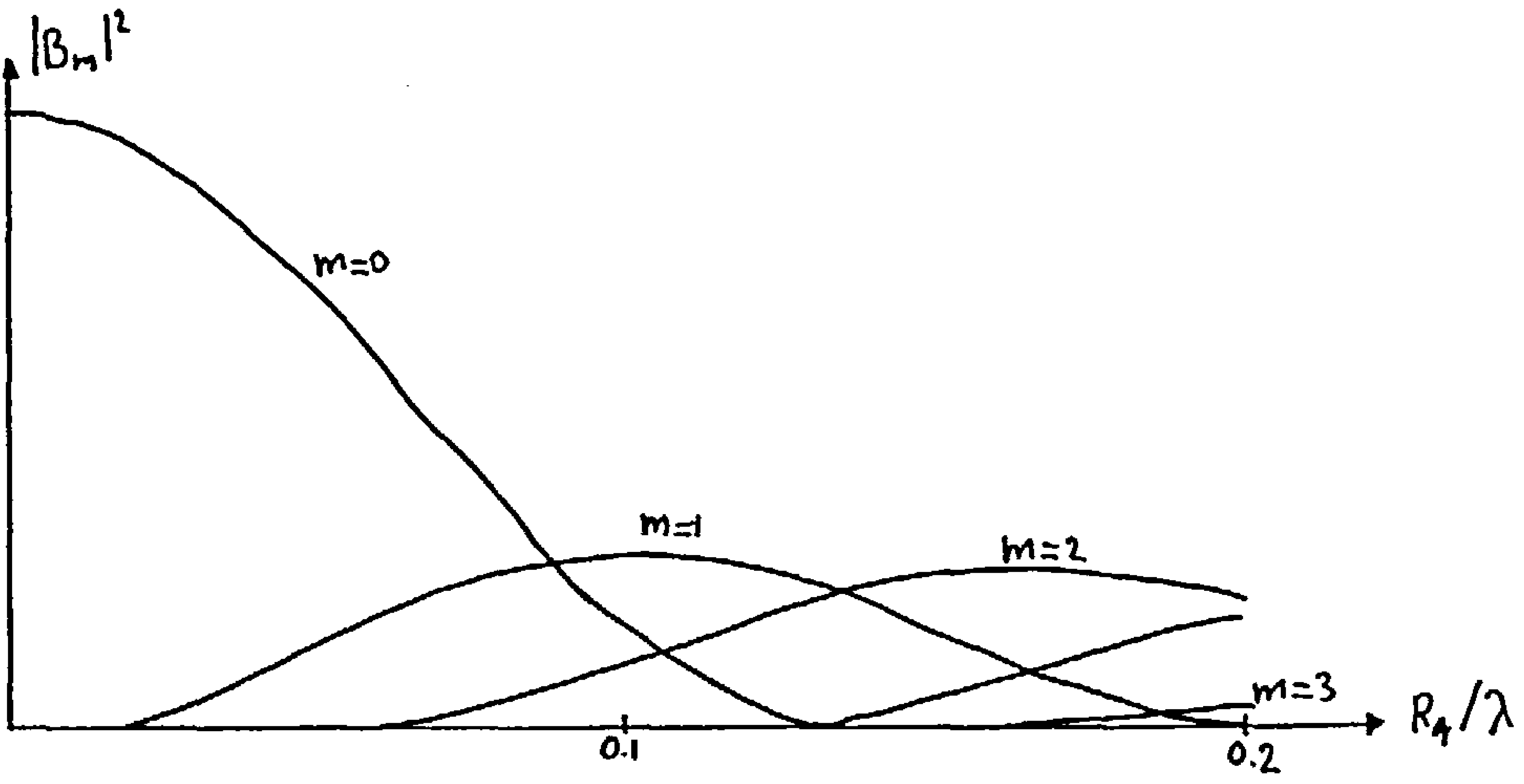


Figure 9.5b Intensity pattern for a triangular surface profile.

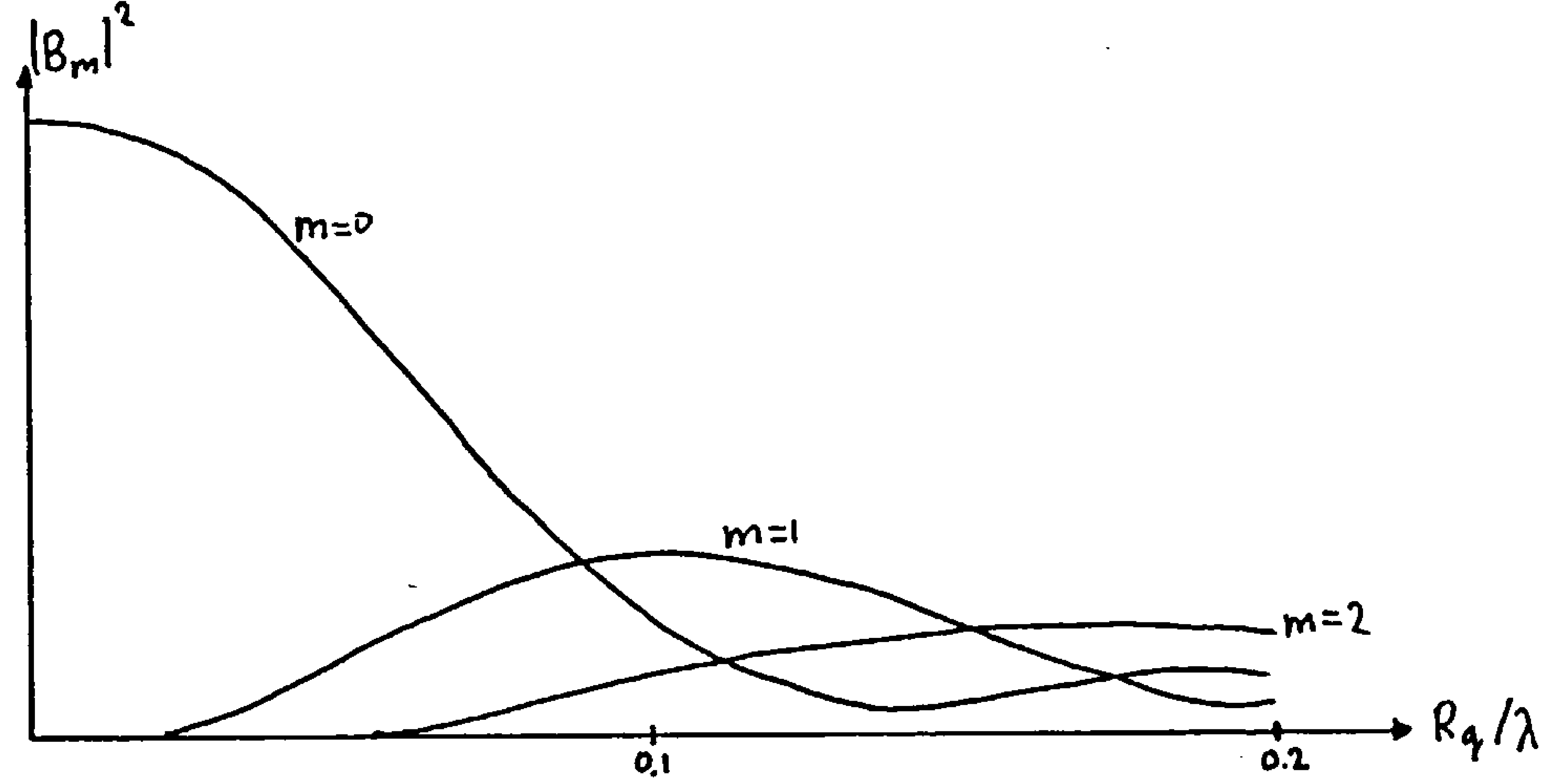


Figure 9.5c Intensity pattern for a rectified sinusoidal profile.

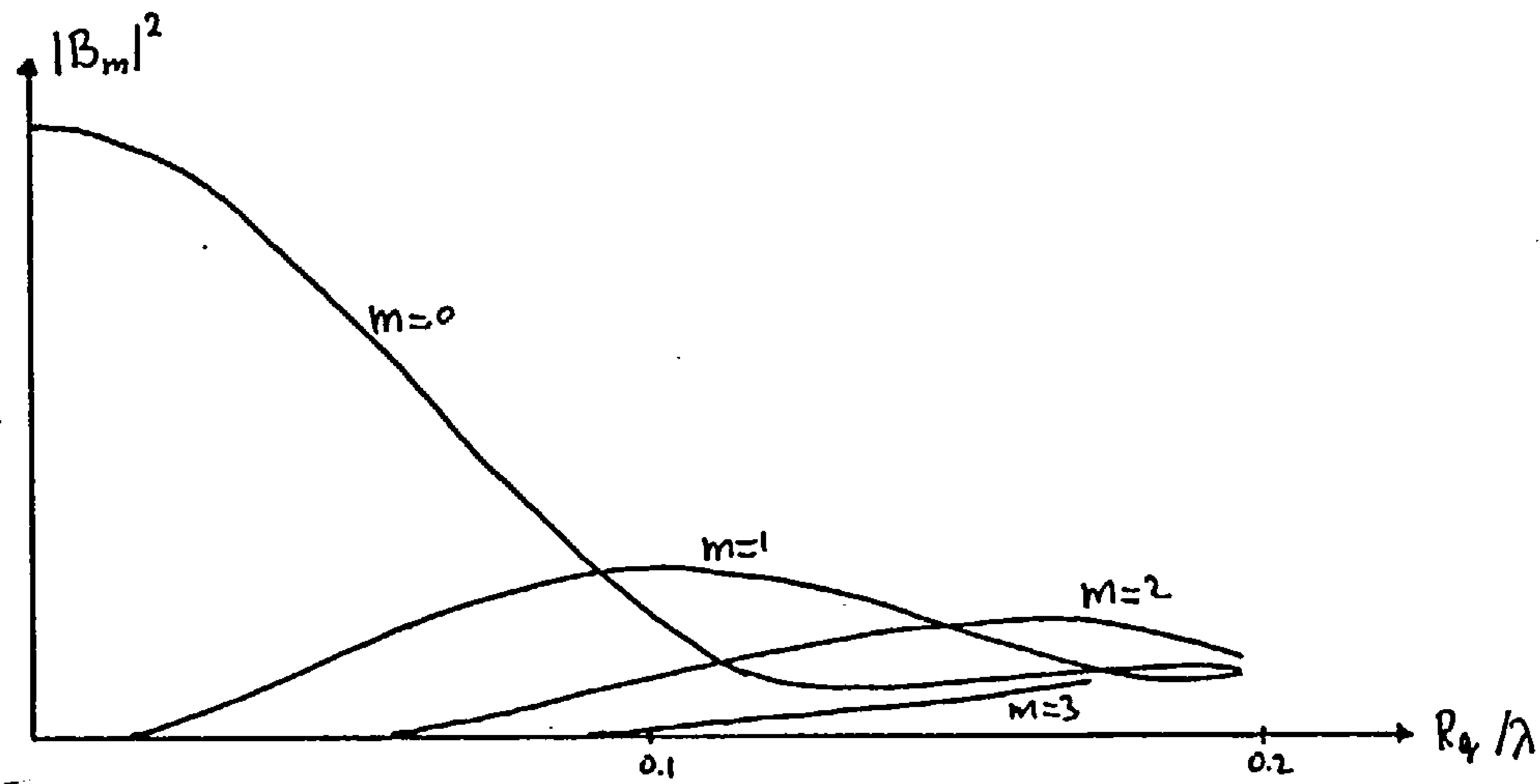


Figure 9,5d Intensity pattern for a parabolic surface profile.

Comparison of the curves of the figures 9.5.a-d shows that for $R_q/\lambda < 0.10$ the zero and first orders (m=0 and 1) do not differ much from one another. This similarity between the intensity curves has been employed to derive three algorithms for the determination of R_q/λ by measuring the amplitudes of the diffraction orders of optical diffraction patterns produced by periodic surfaces.

9.4. <u>Derivation of the algorithms for extracting the R_q parameter of periodic surfaces from their diffraction patterns.</u>

The fact that the zero orders produced by the surface profiles, expressed by (9.27) are virtually identical for $R_q/\lambda < 0.1$, can be used to derive three algorithms to extract the R_q value of these profiles. In order to obtain these algorithms, the zeroth order for the cosine profile is calculated using (9.8) and (9.27), which leads to:

$$U(X) \propto \int_{-\infty}^{\infty} e^{i4\pi\sqrt{2}\frac{R_q}{\lambda}\cos(2\pi x/D)} e^{-i\frac{kXx}{F}} dx \tag{9.35}$$

Which reduces for the zeroth order (X=0) to:

$$U(0) \propto \int_{-\infty}^{\infty} e^{i4\pi\sqrt{2}\frac{R_q}{\lambda}\cos(2\pi x/D)} dx \tag{9.36}$$

Substitution of $Z=4\pi\sqrt{2}\,R_q/\lambda$ and $Y=2\pi x/D$ yields:

$$U(0) \propto \int_{-\infty}^{\infty} e^{iZ\cos Y} dY \tag{9.37}$$

which is one of the integral representations of the Bessel function of the first kind of integral order zero, $J_o(Z)$. The amplitude of the light intensity then becomes $I_o(0)=J_o^2(Z)$.

$J_o(Z)$ can also be expressed by the following series expansion:

$$J_o(Z) = \sum_{n=0}^{\infty} \left(-\left(\frac{Z}{2}\right)^2\right)^n / (n!)^2 \tag{9.38}$$

For small values of Z, only a few terms of this series are necessary to approximate $J_o(Z)$ as can be seen from table 9.2. Here the first column represents Z and the second one the values of $J_o(Z)$ as found in the Handbook of Mathematical Functions by Abramowitz and Stegun [1972]. The third, fourth and fifth column represent the values of expression (9.38) for respectively N=1,2 and 3.

Table 9.2.

Z	$J_o(Z)$	N=1	N=2	N=3
0.0	1.000	1.000	1.000	1.000
0.1	0.997	0.997	0.997	0.997
0.2	0.990	0.990	0.990	0.990
0.3	0.977	0.977	0.977	0.977
0.4	0.960	0.960	0.960	0.960
0.5	0.938	0.937	0.938	0.938
0.6	0.912	0.910	0.912	0.912
0.7	0.881	0.877	0.881	0.881
0.8	0.846	0.839	0.846	0.846
0.9	0.807	0.797	0.807	0.807
1.0	0.765	0,749	0.765	0.765
1.1	0.719	0.697	0.720	0.719
1.2	0.671	0.639	0.672	0.671
1.3	0.620	0.577	0.622	0.620
1.4	0.566	0.509	0.570	0.566
1.5	0.511	0.437	0.516	0.511
1.6	0.455	0.359	0.462	0.455
1.7	0.397	0.277	0.408	0.397

From this table it is obvious that a good approximation of $J_o(Z)$ can be achieved for the following ranges of Z.

$$0.0<Z<0.4 \qquad J_o(Z)=1-(Z/2)^2 \tag{9.39a}$$

$$0.4<Z<1.0 \qquad J_o(Z)=1-(Z/2)^2+(Z/2)^4/4 \tag{9.39b}$$

$$1.0<Z<1.7 \qquad J_o(Z)=1-(Z/2)^2+(Z/2)^4/4-(Z/2)^6/36 \tag{9.39c}$$

For Z>1.7 there is no unique relationship anymore between the value of $J_o(Z)$ and Z, which renders the zeroth order useless for the determination of Z and therefore R_q.

Since $J_o(Z)$ can be extracted from the diffraction pattern by the formula:

$$J_o(Z) = \sqrt{\frac{B_o^2}{\sum_{n=-\infty}^{\infty} B_n^2}} \tag{9.40}$$

it follows that $(Z/2)^2$ can be calculated by solving the equations (9. 39a,b and c). The value of Z can then be converted into the roughness value R_q/λ by:

$$R_q/\lambda = Z/(4\pi\sqrt{2}) \tag{9.41}$$

Three algorithms for the estimation of the R_q/λ parameter of periodic surface profiles from their diffraction pattern based upon the formulae (9.39a,b and c), are derived in appendix A4. The results are listed below together with their respective roughness ranges. The

value of J_o is established by formula (9.40), where B_o is the light reflected into the zeroth order and the summation represents the total of the reflected and scattered light.

Algorithm 1 for $0.000<R_q/\lambda<0.022$ or $0.960<J_o<1$

$$R_q/\lambda = \sqrt{2(1-J_o)}\,/4\pi \qquad (9.42a)$$

Algorithm 2 for $0.022<R_q/\lambda<0.056$ or $0.765<J_o<0.960$

$$R_q/\lambda = \sqrt{(1-\sqrt{J_o}}\,/2\pi \qquad (9.42b)$$

Algorithm 3 for $0.056<R_q/\lambda<0.096$ or $0.397<J_o<0.765$

$$R_q/\lambda = \sqrt{2(S+T+3)}\,/4\pi \qquad (9.42c)$$

with $S = \sqrt[3]{R+\sqrt{D}}$ $\qquad T = \sqrt[3]{R-\sqrt{D}}$

and

$$D = 108(1+3J_o+3J_o^2) \qquad R = -9(1+2J_o)$$

The accuracy with which R_q/λ can be obtained depends on the accuracy with which the normalised zeroth order intensity $I_o = \frac{B_o^2}{\sum_{m=-\infty}^{\infty} B_m^2}$ can be determined. Figure 9.6 depicts absolute value of the percentual error in R_q/λ as a result of the percentual error in the determination of I_o. The derivation of these results can be found in appendix A5.

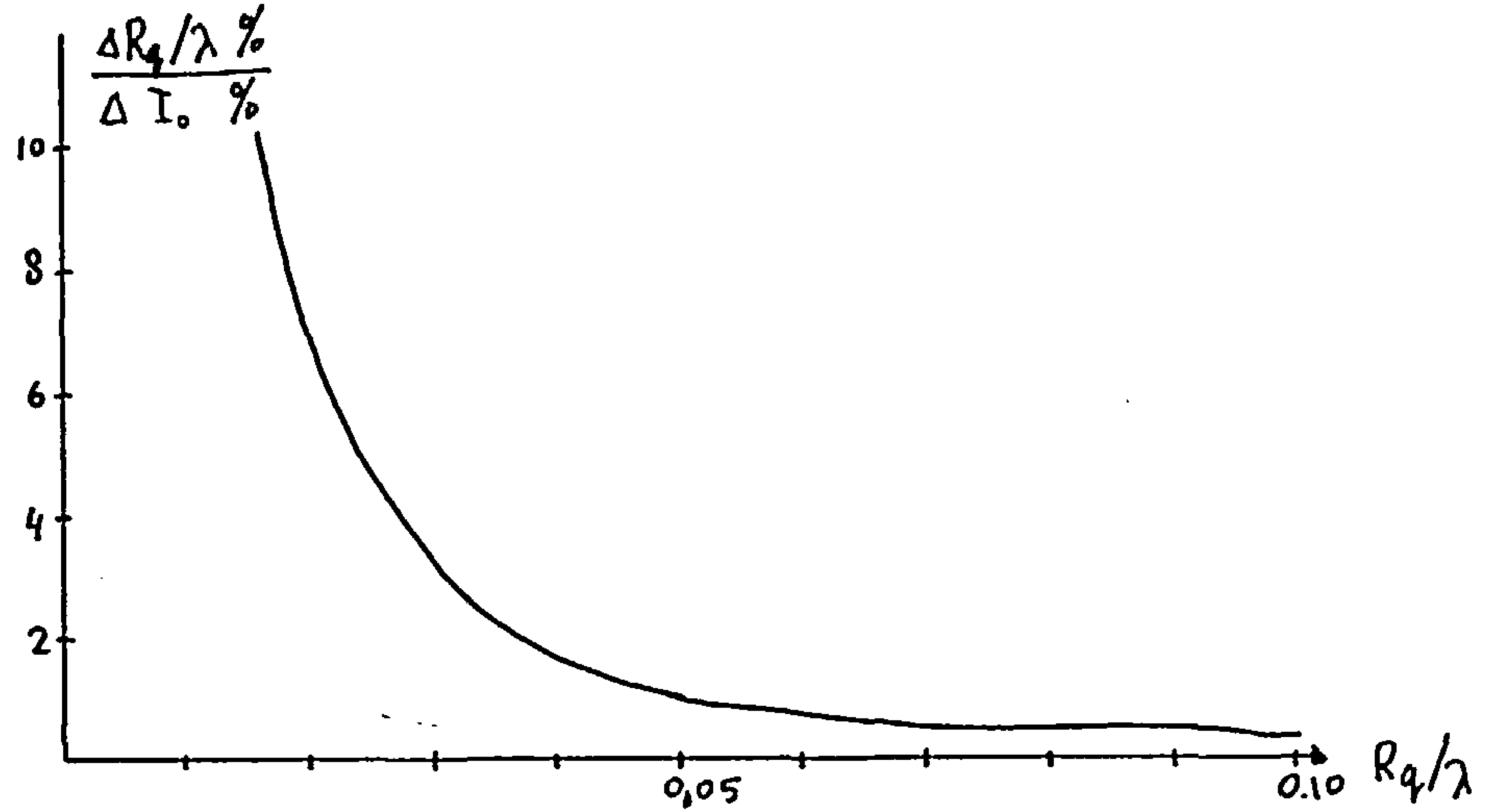


Figure 9.6 Ratio of the accuracy of R_q/λ and the accuracy of I_o.

9.6. Relationships between the second and fourth moment of the diffraction pattern produced by a sinusoidal profile and the constants defining this profile.

9.6.1. The second moment.

Let the profile be defined by $y(x)=h\cos(2\pi x/D)$, where h is the amplitude and D the period length of the profile. This profile produces a diffraction pattern whose intensity amplitudes can be expressed by $J_n(4\pi h/\lambda)$ and which are positioned in the diffraction plane at at $X=n\lambda/D$. $J_n()$ are Bessel functions of the first kind of integral order n. The second moment of this diffraction pattern is:

$$M_2 = \sum_{n=-\infty}^{\infty}\left(\frac{n\lambda}{D}\right)^2 J_n^2\left(\frac{4\pi h}{\lambda}\right) \qquad (9.43)$$

The evaluation of this expression can be found in appendix A6. This gives:

$$M_2 = 2\left(\frac{2\pi h}{D}\right)^2 \tag{9.44}$$

The rms slope Δ_q can be related to this expression, since

$$\Delta_q^2 = \frac{1}{D}\int_0^D \left(\frac{d}{dx} h\cos(2\pi x/D)\right)^2 dx = \frac{1}{2}\left(\frac{2\pi h}{D}\right)^2 \tag{9.45}$$

From which the relationship between the second moment of the diffraction pattern and the rms slope follows

$$\Delta_q = \sqrt{M_2}/2 \tag{9.46}$$

9.6.2. <u>The fourth moment.</u>

The fourth moment of the diffraction pattern is:

$$M_4 = \sum_{n=-\infty}^{\infty}\left(\frac{n\lambda}{D}\right)^4 J_n^2\left(\frac{4\pi h}{D}\right) \tag{9.47}$$

which leads to the following expression (see appendix A6)

$$M_4 = M_2\left(3M_2/2 + (\lambda/D)^2\right) \tag{9.48a}$$

or

$$(\lambda/D)^2 = \frac{2M_4 - 3M_2^2}{2M_2} \tag{9.48b}$$

This can be related to R_q/λ since for a sinusoidal profile the following expression can be derived:

$$R_q^2 = \frac{D^2(\Delta q)^2}{(2\pi)^2} \qquad (9.49)$$

Insertion of (9.46) and (9.48b) into (9.49) yields:

$$\left(\frac{R_q}{\lambda}\right)^2 = \frac{M_2^2}{8\pi^2(2M_4 - 3M_2^2)} \qquad (9.50)$$

The expressions (9.46) and (9.50), which have been derived for a sinusoidal profile, have been evaluated for the four profiles listed in table 9.1. The values of $J_n^2()$ have been replaced by B_m^2, whose values are plotted in the figures 9.5a,b,c and d. The results are plotted in the figures 9.7 and 9.8.

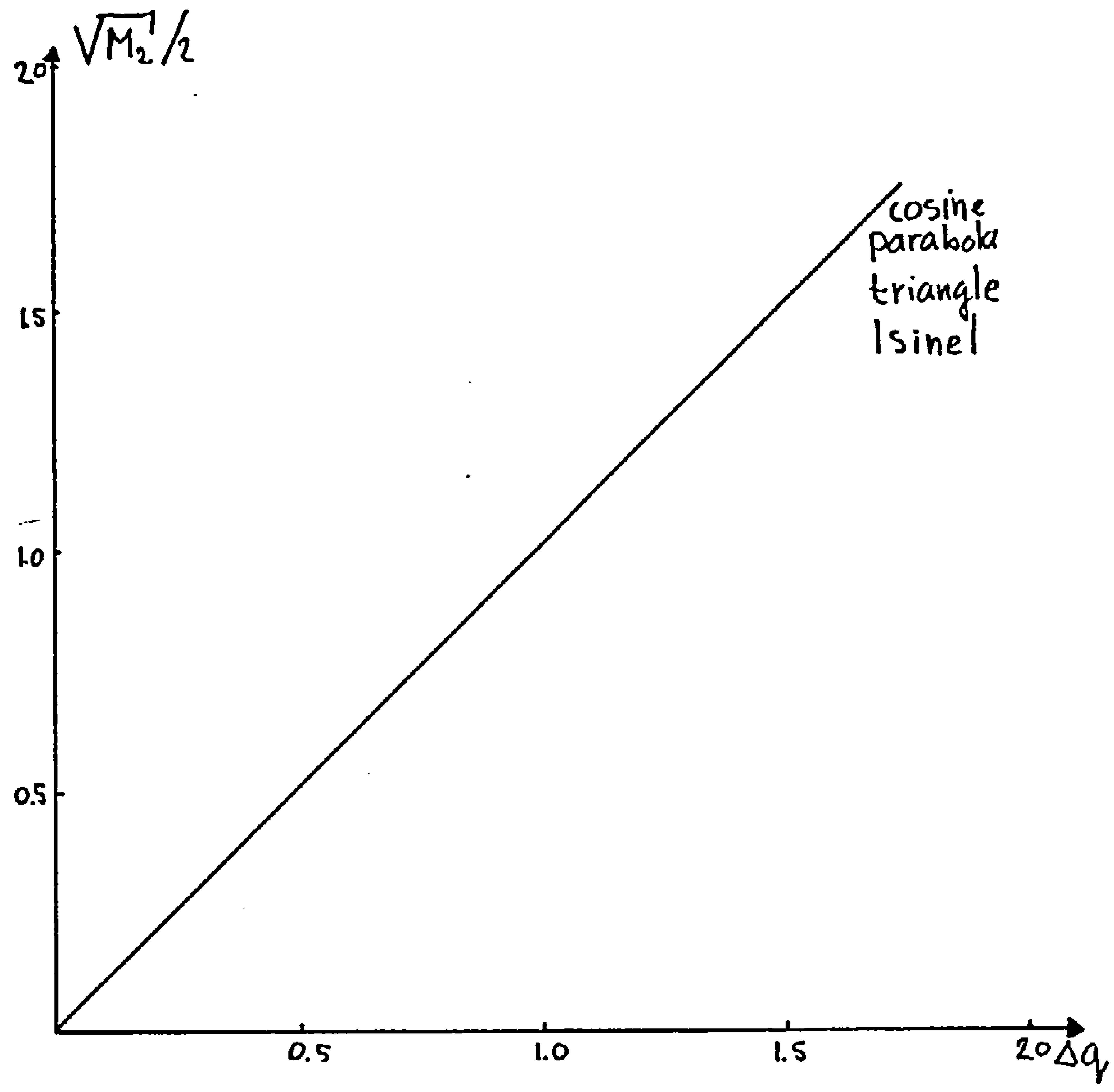


Figure 9.7 Relationship between the second moment and Δ_q.

In figure 9.7 the horizontal axis represents $\sqrt{M_2}/2$ and the vertical axis the rms slope value obtained by evaluating the definition for Δq (see page 17) for $0<R_q<0.26$. From figure 9.7 it can be concluded that formula (9.46) is not only valid for a sinusoidal profile but for the others as well. So (9.46) can be employed to obtain the rms surface slope from the second moment of the diffraction pattern.

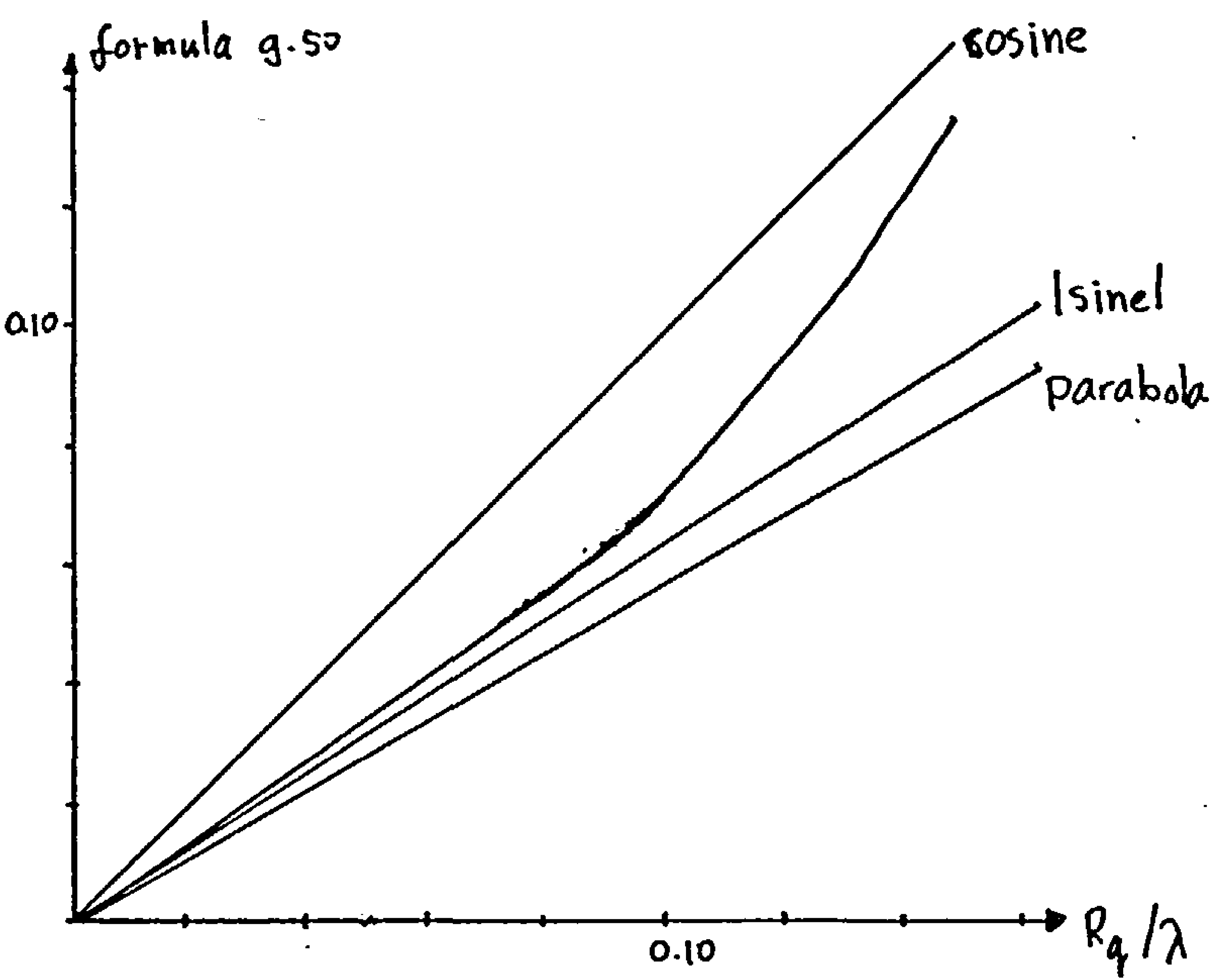


Figure 9.8. Relationship between R_q/λ and equation (9.50).

Figure 9.8 shows the relation between R_q/λ and the values obtained by evaluating (9.50). The horizontal axis represents the values from (9.59) and the vertical axis the values of R_q/λ obtained by:

$$\frac{R_q}{\lambda} = \sqrt{\frac{1}{D}\int_0^D \left(\frac{h}{\lambda}\sum_{n=1}^{7} a_n \cos(2\pi n x/D)\right)^2 dx} \qquad (9.51)$$

From figure 9.8 it can be concluded that formula (9.50) is only valid for the sinusoidal profile. The graph for the triangular

profile is curved due to the fact that for $R_q/\lambda > 0.1$ not enough diffraction orders are used to evaluate the fourth moment. The differences in the slopes of the graphs are partly due to the fact that the effective rms wavelength of these profiles differ from the period length D.

9.7. Validity range of the Kirchhoff diffraction integral for sinusoidal profiles.

In the derivation of the Kirchhoff diffraction integral, it has been assumed that no secondary scattering occurs, i.e. the diffracted rays do not hit the surface. This condition is true when the maximum surface slope is smaller than the tangent of the light reflected in the highest diffraction order m. This is a function of the amplitude h and the period length D of the surface profile $y(x)=h\cos(2\pi x/D)$. In order to estimate the validity condition, firstly the number of diffraction orders M necessary to contain 99.9%, 99% and 95% of the total reflected light, have been calculated according to the formula:

$$I(M) = \sum_{m=-M}^{M} J_m^2(x) > 0.999 \quad \text{or } 0.99 \quad \text{or } 0.95 \tag{9.52}$$

where $x=4\pi h/\lambda$. The results are listed in table 9.3.
Secondly, $\tan(\theta'_M)$ is calculated, where θ'_M is the angle under which the Mth order is reflected. Consequently $\tan(\theta'_M)$ is equated to the maximum surface slope $\tan(\phi)_{max}$. The various angles involved are depicted in figure 9.9 . The relevant expressions for these

quantities are:

$$\tan \theta'_M = \sqrt{\left(\frac{D}{M\lambda}\right)^2 - 1} \tag{9.53}$$

and

$$\tan \phi_{max} = \frac{2\pi h}{D} \tag{9.54}$$

Equating these expressions yields:

$$\sqrt{\left(\frac{D}{M\lambda}\right)^2 - 1} = \frac{2\pi h}{D} \tag{9.55}$$

The M and h values can be obtained from table 9.3 where $h = \lambda x / 4\pi$. Insertion into (9.55) results in

$$\sqrt{\left(\frac{D}{M\lambda}\right)^2 - 1} = \frac{\lambda x}{2D} \tag{9.56}$$

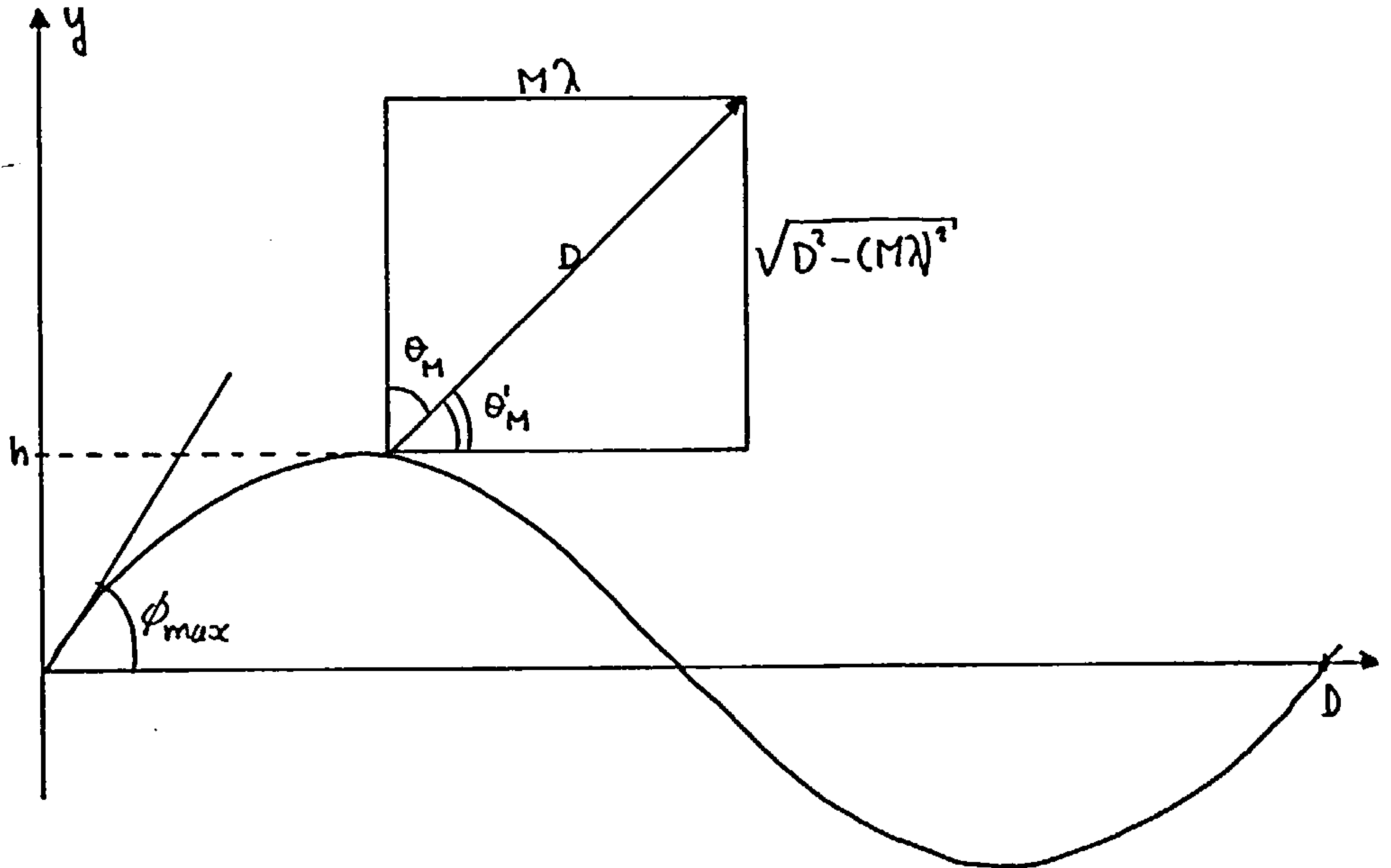


Figure 9.9 Maximum angle of diffraction and maximum surface slope.

Table 9.3

x	I(M)>0.999 M	I(M)>0.99 M	I(M)>0.95 M
0.2	1	1	0
0.4	2	1	1
0.6	2	1	1
0.8	2	2	1
1.0	3	2	1
1.2	3	2	2
1.4	3	2	2
1.6	3	2	2
1.8	4	3	2
2.0	4	3	2
2.2	4	3	3
2.4	5	3	3
2.6	5	4	3
2.8	5	4	3
3.0	5	4	3
3.2	6	4	4
3.4	6	5	4
3.6	6	5	4
3.8	6	5	4
4.0	7	5	4
4.2	7	5	5
4.4	7	6	5
4.6	7	6	5
4.8	7	6	5
5.0	8	6	5
6.0	9	7	6
7.0	10	8	7
8.0	11	9	8
9.0	12	10	9
10.0	13	11	10

Solving equation (9.56) for $(\lambda/D)^2$ gives

$$\left(\frac{\lambda}{D}\right)^2 = \frac{2}{M^2 + \sqrt{M^4 + M^2 x^2}} \tag{9.57a}$$

or

$$\frac{\lambda}{D} = \sqrt{2/(M^2 + \sqrt{M^4 + M^2 x^2})} \tag{9.57b}$$

and

$$\frac{h}{D} = \frac{h}{\lambda} \cdot \frac{\lambda}{D} = \frac{x}{4\pi} \sqrt{2/(M^2 + \sqrt{M^4 + M^2 x^2})} \tag{9.58}$$

In figure 9.10 H/d=2h/D is plotted versus λ/D using the formulae (9.57) and (9.58). The paired values M and x are obtained from table 9.3 for the case $I_M > 0.999$. H/D tends to a limit for for an increasing value of M. This limit equals $\sqrt{2/(1+\sqrt{2})}/2\pi = 0.1448$. This value is comparable with the value obtained by Millar (see section 7.1.3) for the validity of the Rayleigh hypothesis, there H/D equals 0.142.

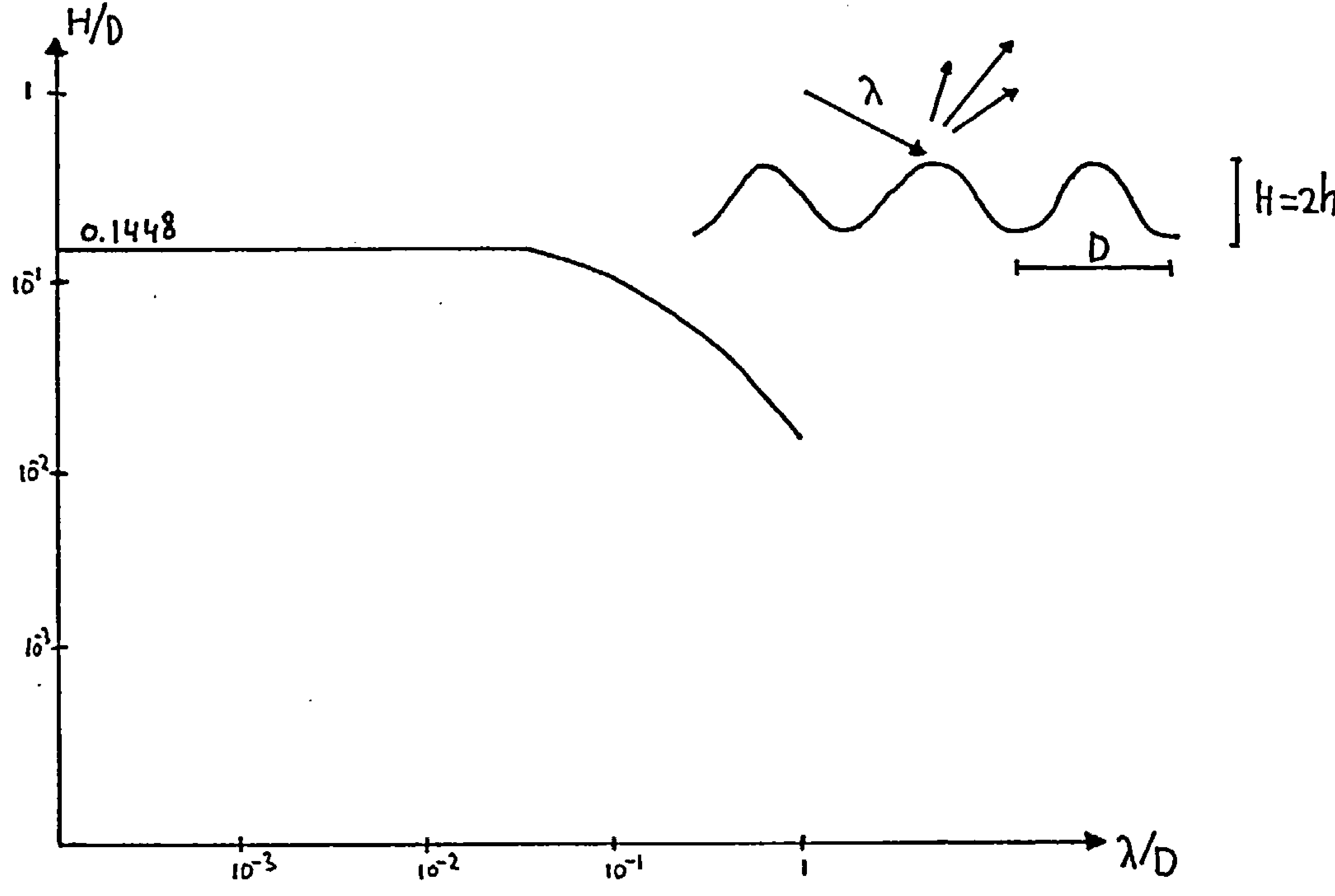


Figure 9.10 Validity range of the Kirchhoff diffraction approach.

Millar's analysis involves the location of the singularities in the representation of the solution of the wave problem and he relates these to the singularities and the critical points in the Green's function of the corresponding potential problem. The value 0.142 is the approximate solution of a certain transcendental equation.

Another interesting observation is that for $I(M)>0.95$ $M=x$, in other words, the diffraction orders necessary to contain 95% of the reflected light equals $x=4\pi h/D$.

9.8. Conclusions.

For $R_q/\lambda<0.03$ the diffraction pattern can be treated as the power-spectrum of the surface profile and therefore various statistics of the surface profile can be extracted from the diffraction pattern in the same way as outlined in section 4.1.3.

For $R_q/\lambda<0.096$ it is possible to obtain the R_q/λ value from the diffraction pattern by determination the normalised value of the light intensity reflected into the specular direction. When $R_q/\lambda>0.096$ then the normalised amount of light reflected into the specular direction is not an uniquely valued function of R_q/λ.

The algorithms for the extraction of the R_q/λ value are reliable for the roughness range $0.015<R_q/\lambda<0.096$.

The second moment of the diffraction pattern equals four times the mean square value of the surface slope. This relationship holds over the total range as long as the Kirchhoff approach is valid. This approach is valid as long as $h/D<0.0724$, where h is the amplitude and

D the period length of the surface profile.

From the R_q/λ and the rms slope Δ_q the rms surface wavelength can be determined by the obvious relationship:

$$\lambda_{rms} = \frac{2\pi R_q}{\Delta_q} \tag{9.59}$$

The diffraction orders necessary to contain 95% of the reflected light equals $M=4\pi h/D$. The maximum diffraction angle related to this value of M is determined by the grating formula:

$$\sin\theta_M = M\lambda/D \tag{9.60}$$

This angle θ_M can be related to the back focal-length, F, of a transform lens and the half-length, h/2, of the diffraction plane by:

$$\tan\theta_M = \frac{h/2}{F} = \frac{h}{2F} \tag{9.61}$$

10 Computer simulation of the optical diffraction system.

10.1 Introduction.

In chapter 9 the one dimensional Kirchhoff-Fresnel integral has been derived which relates the intensity distribution I(X) in the diffraction plane to the surface profile y(x) by:

$$I(X) = \left| \int_{-\infty}^{\infty} e^{-2iky(x)} \, e^{-ikXx/F} \, dx \right|^2 \qquad (10.1)$$

where $k=2\pi/\lambda$ is the wavenumber , λ the wavelength of the illuminating plane wavefront and F is the focal-length of the optical system to produce the Fraunhofer diffraction pattern I(X). The geometry related to this expression is shown in figure 10.1.

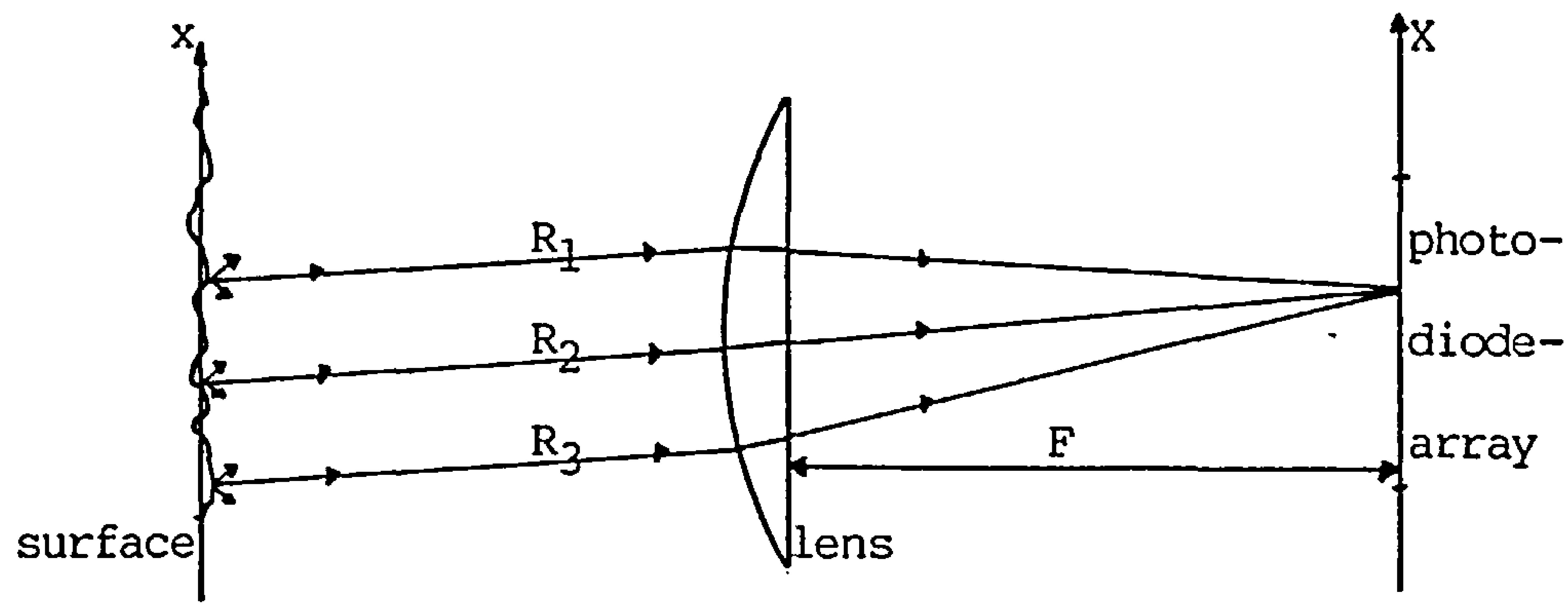


Figure 10.1. Geometry of the optical diffraction system.

In the derivation of the above integral the following assumptions have been made:

i the illuminating wavefront is plane, with uniform intensity and infinitely wide.

ii the profile y(x) exists for $-\infty < x < \infty$.

iii the transform lens with focal-length F has no abberrations.

In the actual system these conditions cannot be met. The laser output will normally produce a plane wavefront whose intensity distribution is Gaussian bell shaped. The surface profile has a finite length and the lens has abberations. In order to investigate the influence of these deviations from the ideal case a computer simulation program has been devised which simulates the actual sensor complete with the optical components involved. Whereas expression (10.1) can be easily evaluated for simple analytical expressions of y(x), this computer program has the additional advantage that diffraction patterns caused by complicated surface profiles can be generated. Furthermore this program can be utilised to produce an atlas which maps characteristic diffraction patterns versus machine-tool errors.

10.2 The computer program.

The computer program is based upon Huygens' principle which states that each point on a wavefront acts as a source of new spherical waves. In this particular case, the points from which the new waves or reflected rays are emanating, is the surface of the turned component. These rays travel via the transform lens to the linear

photo-diode array, positioned in the back-focal plane of the transform lens, which captures the diffraction pattern, see figure 10.1. The change of direction of the rays when crossing the boundaries of the lens are calculated by Snell's law. The relative light intensity in the diode-array plane is calculated by addition of the complex amplitudes of each individual ray, taking into account their relative phases. It is assumed that the surface is illuminated by a plane wavefront with a finite width (5mm). The intensity profile of this wavefront is expressed by I(x) in formula (10.2). The profile y(x) can either be generated by the computer or it can be the digitized output of a stylus instrument sampled at a lum interval. The surface profile is represented as a histogram, i.e. the slope of each facet, making up the profile is zero and there are discontinuities between the adjacent facets. The computer program inserts the surface profile into the mean surface plane y_o, which extends between the limits from $-x_o$ to $+x_o$. The light intensity I(X) in a point $P(X,Y_o)$, with $-X_o<X<X_o$, is then computed by the formula:

$$I(X) = \left[\sum_{-XD}^{XD} I(x)\cdot\cos(-k\,y(x)+R)\right]^2 + \left[\sum_{-XD}^{XD} I(x)\cdot\sin(-k\,y(x)+R)\right]^2 \qquad (10.2)$$

where R is the optical path length between the points (x,y(x)) on the surface and the point (X,Y_o) in the detector plane, whilst passing through the transform lens. This program leaves the user with the following options:

i any surface profile can be inserted.

ii the wavelength of the the laser can be changed.

iii the intensity profile of the illuminating wavefront can be adjusted.

iv the back-focal length of the lens can be changed.

v the spacing between the photo-diodes can be changed.

The program is written in FORTRAN7. It is run on a PRIME 750 computer. The CPU time taken by the program is 36 minutes. It calculates 16000 ray paths per 250 diodes, which constitutes one half photo diode array. To obtain a complete diffraction pattern the program has to be executed twice. The two half spectra are joined together by another program. The various variables used in the program can be traced in figure 10.2. Figure 10.3 shows the flow diagram of the computer diagram.

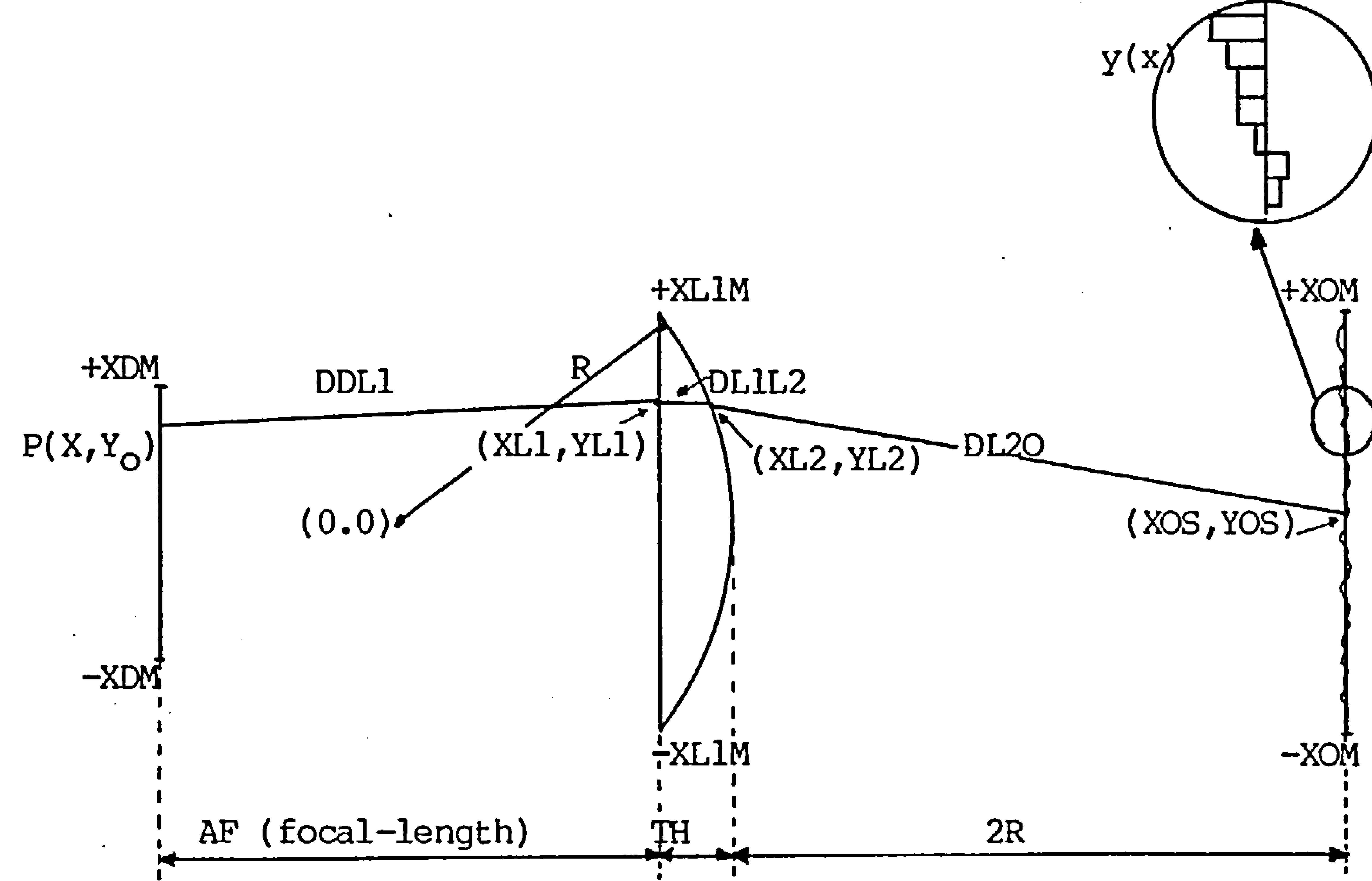


Figure 10.2. Diagram of the diffraction system showing the variables used in the computer program.

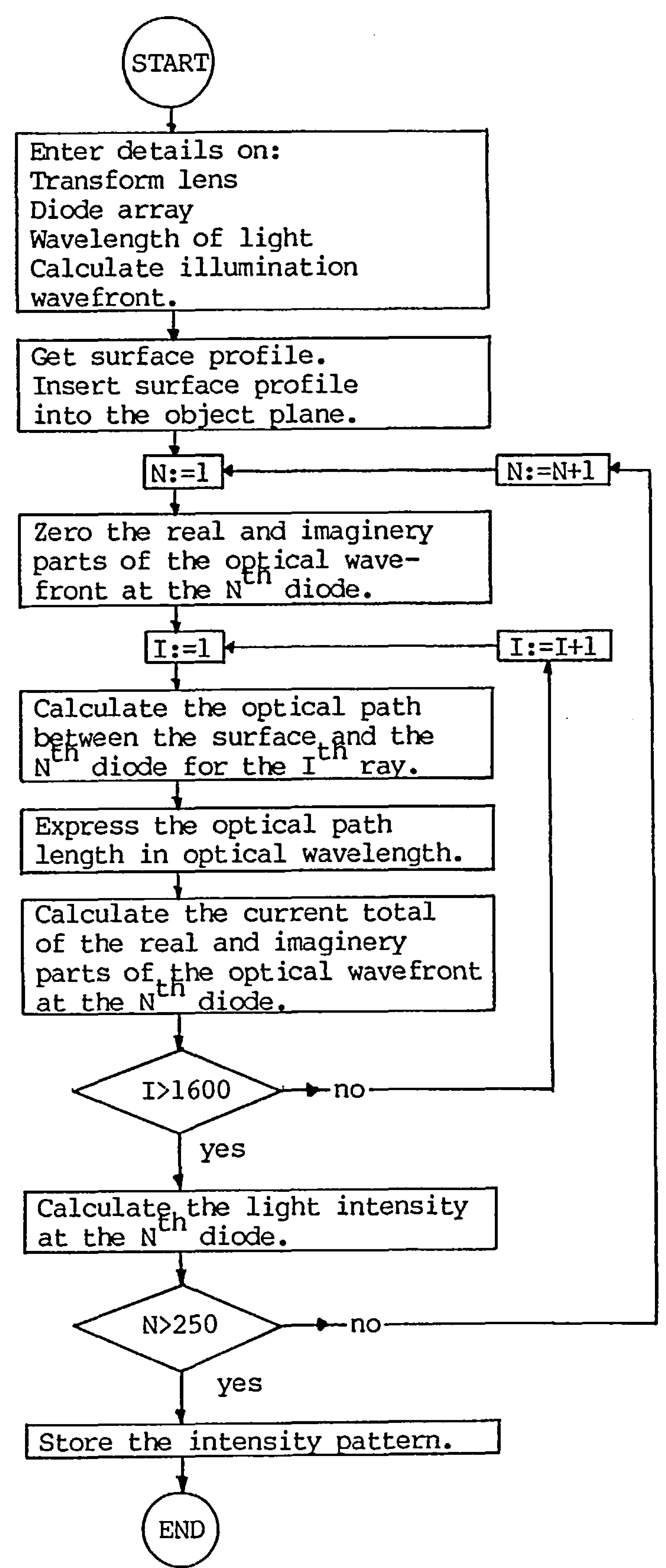


Figure 10.3. Flow diagram of the simulation program.

program listing and comments.

LISTING   comments

```
      DIMENSION YO(5000),SI(250),AMP(5000)
      ANS=1.52                      refractive index of the lens
      AF=270000.                    focal-length of the lens (um)
      TH=6000.                      lens thickness (um)
      R=(ANS-1.)*AF                 radius of the lens
      YL1=R-TH                      x ordinate of the flat lens surface
      PI=3.1416
      XL1M=8000.                    half height of the lens (um)
      XDM=5000.                     half height of the diode array (um)
      WV=.870                       wavelength of the laser (um)
      NL1M=16000                    number of rays per diode
      DXL1M=2.*XL1M/NL1M
      XOM=2500.                     half length of the surface profile
      YOB=R+AF                      x ordinate of the diode array
      DO 20 I=1,5000                define the intensity of the illumi-
20    AMP(I)=EXP(-((I-2500.)/2500.)**2)      nating wavefront
      GOTO 14                       leave in for defining the surface
      CALL FOPEN(6,7,'PROFILE',1,IERR)
      READ(6,10) (YO(I),I=1,5000)      reads in external profile data
      CALL FCLOSE(6,IERR)
10    FORMAT(F19.1)
14    DO 15 I=1,5000                   generates its own profile data
15    YO(I)=YOB+0.070*SIN(.01*PI*I)
```

```
DO 30 N=1,250
SINE=0.                       zero the real and imaginery parts
CO=0.                         of the optical wavefront.
DO 40 I=1,NL1M
XD=XDM*(2.*N/500.-1)          calculation of the ray paths
XL1=-XL1M+I*DXL1M             between the diodes and the flat
AL1=(XL1-XD)/AF               lens surface.
ARS=(AL1)/ANS
AL2=(ARS/SQRT(1.-ARS**2))
T=AL2
Q=XL1-T*YL1
YL2=(-Q*T+SQRT((Q*T)**2-(1.+T**2)*(Q**2-R**2)))/(1.+T**2)
XL2=Q+T*YL2                   calculation of the ray inside the
AL3=ATAN(XL2/YL2)             lens
AL4=AL2-AL3
ARS=ANS*SIN(AL4)
AL5=ATAN(ARS/SQRT(1.-ARS**2))
AL6=AL3+AL5                   calculation of the rays between
XO=XL2+(YOB-YL2)*AL6          the curved lens surface and the
IF (ABS(XO) .GE. XOM) GO TO 40     profile
IH=INT(XO+XOM)
YOS=YO(IH)
XOS=XL2+(YOS-YL2)*AL6
DDL1=SQRT(AF**2+(XL1-XD)**2)-AF distance between diode and lens
DDL1=DDL1-WV*INT(DDL1/WV)       fractional distance
DL1L2=ANS*SQRT((XL2-XL1)**2+(YL2-YL1)**2)
DL1L2=DL1L2-WV*INT(DL1L2/WV)    distance travelled in the lens
```

```
      DL2O=SQRT((XOS-XL2)**2+(YOS-YL2)**2)-AF
      DL2O=DL2O-WV*INT(DL2O/WV)        distance between the lens  and
      DTO=(DDL1+DL1L2+DL2O)*2.*PI      total distance.  (the profile
      SINE=SINE+AMP(IH)*SIN(DTO/WV)    update the current values
      CO=CO+AMP(IH)*COS(DTO/WV)        of the optical wavefront
40    CONTINUE
      SI(N)=SINE**2+CO**2              calculation of the intensity
30    CONTINUE
      CALL FOPEN(6,8,'SPECTRUM',3,IERR)
      WRITE(6,80) (SI(J),J=1,250)
      CALL FCLOSE(6,IERR)
80    FORMAT(F11.1)
      STOP
      END
```

10.3. Results obtained by the computer simulation program.

Firstly, the simulation program has been tested by insertion of a flat surface, $y(x)=0$, into the object plane XOB, under two different illumination conditions:

i $I(x)=1$ an illumination wavefront with uniform light intensity.

ii $I(x)=e^{-\left(\frac{x}{a}\right)^2}$ a Gaussian-bell shaped intensity distribution, with $a=2.5mm$.

The theoretical results for these two cases are:

$$I(X) = \left|\int_{-a}^{a} e^{-\frac{ikXx}{F}} dx\right|^2 \propto \sin^2\left(\frac{kXa}{F}\right) \Big/ \left(\frac{kXa}{F}\right)^2 \qquad (10.3a)$$

and

$$I(X) \approx \left|\int_{-\infty}^{\infty} e^{-\left(\frac{x}{a}\right)^2} e^{-\frac{ikXx}{F}} dx\right|^2 \propto e^{-2\left(\frac{kXa}{2F}\right)^2} \qquad (10.3b)$$

Expression (10.3a) represents a light intensity distribution consisting of one principal maximum with neighbouring higher order maxima. The intensity amplitudes of these subsidiary maxima are less than 4.7% of the value of the principal maximum. Expression (10.3b) represents a single Gaussian-bell shaped intensity peak. By defining the half-width, W/2, of the intensity peaks as the value of X over which the amplitude of the diffraction peaks have decayed to 10% of their maximum value, leads to:

$$W_a=2\cdot 2.32F/ka=69um$$

and

$$W_b=2\sqrt{2.3/2}\,2F/ka=64um$$

Here the focal-length F=270mm, a=2.5mm and λ=.870um. The peaks produced by the simulation program, figures 10.4a and b, have 10% width of approximately 180um. This extra broadening of the peaks can be attributed to the abberrations of the lens. The figures 10.4c and d are enlargements of the bases of the two diffraction peaks. From this it is clear that the Gaussian shaped intensity illumination produces a less noisy diffraction pattern than the unity intensity illumination. Figure 10.4e shows the diffraction obtained by an

Figure 10.4a.

Figure 10.4b.

Figure 10.4c.

Figure 10.4d.

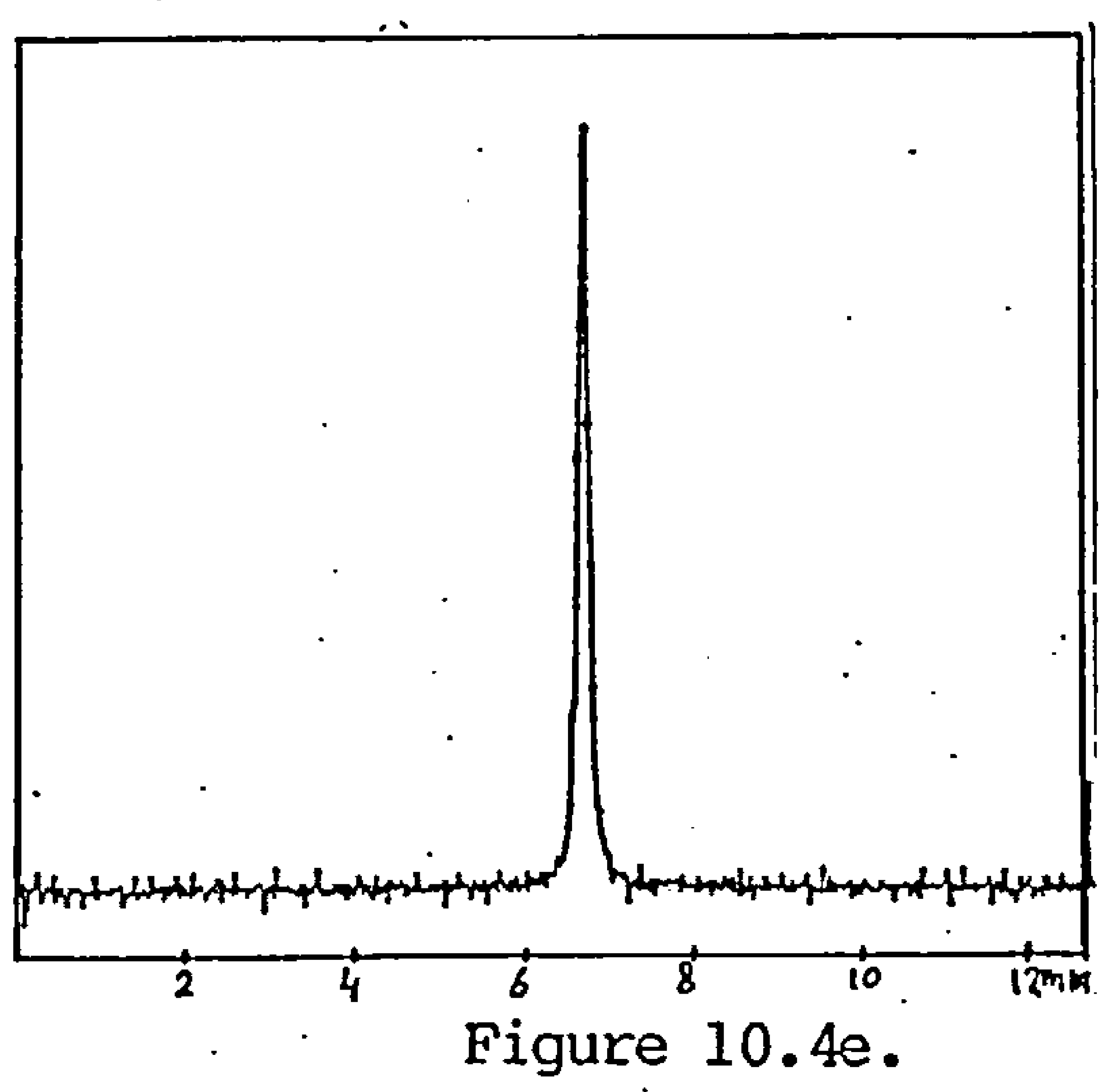


Figure 10.4e.

Fig.10.4a Diffraction pattern produced by an uniform light intensity.

Fig.10.4b Diffraction pattern produced by a Gaussian light intensity.

Fig.10.4c Enlargement of fig. 10.4a

Fig.10.4d Enlargement of fig. 10.4b

Fig.10.4e Real diffraction pattern using a mirror

actual optical system, by illumination of a mirror by a truncated Gaussian-shaped light intensity distribution. This pattern is in close agreement with the computer generated ones, here W=250um.

The second test consisted of inserting sinusoidal surface profiles into the object plane. This tests the correctness of the spacing and amplitudes of the diffraction orders. The theoretical spacing equals $F\lambda/D$, where D=200um is the surface wavelength of the sinusoid $y(x)=h\cdot\cos(2\pi x/D)$.

The theoretical amplitudes equal $J_n^2(4\pi h/\lambda)$ (see section 9.6.1.), where h is the amplitude of the profile.

The theoretical spacing between the diffraction orders is 1.175mm, while the computer generated one is 1.19mm. Table 10.1 lists the amplitudes, h, of the surface profiles together with their theoretical and simulated intensity amplitudes. The values in the table are normalised by $I_n=I_n/\sum_{m=-\infty}^{\infty}I_m$. The computer results are shown in figure 10.5a-e. Notice that in figure 10.5e the zero order peak is smaller than the higher order diffraction peaks.

Table 10.1. Theoretical and simulated intensity amplitudes.

amp	theoretical intensity				simulated intensity			
h	I_0	I_1	I_2	I_3	I_0	I_1	I_2	I_3
14	0.98	0.01	0.000	0.000	0.96	0.02	0.000	0.000
35	0.87	0.06	0.000	0.000	0.86	0.07	0.000	0.000
70	0.58	0.19	0.013	0.000	0.57	0.20	0.019	0.000
84	0.45	0.25	0.025	0.000	0.46	0.25	0.025	0.000
140	0.05	0.33	0.124	0.017	0.03	0.33	0.140	0.018

Figure 10.5a.

Figure 10.5b.

Figure 10.5c.

Figure 10.5d.

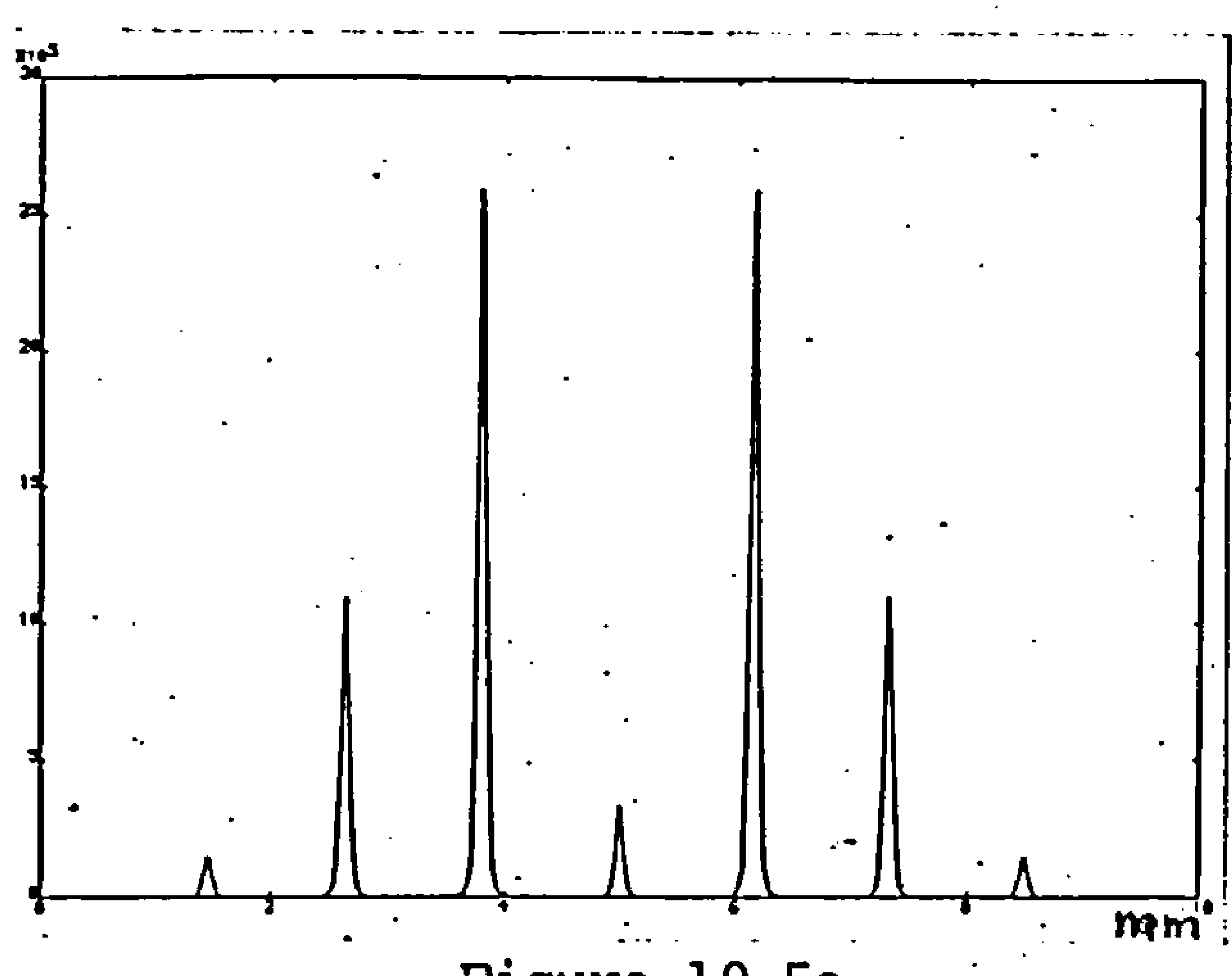


Figure 10.5e.

Fig.10.5 Diffraction patterns by profiles $y(x)=h\cdot\cos(2\pi x/D)$.

Fig.10.5a	Amplitude	h= 14nm
fig.10.5b	Amplitude	h= 35nm
Fig.10.5c	Amplitude	h= 70nm
Fig.10.5d	Amplitude	h= 84nm
Fig.10.5e	Amplitude	h=140nm

From these two tests it can be concluded that the computer simulation program produces results which are in close agreement with theory. The small deviations between the two are mainly caused by the accuracy with which the results can be taken off the plots.

Finally, the program was tested by insertion of actual surface traces of diamond turned drums into the computer program. The computer generated patterns are shown in the figures 10.6a-e. The figures 10.7a-e depict the diffraction patterns obtained by an actual diffraction device. Here an exact one-to-one relationship between the simulated and real diffraction patterns should not be expected, since the computer program uses only a single surface trace (infinitely small width), whereas the real device produces a diffraction pattern caused by a finite area, so it produces a diffraction pattern which can be considered to be an averaged pattern caused by many surface traces. Another practical problem is the impossiblity to use exactly the same area for the surface traces and the optical measurement, since the illumination is in the infra-red, i.e. invisible to the human eye. Visual inspection of the results, however, show a close similarity between the simulated and actual diffraction patterns. The simulated and actual diffraction patterns show an increased number of diffraction orders for the figures c,d and e. This is caused by the increased amplitudes of the corresponding surface profiles. They also exhibit an asymmetry in the diffraction patterns, in grating theory this is referred to as "blazing". This is caused by a skewed position of the tool-tip during the turning process. The figures 10.8a-e show the surface profiles corresponding to the optical spectra.

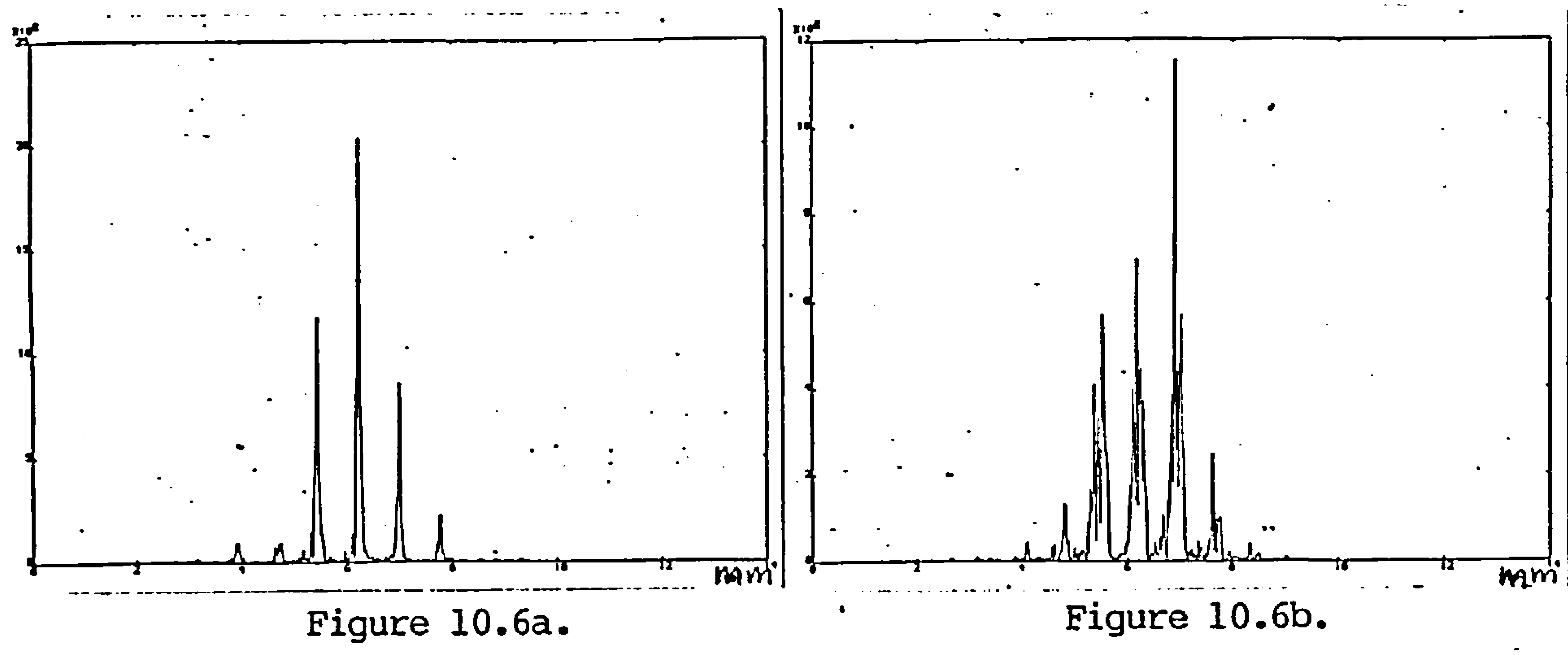


Figure 10.6a.

Figure 10.6b.

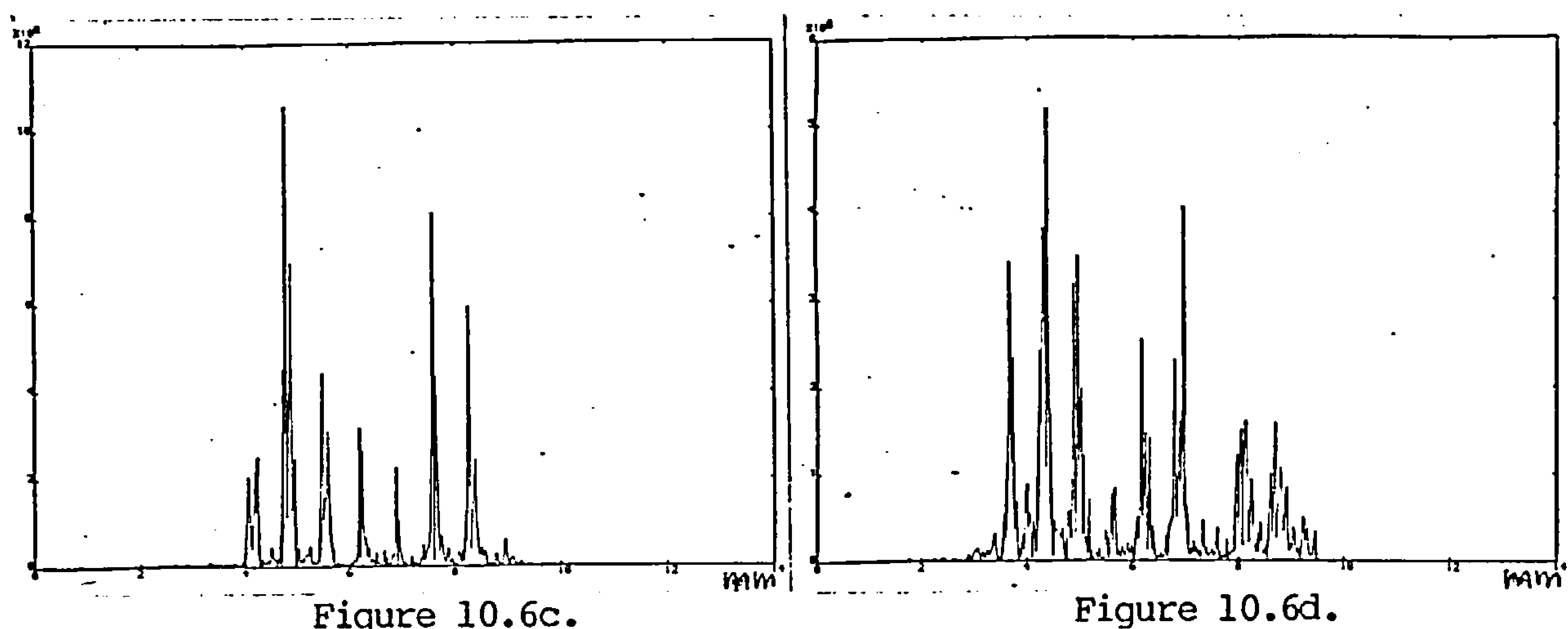


Figure 10.6c.

Figure 10.6d.

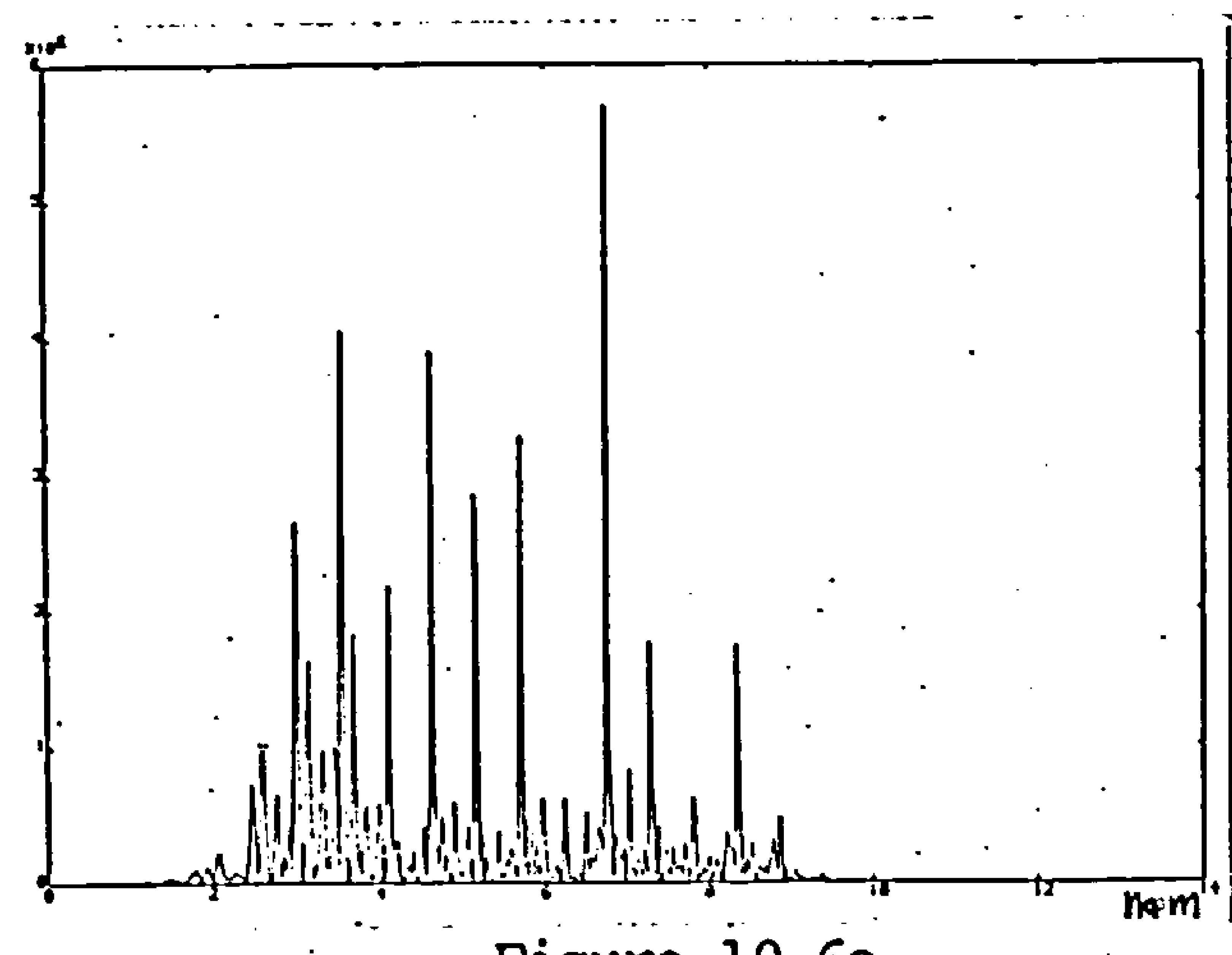


Figure 10.6e.

Fig.10.6 Computer generated diffraction patterns by insertion of real profiles.

Fig.10.6a	Amplitude	h=150nm
fig.10.6b	Amplitude	h=150nm
Fig.10.6c	Amplitude	h=240nm
Fig.10.6d	Amplitude	h=300nm
Fig.10.6e	Amplitude	h=600nm

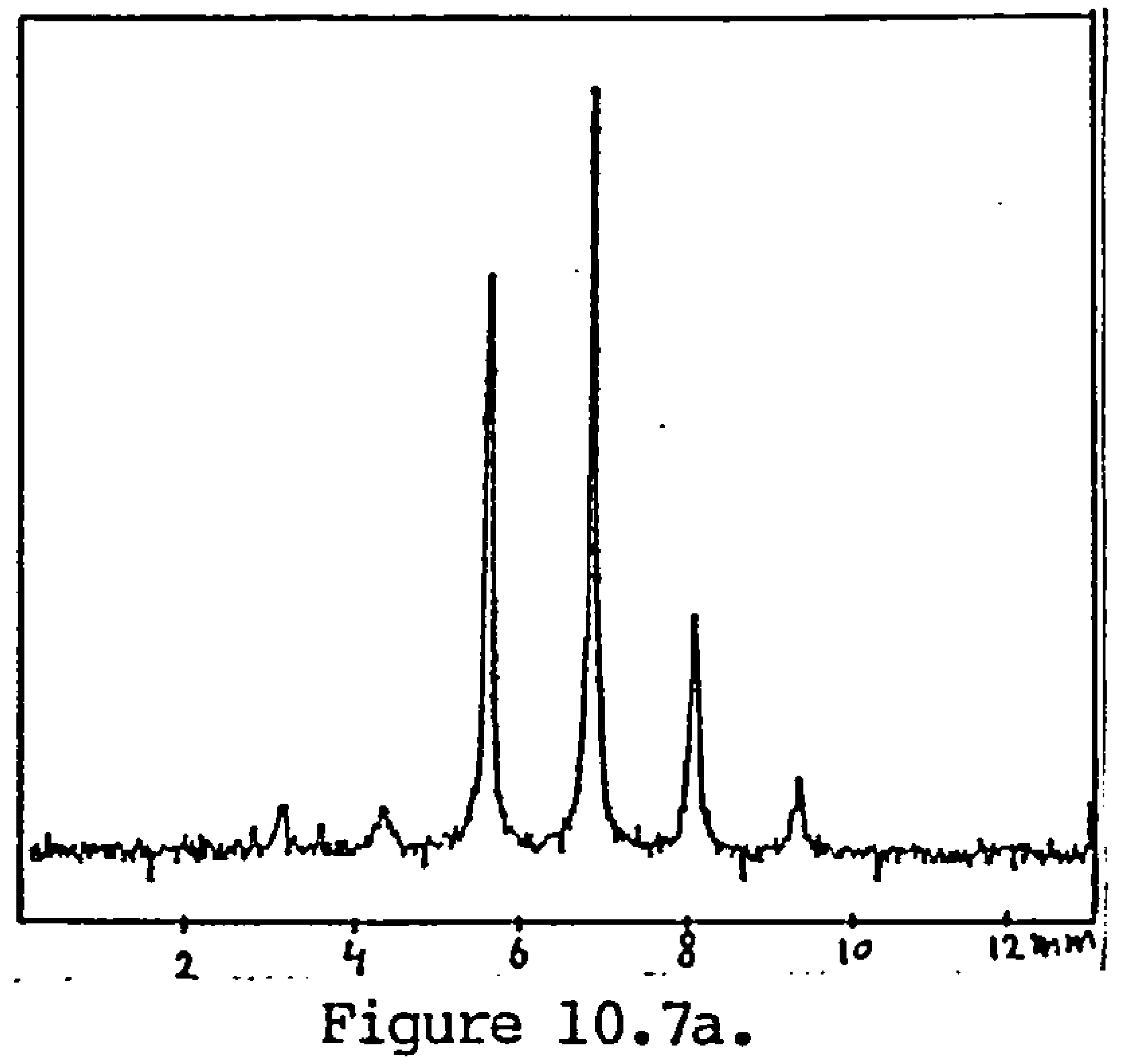


Figure 10.7a.

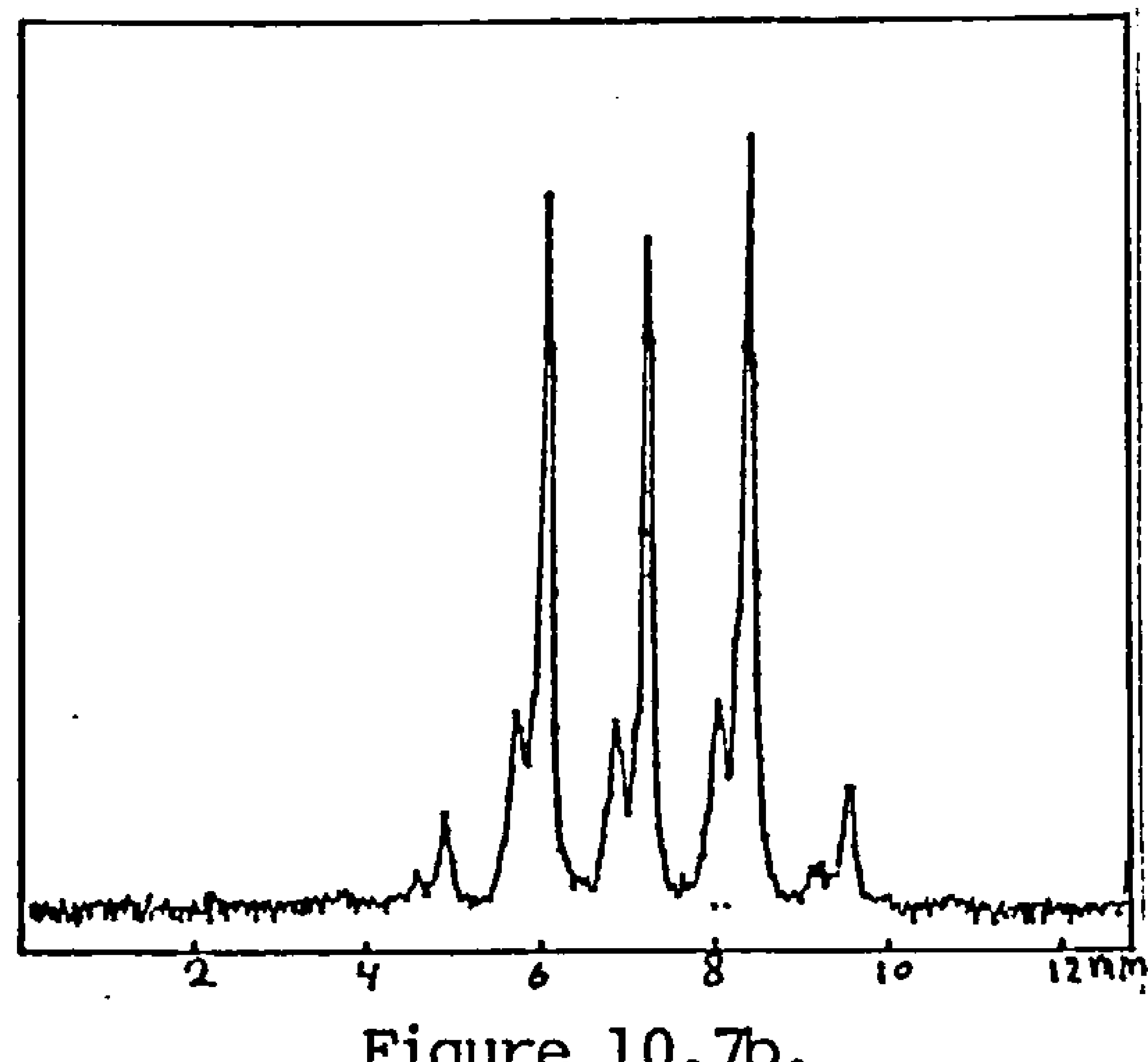


Figure 10.7b.

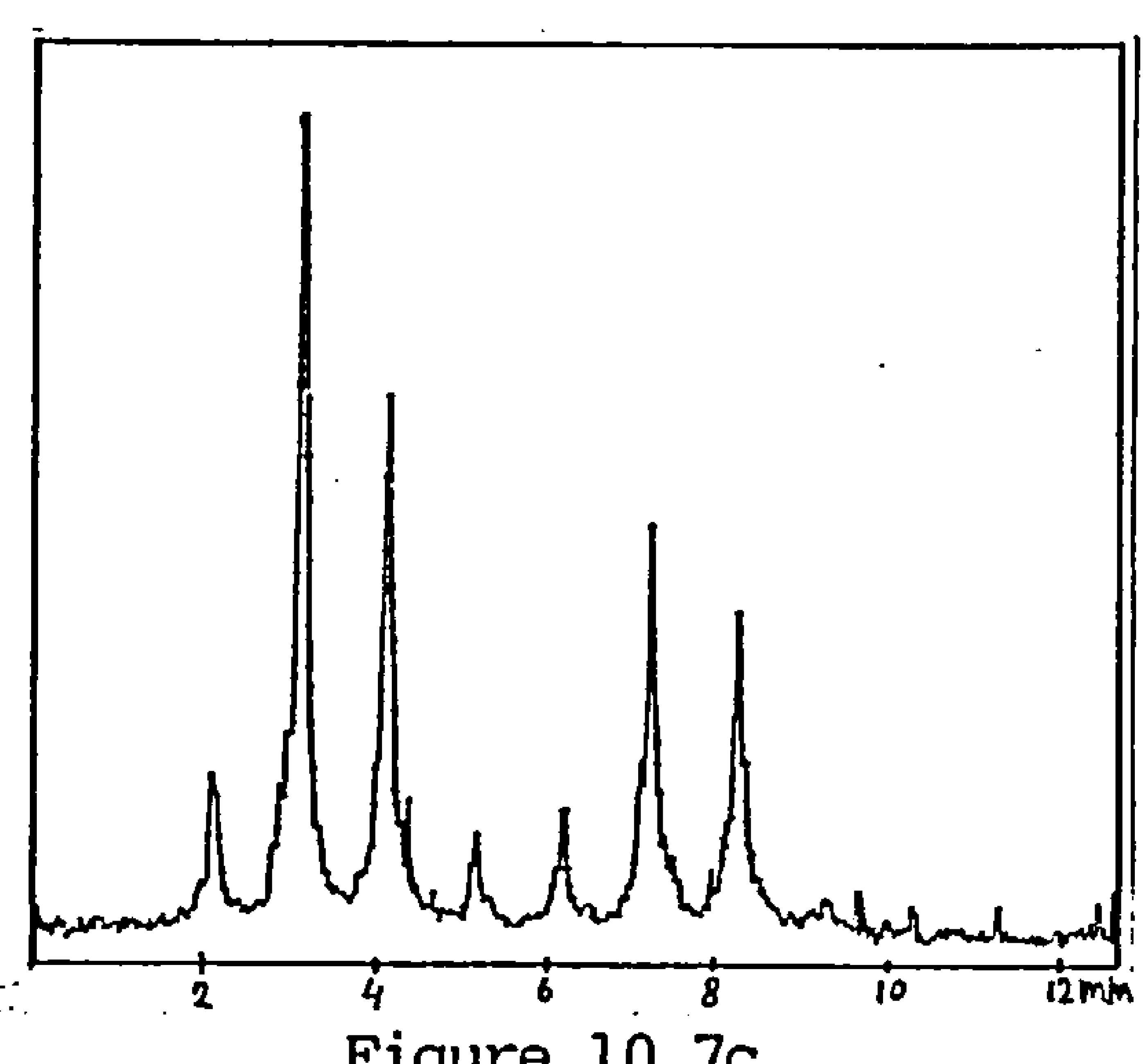


Figure 10.7c.

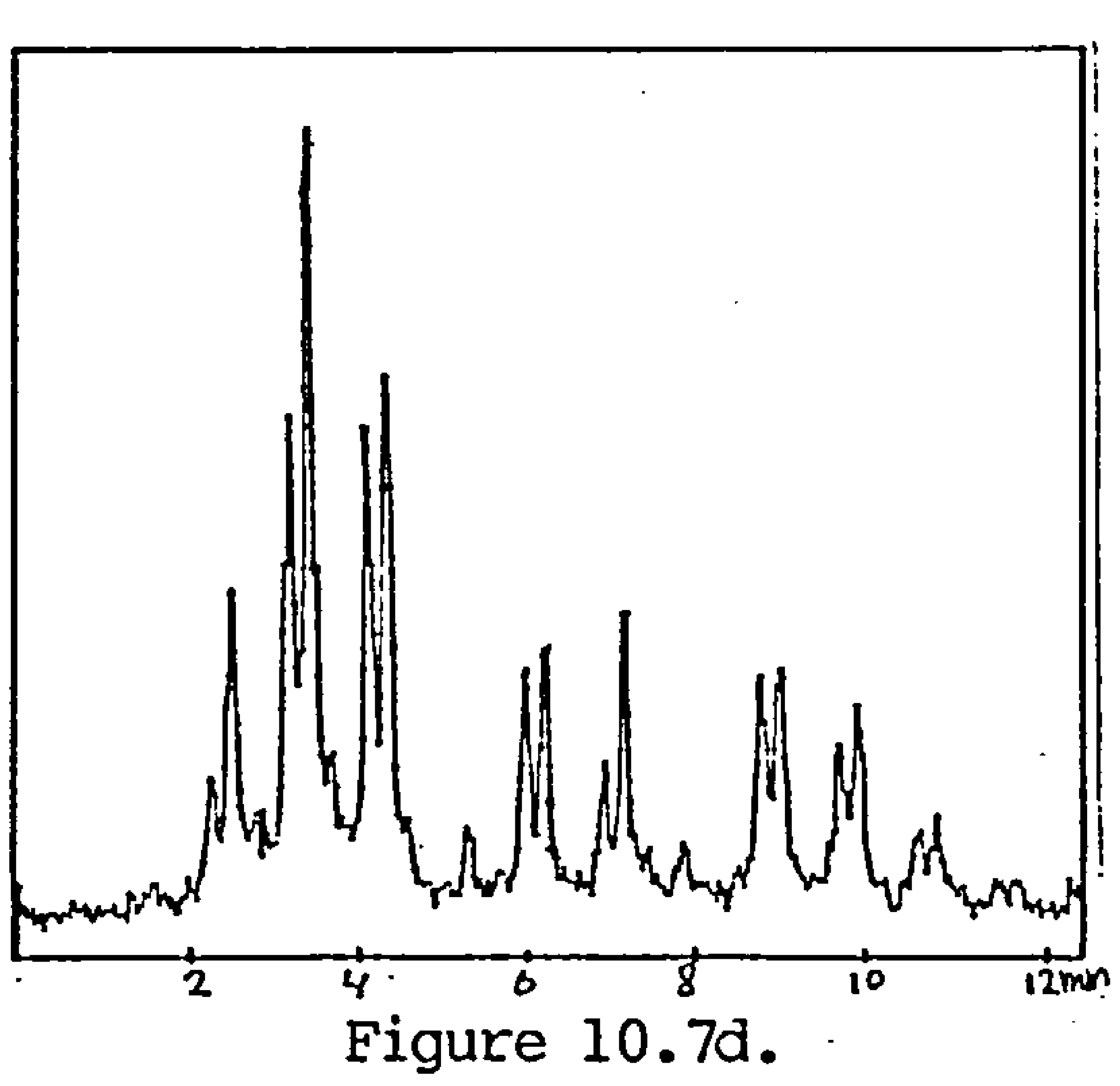


Figure 10.7d.

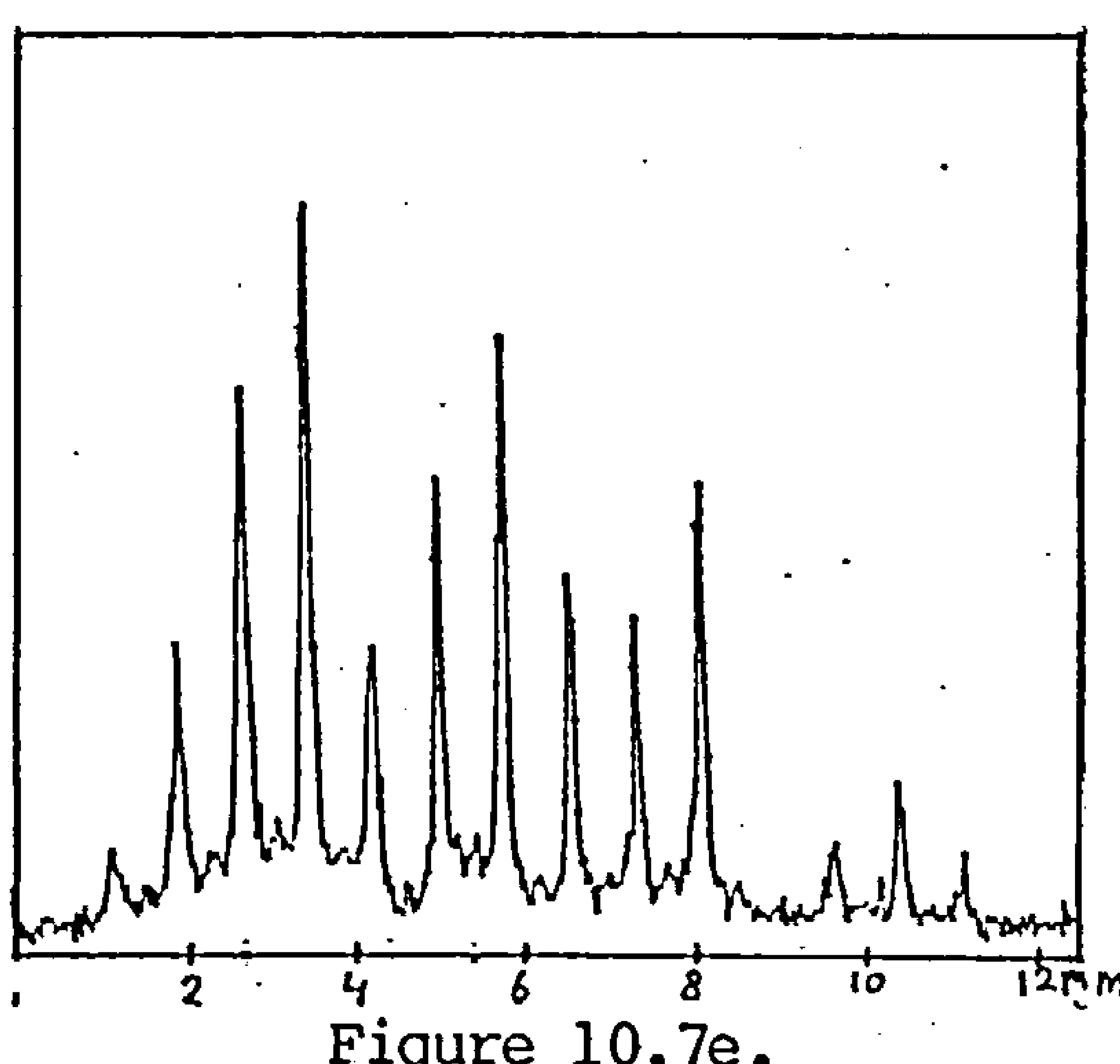


Figure 10.7e.

Fig.10.7 Actual diffraction patterns obtained from diamond turned drums with:

Fig.10.7a	Amplitude	h=150nm
fig.10.7b	Amplitude	h=150nm
Fig.10.7c	Amplitude	h=240nm
Fig.10.7d	Amplitude	h=300nm
Fig.10.7e	Amplitude	h=600nm

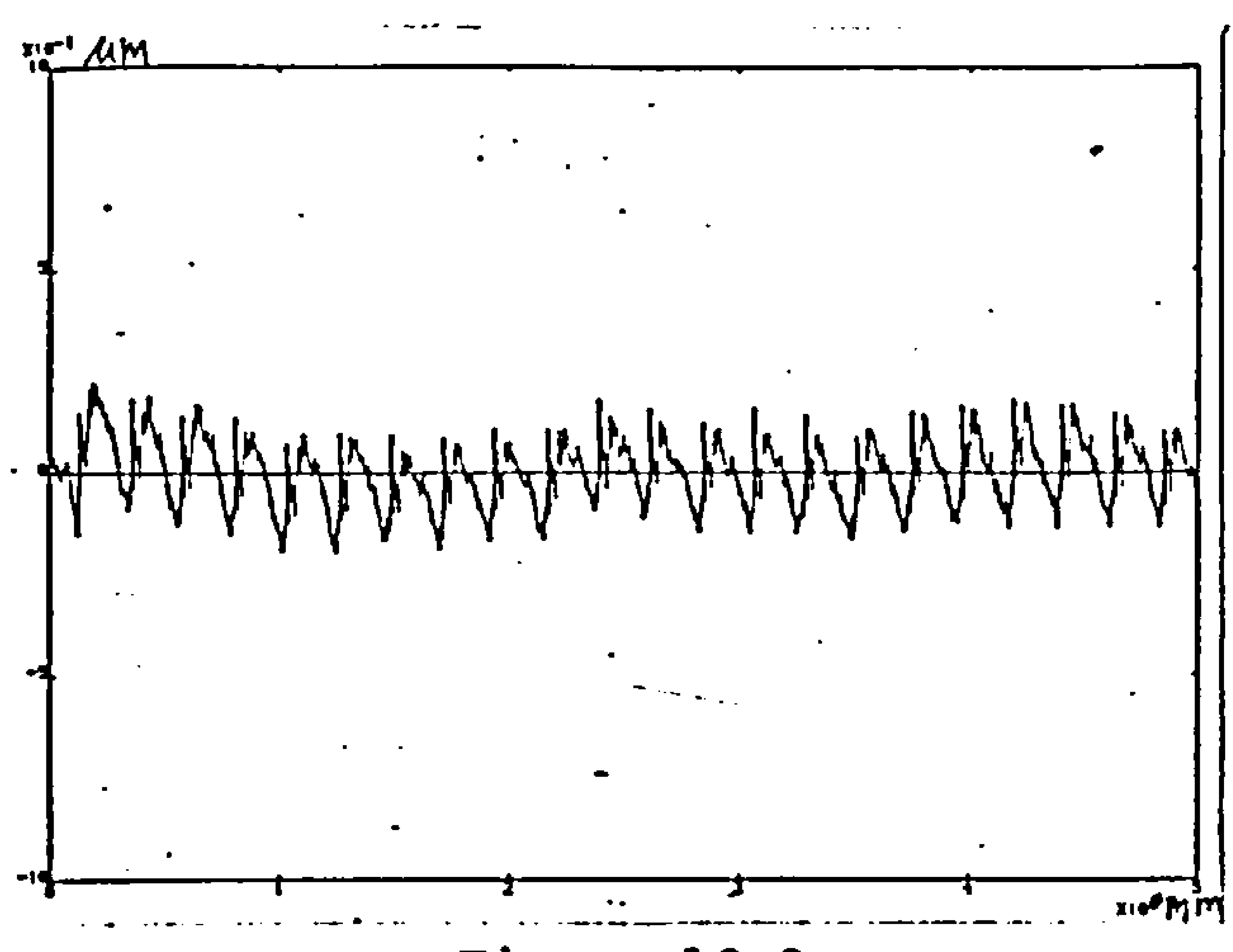


Figure 10.8a.

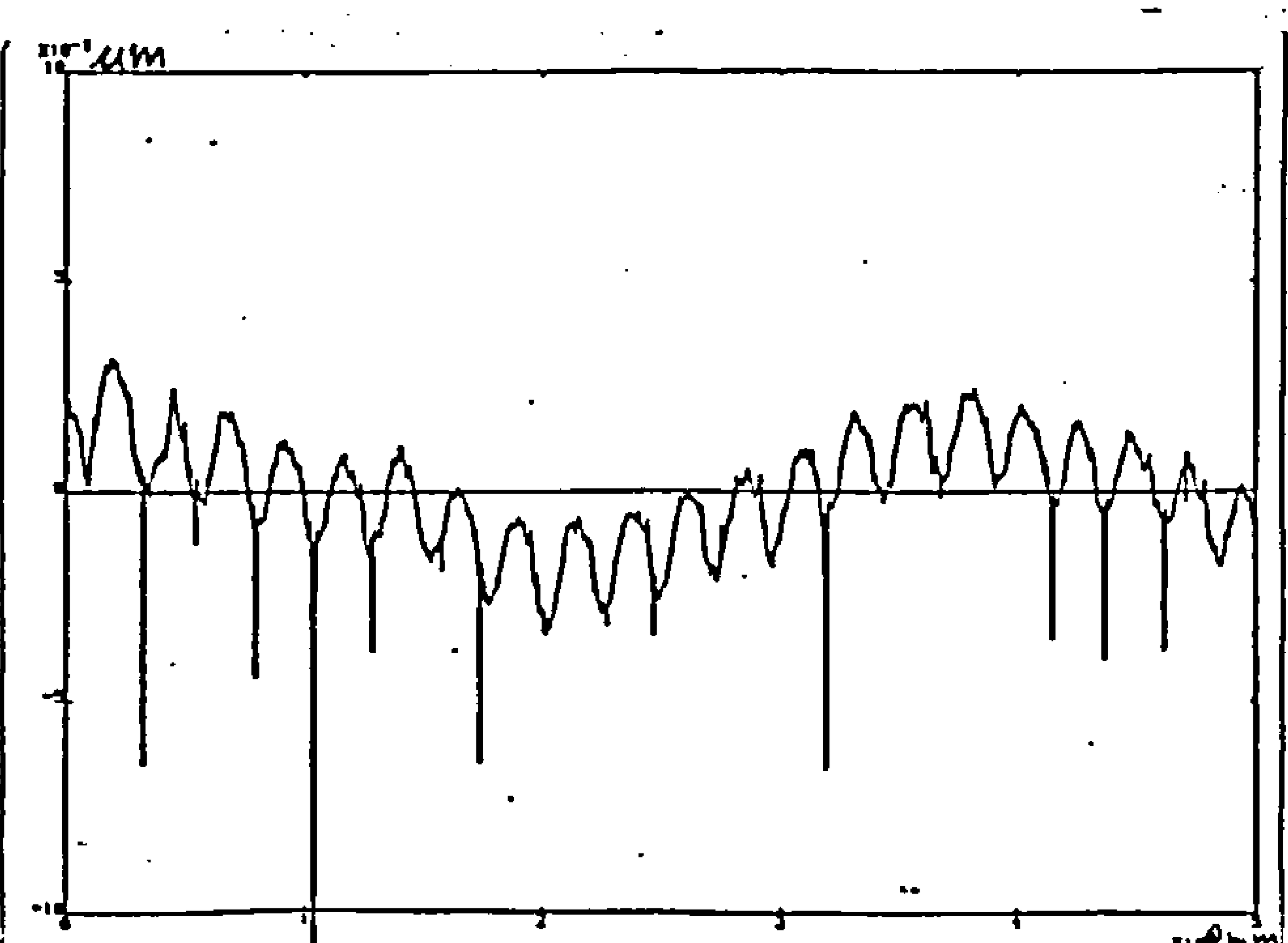


Figure 10.8b.

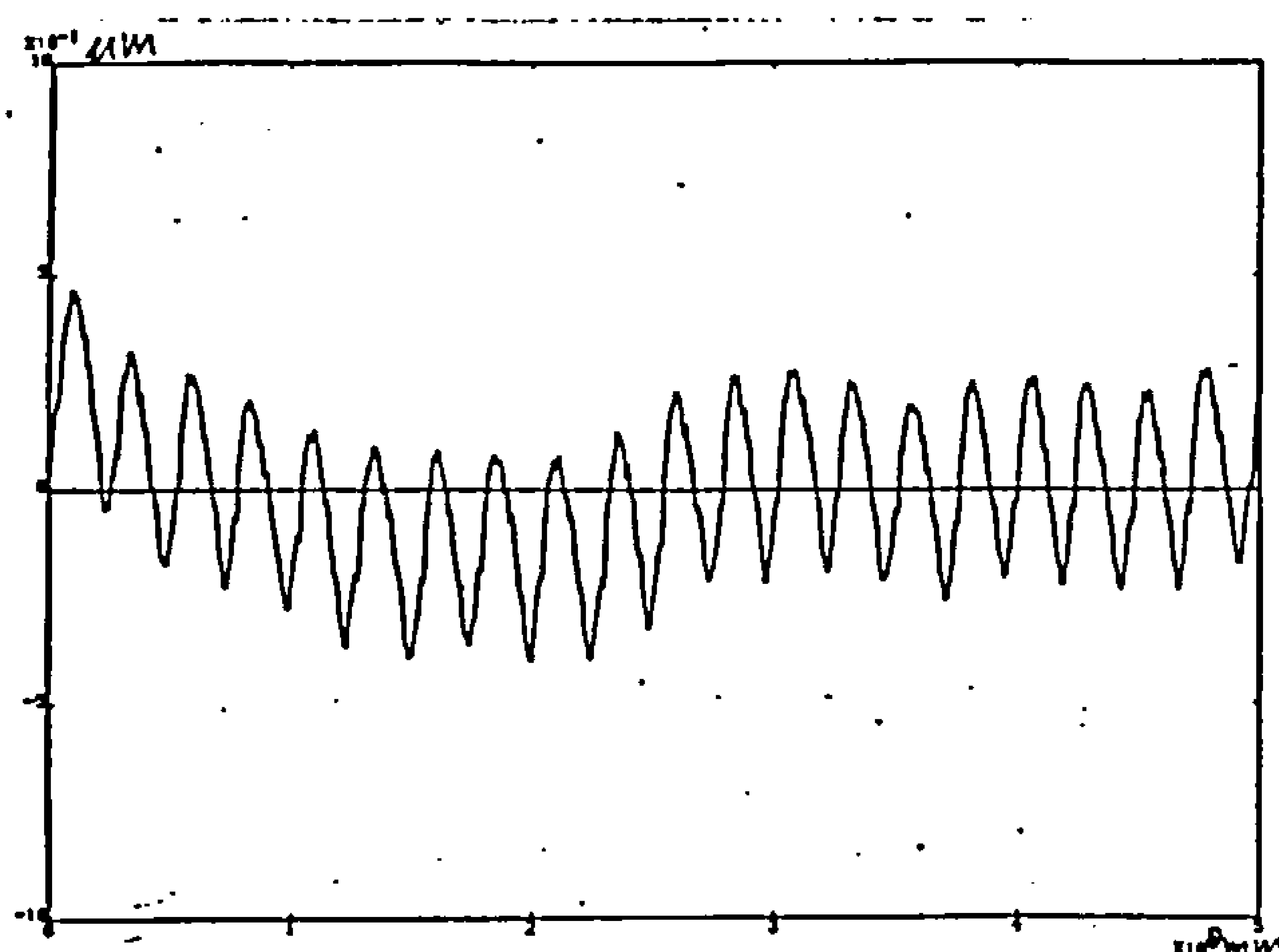


Figure 10.8c.

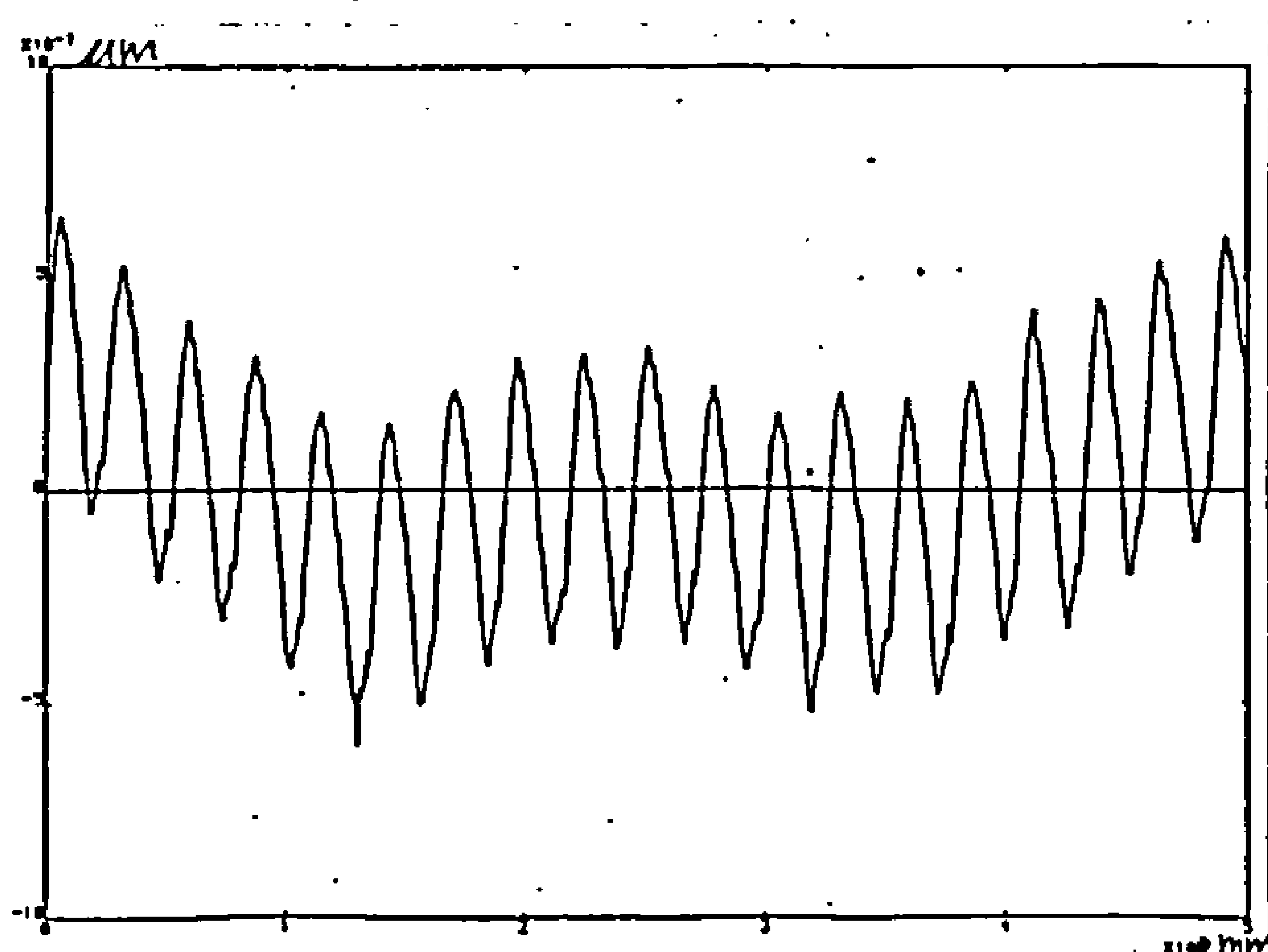


Figure 10.8d.

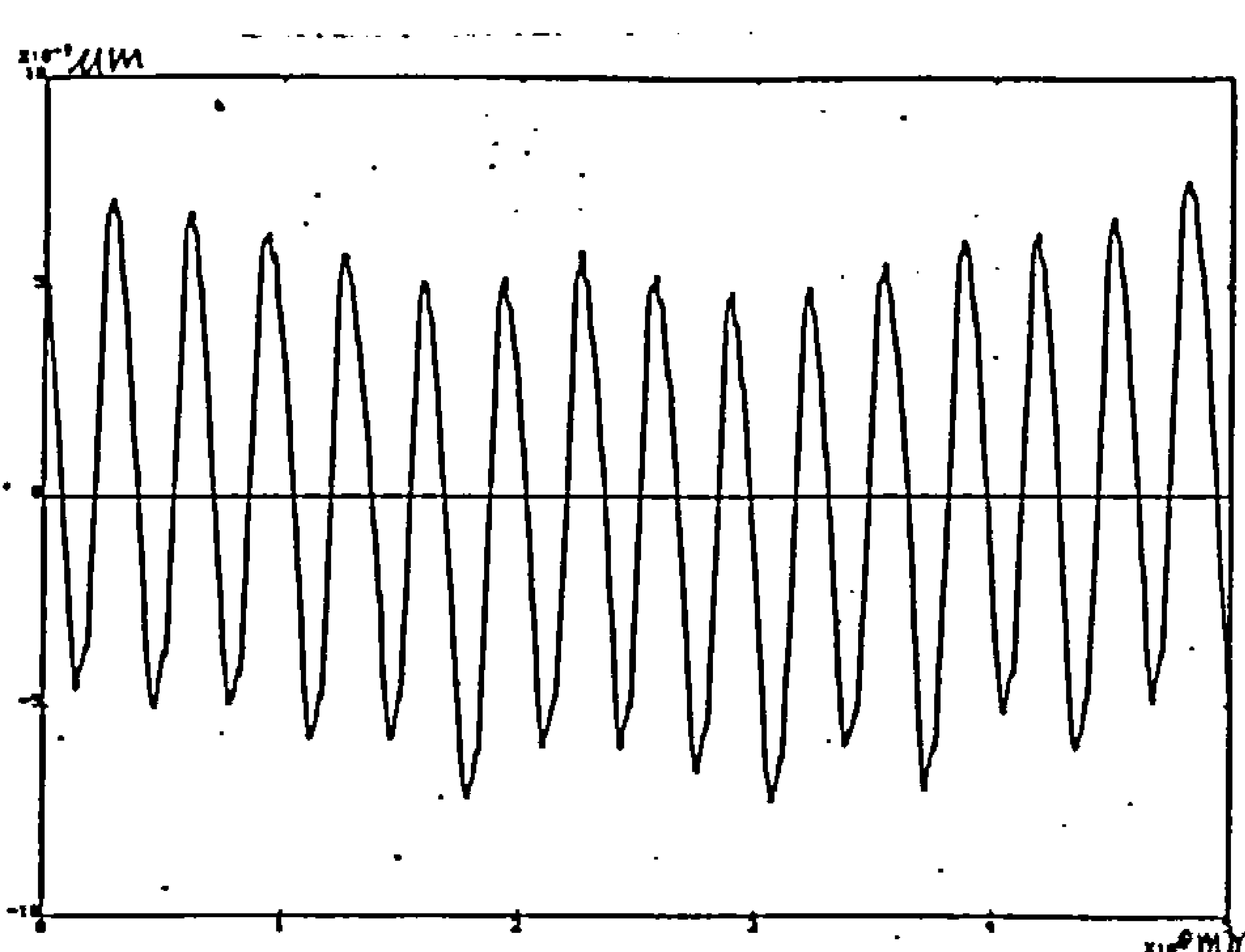


Figure 10.8e.

Fig.10.8 The surface profiles used to produce the actual and simulated optical spectra.

Fig.10.8a Ampli=150nm, feed=25um

fig.10.8b Ampli=150nm, feed=25um

Fig.10.8c Ampli=240nm, feed=25um

Fig.10.8d Ampli=300nm, feed=25um

Fig.10.8e Ampli=600nm, feed=33um

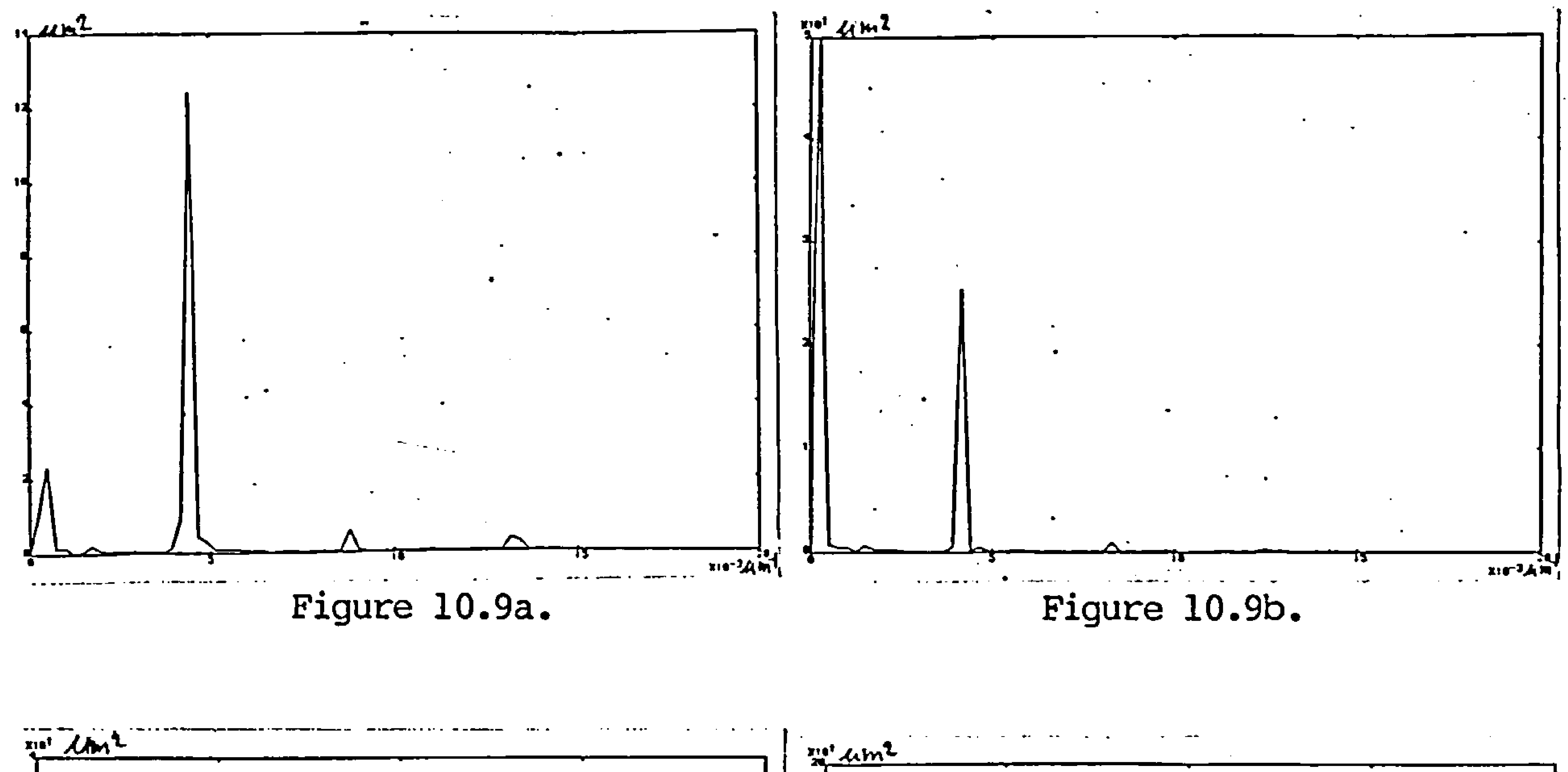


Figure 10.9a. Figure 10.9b.

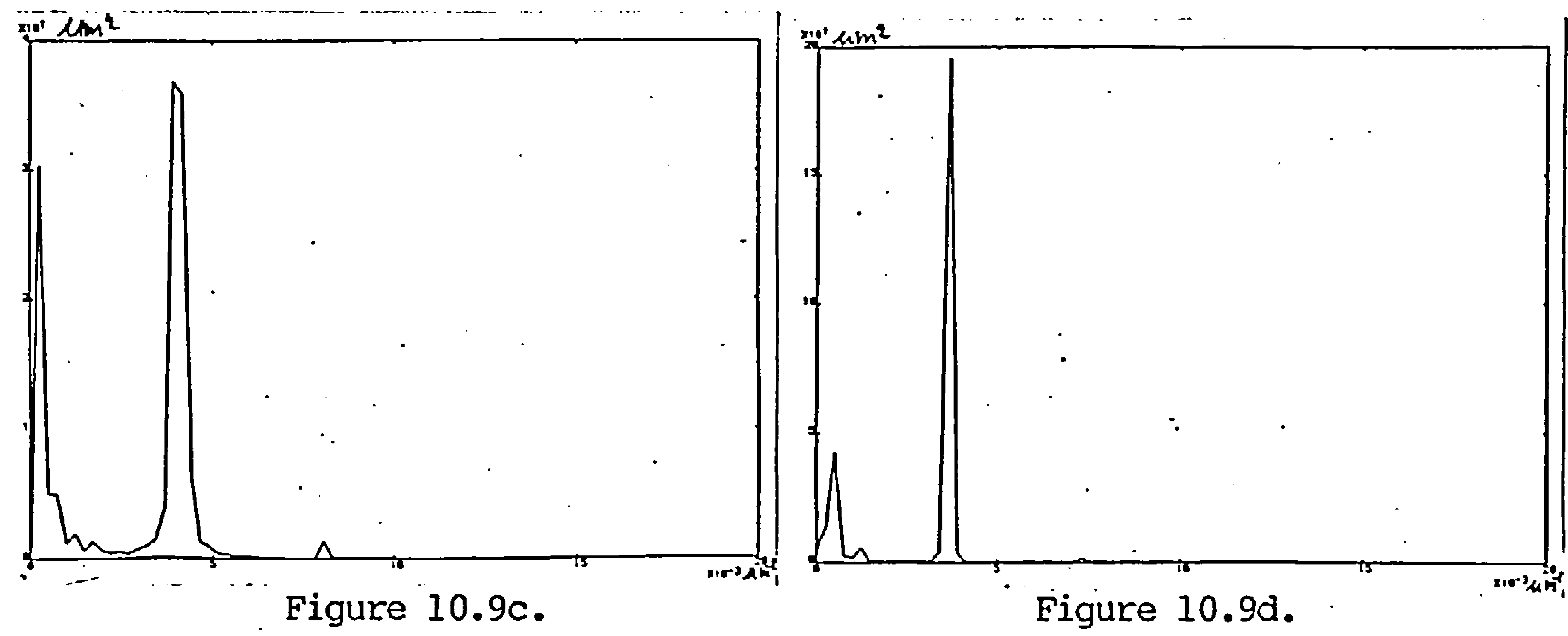


Figure 10.9c. Figure 10.9d.

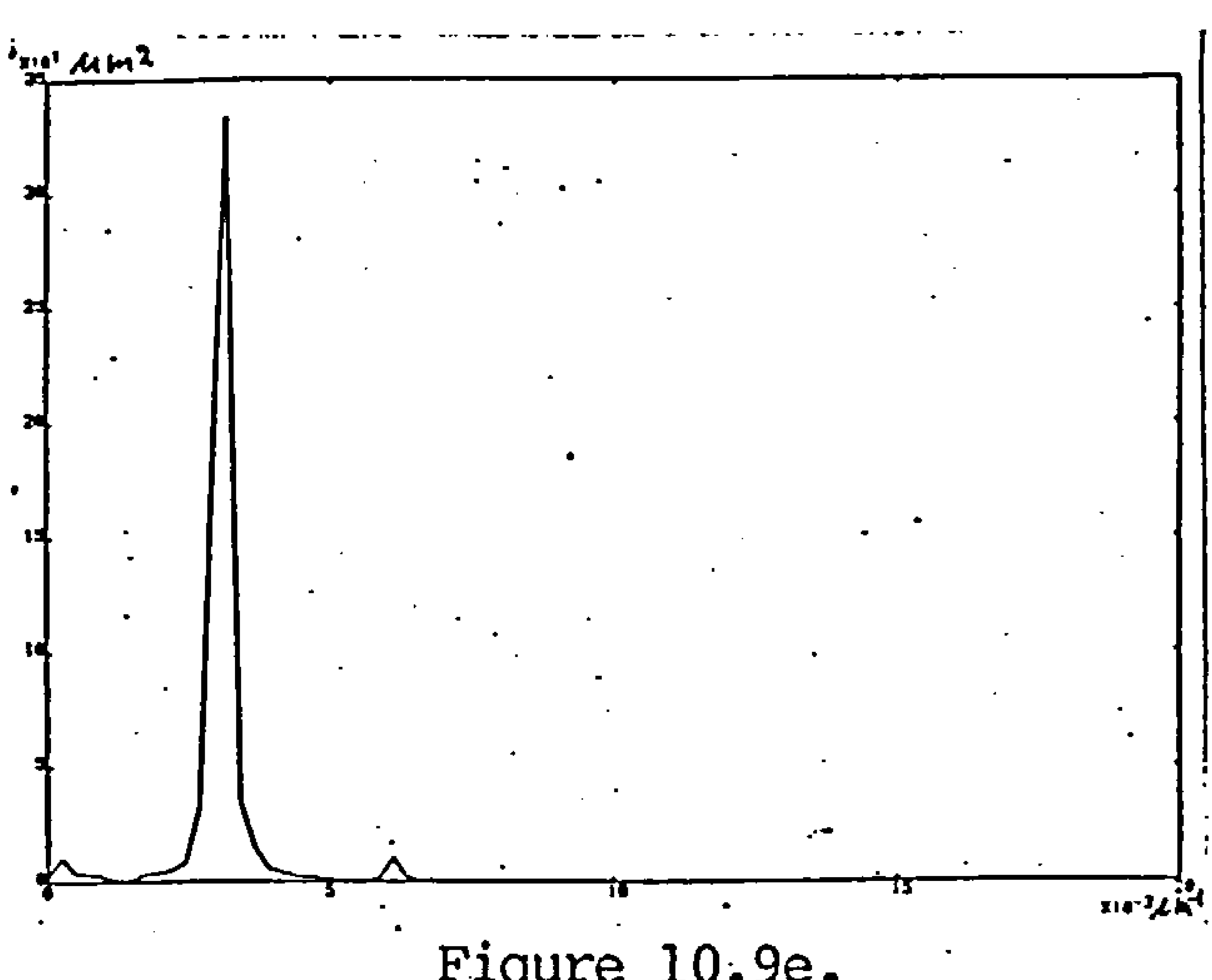


Figure 10.9e.

Fig.10.9 The digitally computed power spectral density functions of the profiles

Fig.10.9a Ampli=150nm, feed=25um

fig.10.9b Ampli=150nm, feed=25um

Fig.10.9c Ampli=240nm, feed=25um

Fig.10.9d Ampli=300nm, feed=25um

Fig.10.9e Ampli=600nm, feed=33um

The figures 10.9a-e show the digitally computed power spectral density functions (PSDF) of these profiles. From these PSDF's it is obvious that most of these surfaces contain a long surface wavelength component (waviness). This manifests itself in the diffraction pattern as a splitting up of the individual diffraction orders into two or more peaks. In diffraction grating theory this is often referred to as "ghosting". This effect can be explained by considering the surface to consist of two profiles whose mean lines meet one another at an angle ′ . This total profile can be expressed by:

$$y(x)=f(x) \qquad 0<x<a/2$$

$$y(x)=f(x)+x\cdot\tan\alpha \qquad a/2<x<a$$

Assuming that the maximum surface wavelength of f(x) and the illumination wavelength are much smaller than "a", then the diffraction pattern consists of a combination of the patterns produced by each half surface profile individually. The pattern produced by the first half can be expressed by:

$$I_1(X) = \left| \int_{-\infty}^{\infty} e^{-i2f(x)\cdot k} \, e^{-\frac{ikXx}{F}} \, dx \right|^2 \tag{10.4a}$$

and the one by the latter half by:

$$I_2(X) = \left| \int_{-\infty}^{\infty} e^{-i2 f(x)\cdot k} \, e^{-\frac{ikx}{F}(X+2F\tan\alpha)} \, dx \right|^2 = I_1(X+2F\tan\alpha) \tag{10.4b}$$

In words, the pattern produced by the latter half is the same as the one produced by the first half, but it is shifted in the diffraction plane by an amount $2F \cdot \tan\alpha$ relative to the position of the first diffraction pattern. The relative amplitudes of each diffraction pattern are proportional to the length of each flat surface profile making up the total profile containing waviness. The two effects, the positional shift in the diffraction plane and the relative amplitudes of the patterns, can be used to quantify the waviness of the surface profile.

10.4 Application of the simulation program.

One of the features of the simulation program is its ability to determine the effects of light scattering from real and simulated surface profiles. The ability to simulate realistic surface profile data offers the advantage of investigating the effects of an optical assessment for a range of controlled surface conditions, as can be produced by malfunction of the machine tools producing these surfaces. As a demonstration, the program has been applied to more or less realistic, periodic surface profiles as produced in turning. The profile 10.10a shows an ideal turned surface profile. The figures 10.10b-e show profiles exhibiting errors in the cutting tool movement. The figures 10.11a-e show their simulated optical diffraction patterns and the figures 10.12a-e their digital PSDF's. Figure 10.10b shows a change of the depth of cut every third revolution , 10.10c a change in the traversing speed of the tool

every third revolution and 10.10d the combination of the before mentioned errors. Both errors produce subharmonics in the diffraction patterns and their spacing is related to a surface profile wavelength equal to 3Xfeed. The change in depth of cut still produces a symmetric pattern but the defect in traversing speed produces an asymmetric pattern. The profiles 10.10a,b,c and d are defined by:

$$h(x)=\min[amp\ x(feed-x), amp(x-feed-horoff)(2feed+horoff-x)+ +veroff, amp(x-2feed)(3feed-x)] \quad (10.4)$$

with $amp=-4c/feed^2$, c=200nm (c is approximately the peak to valley height) and horoff introduces the change in the traversing speed of the tool and veroff the change in the depth of cut.

	feed	horoff	veroff
Fig.10.10a	200um	0.0	0.0
Fig.10.10b	200um	0.2·c	0.0
Fig.10.10c	200um	0.0	0.2·feed
Fig.10.10d	200um	0.2·c	0.2·feed

Profile 10.10e is an amplitude modulated surface profile produced by multiplying profile 10.10a with $(1+0.2\cos(2\pi x/Dm))$, with Dm=1mm (figure 10.13). Its optical and digital spectra shows typical convolution between the original spectra (figures 10.11a and 10.12a) and the power-spectrum which would have been produced by the amplitude modulation function shown in figure 10.14. This convolution exhibits itself by the presence of small sidelobes neighbouring the

main diffraction pattern.

The figures 10.15a-e show the same plots as the figures 10.11a-e, except that here the intensity is plotted on a logarithmic scale. This enhances the smaller intensity amplitudes, thereby accentuating the machine tool defects.

The PSDF's shown in the figures 10.12 confirm that malfunctioning of the machine can only contribute to signals in the low frequency range up to the feed frequency in the PSDF, as predicted in section 5.1., and shown in figure 5.3.

It is obvious that this simulation program can be exploited to produce an atlas of optical diffraction spectra related to certain machine tool defects. In this way the optical roughness sensor can be exploited for the monitoring of machine tool behaviour.

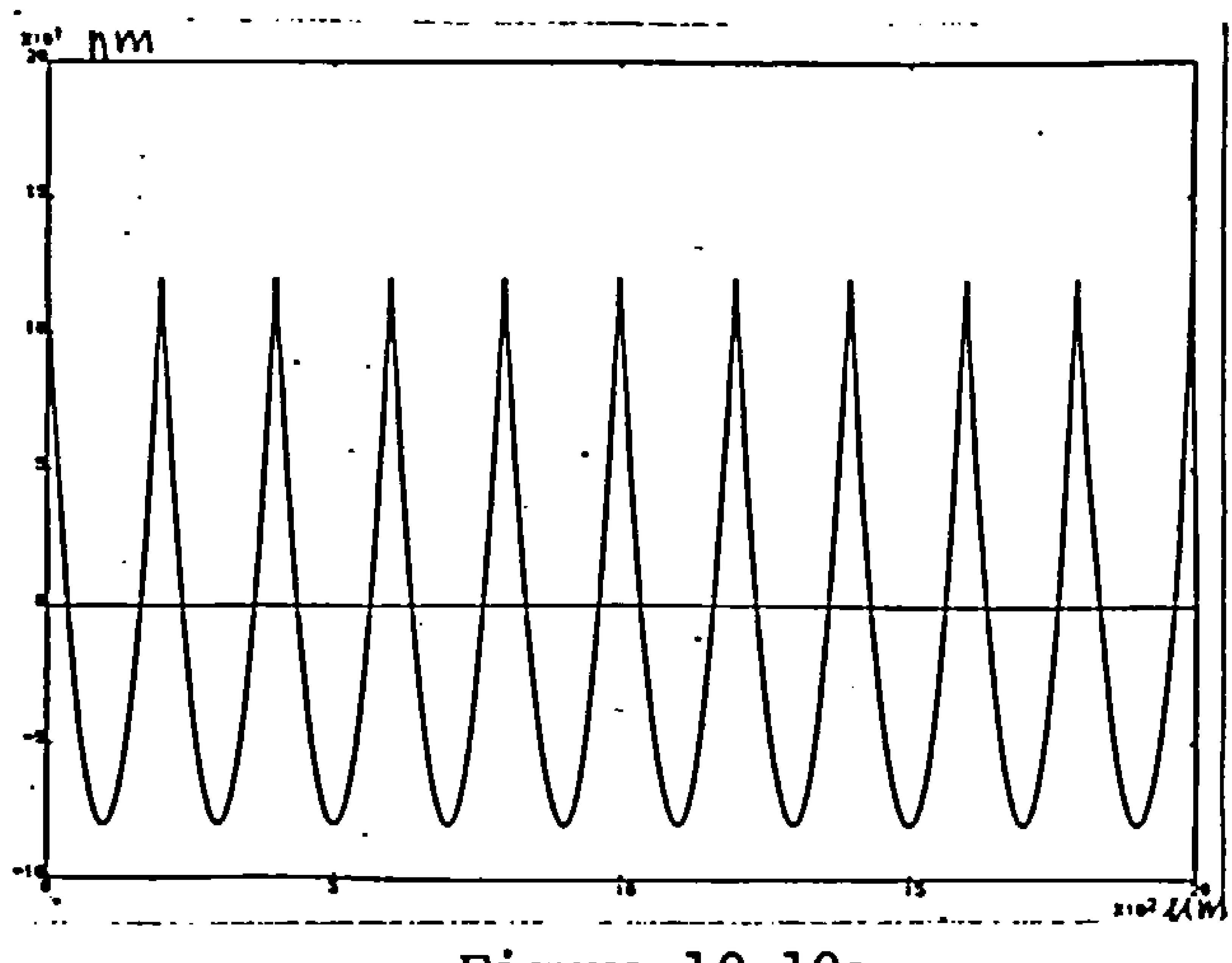


Figure 10.10a.

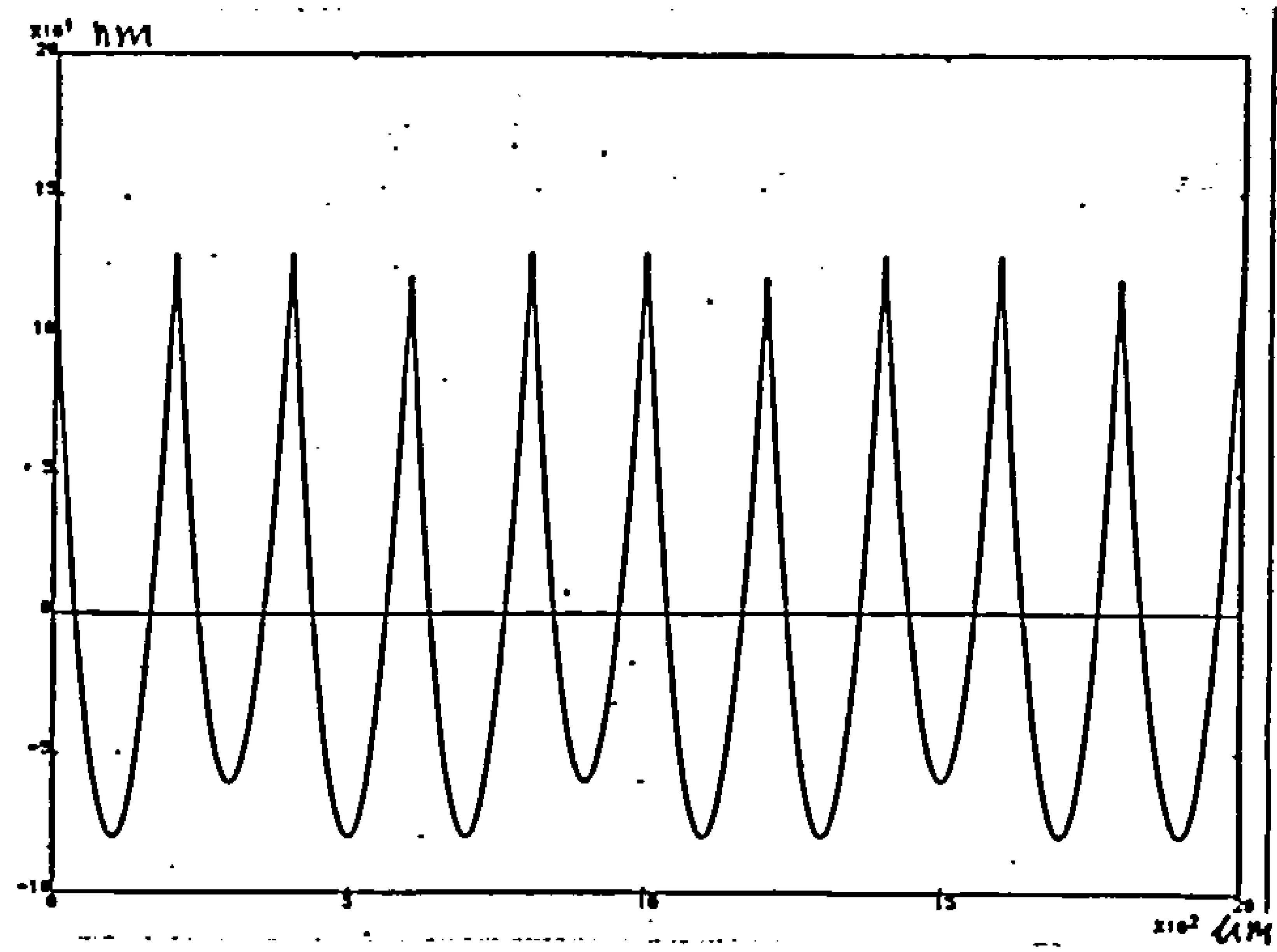


Figure 10.10b.

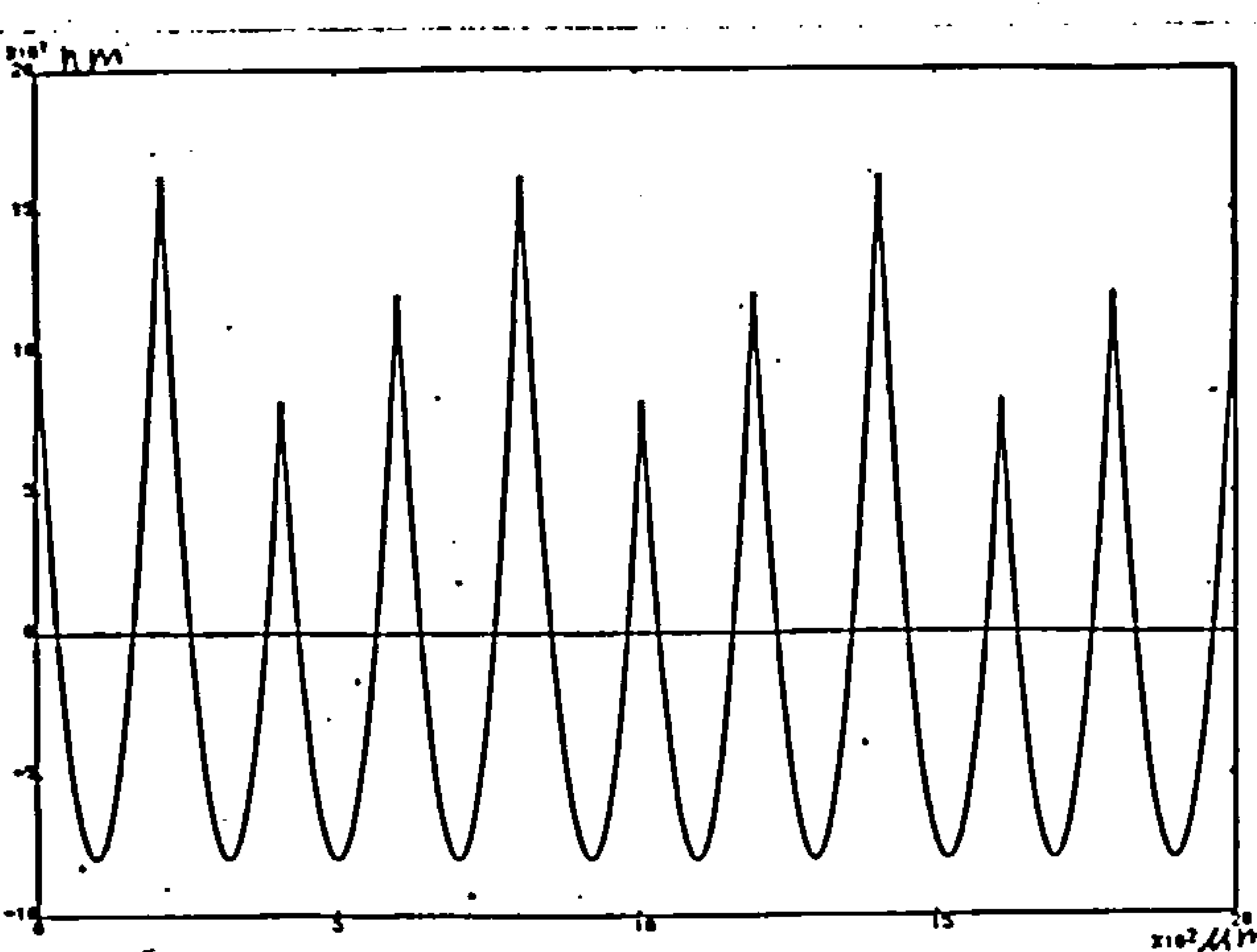


Figure 10.10c.

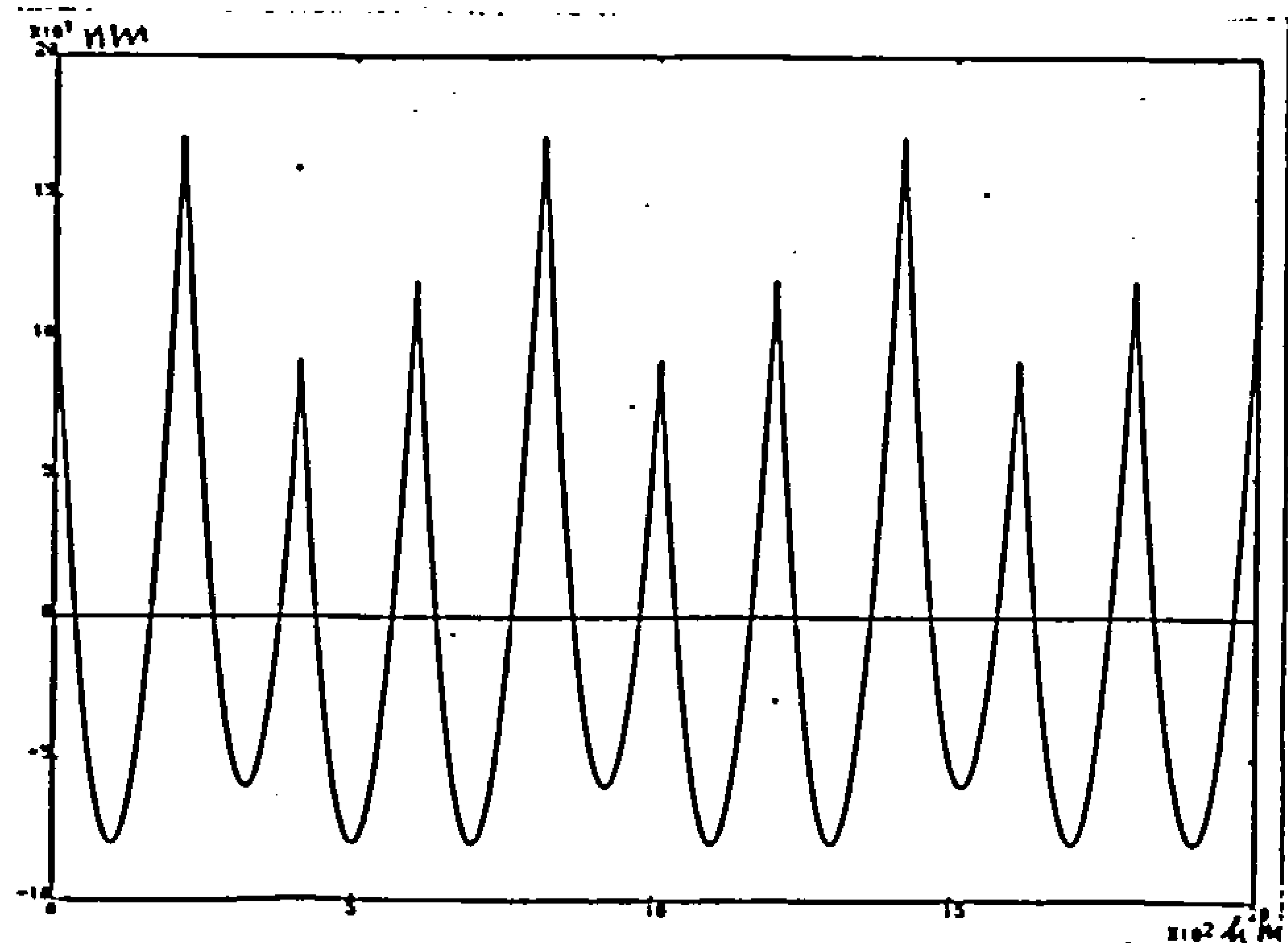


Figure 10.10d.

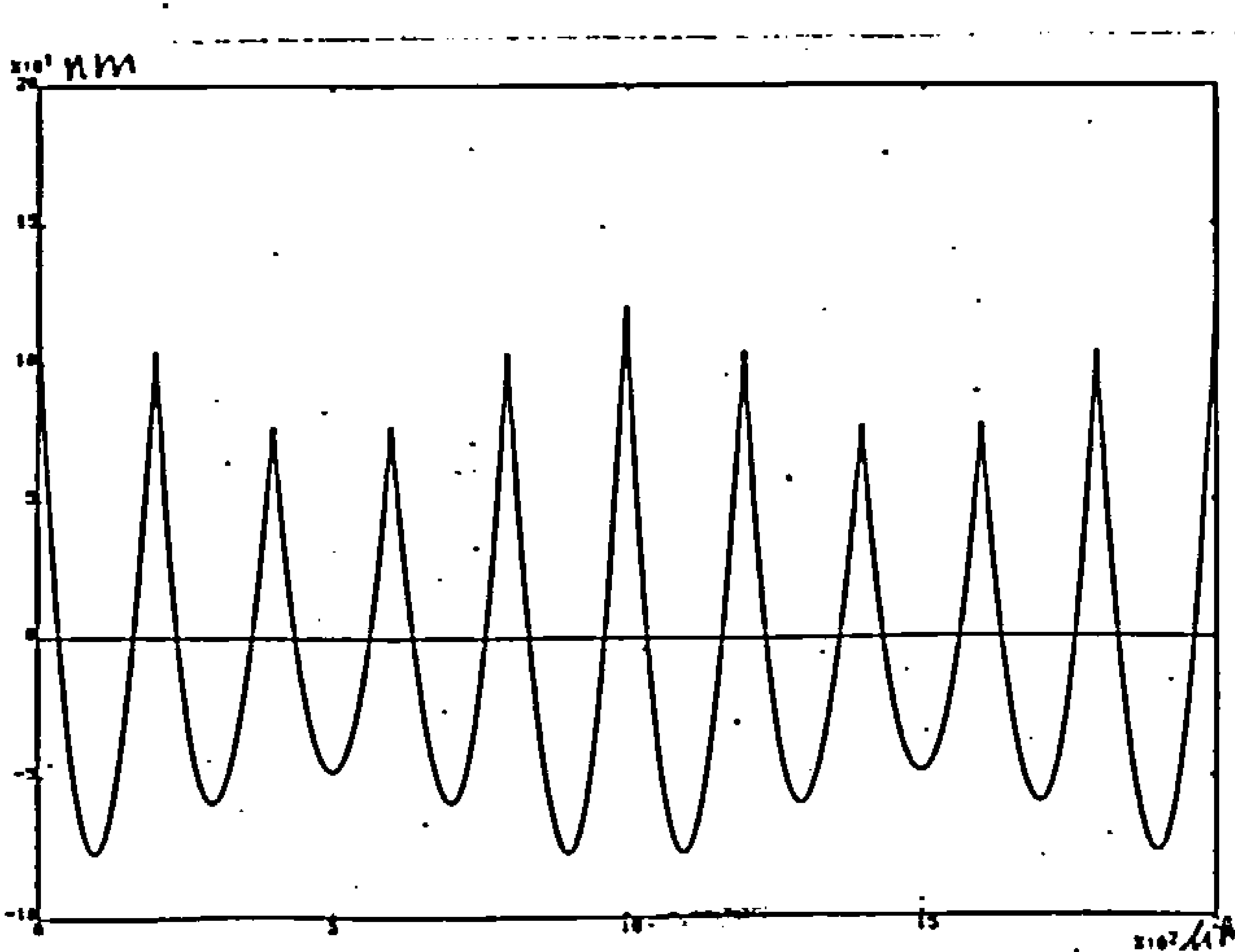


Figure 10.10e.

Fig.10.10 Artificial surface profiles exhibiting certain periodic machining errors.

Fig.10.10a Perfect profile.

fig.10.10b Change of depth of cut.

Fig.10.10c Change of traverse.

Fig.10.10d Combination of above.

Fig.10.10e Amplitude modulation.

Figure 10.11a.

Figure 10.11b.

Figure 10.11c.

Figure 10.11d.

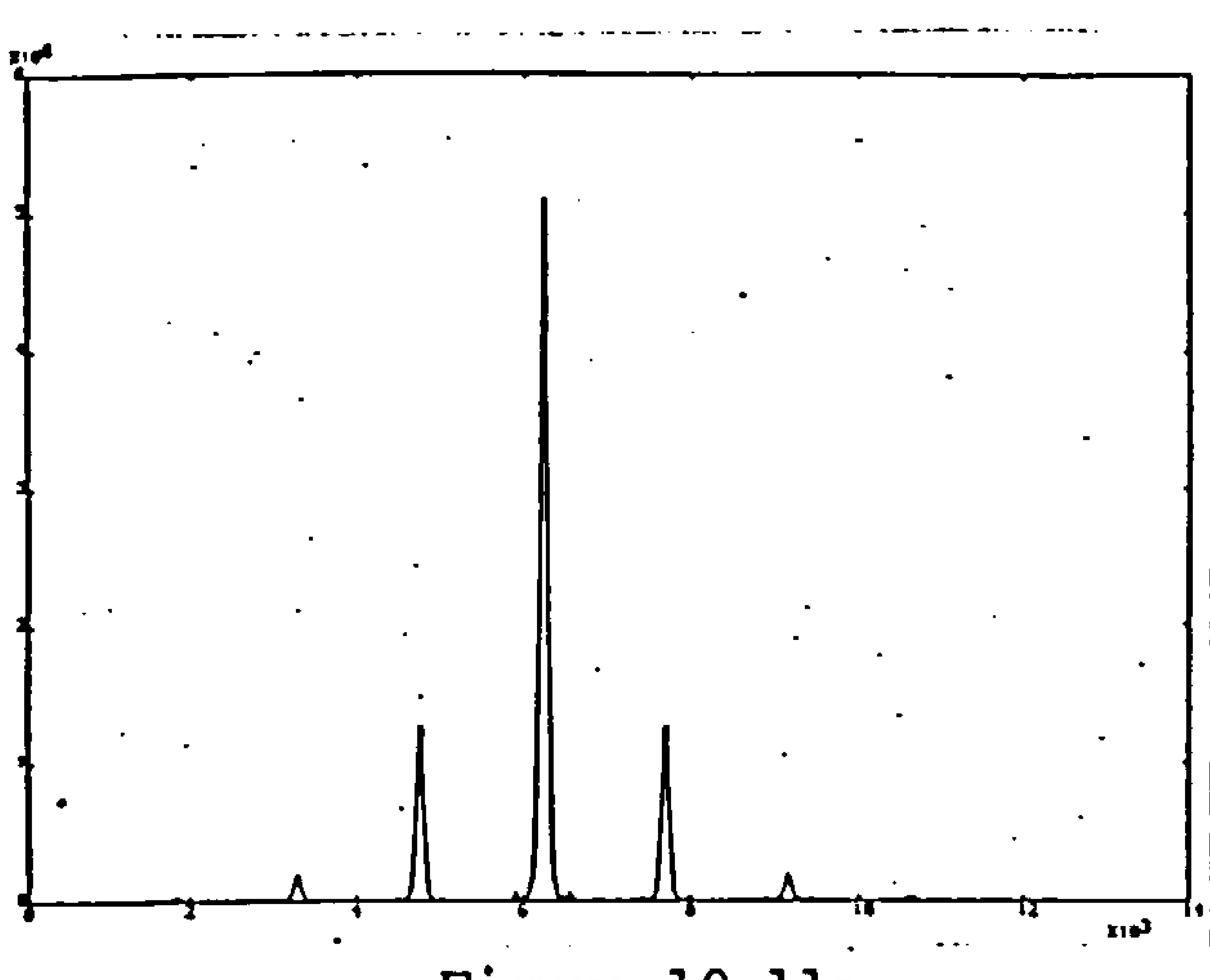


Figure 10.11e.

Fig.10.11 Simulated diffraction patterns of surface profiles with certain periodic machining errors

Fig.10.11a Perfect profile.

fig.10.11b Change of depth of cut.

Fig.10.11c Change of traverse.

Fig.10.11d Combination of above.

Fig.10.11e Amplitude modulation.

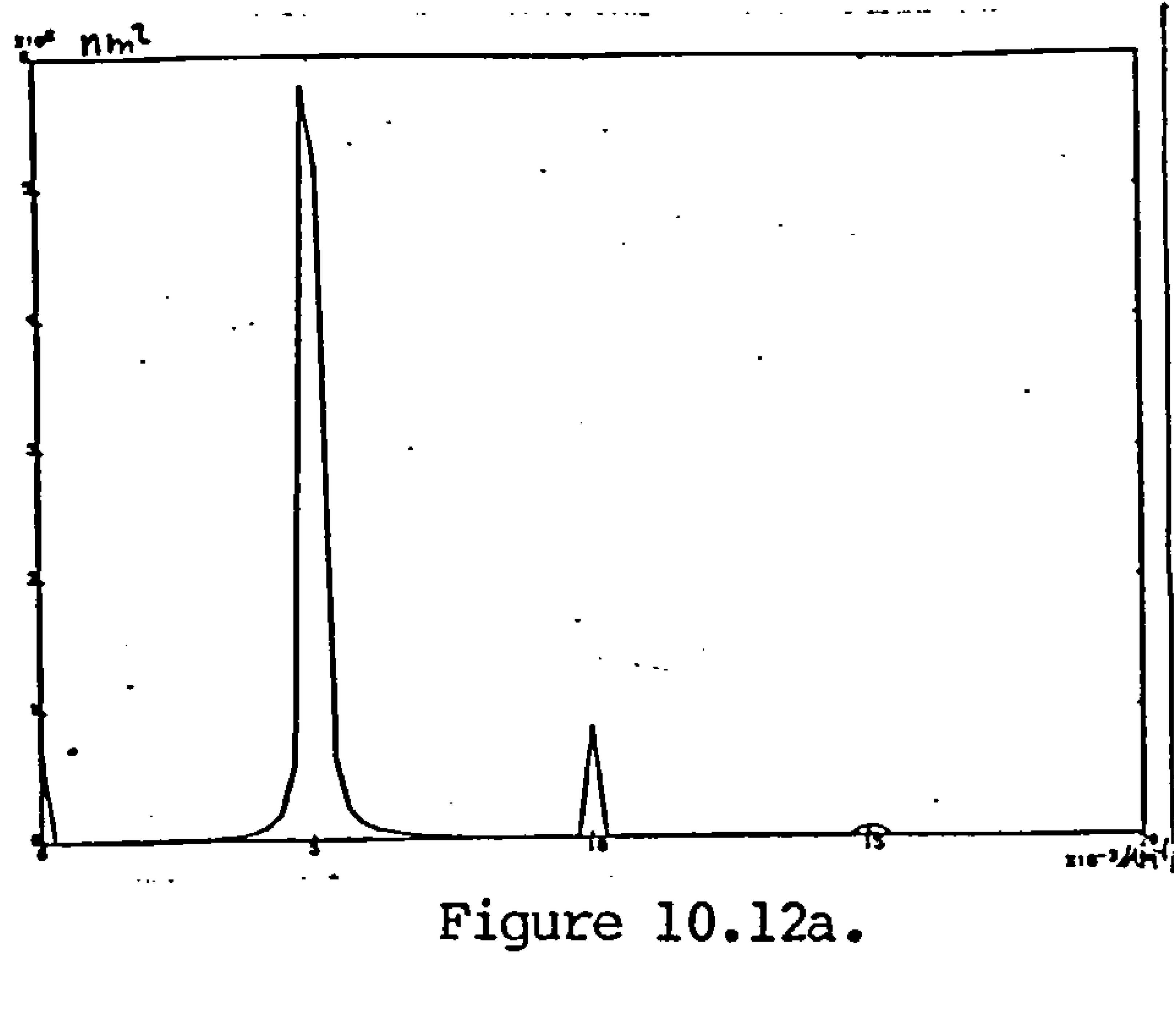


Figure 10.12a.

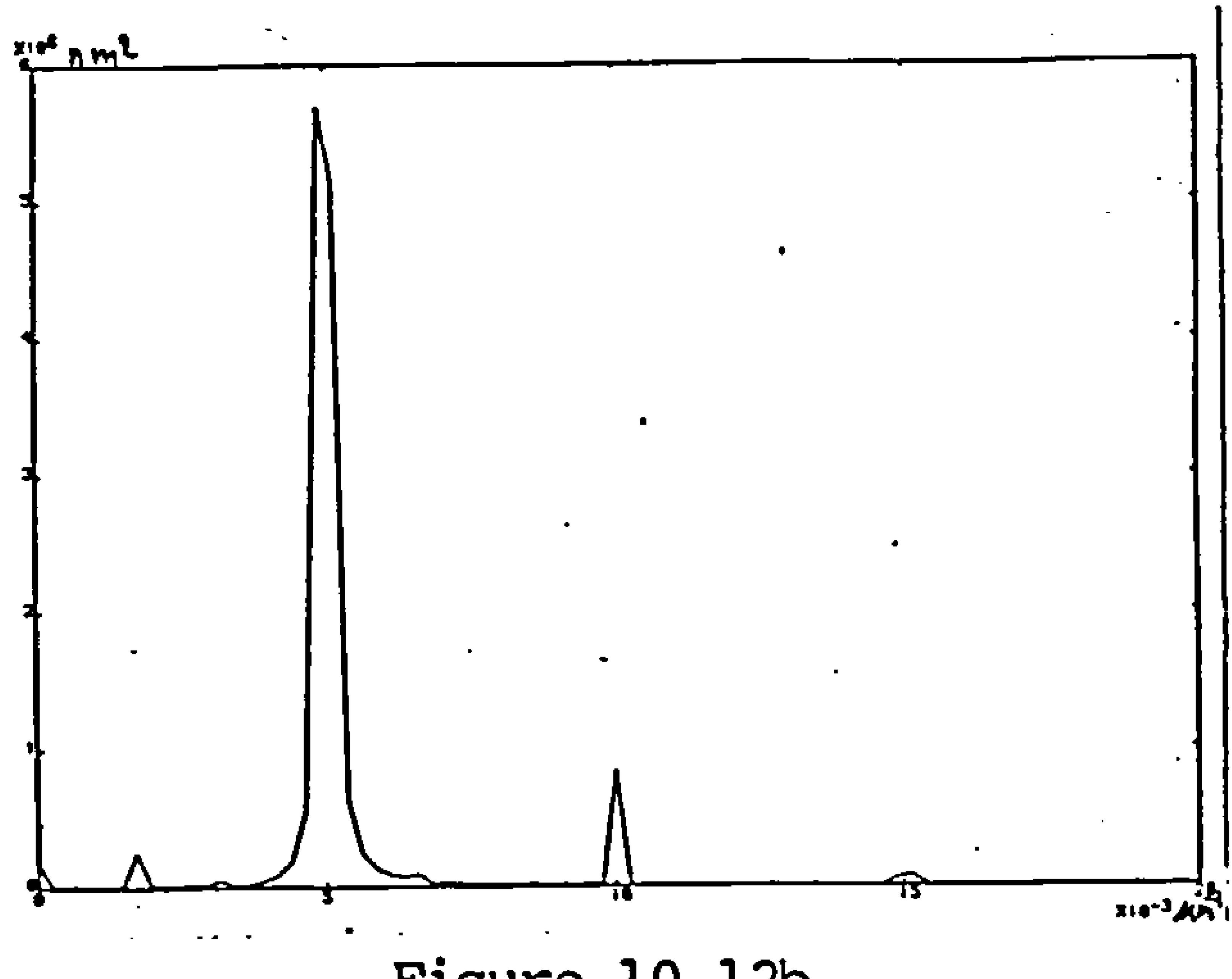


Figure 10.12b.

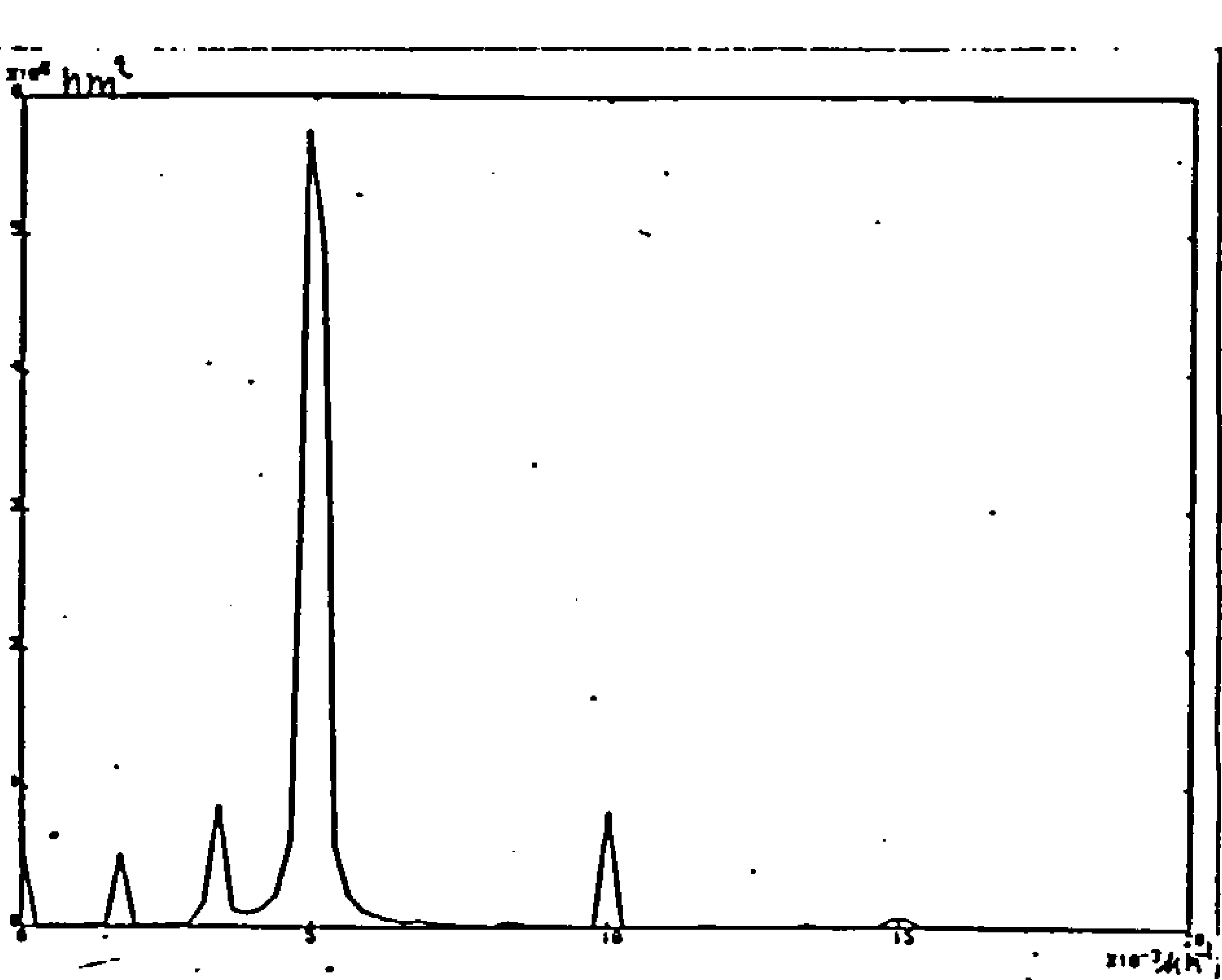


Figure 10.12c.

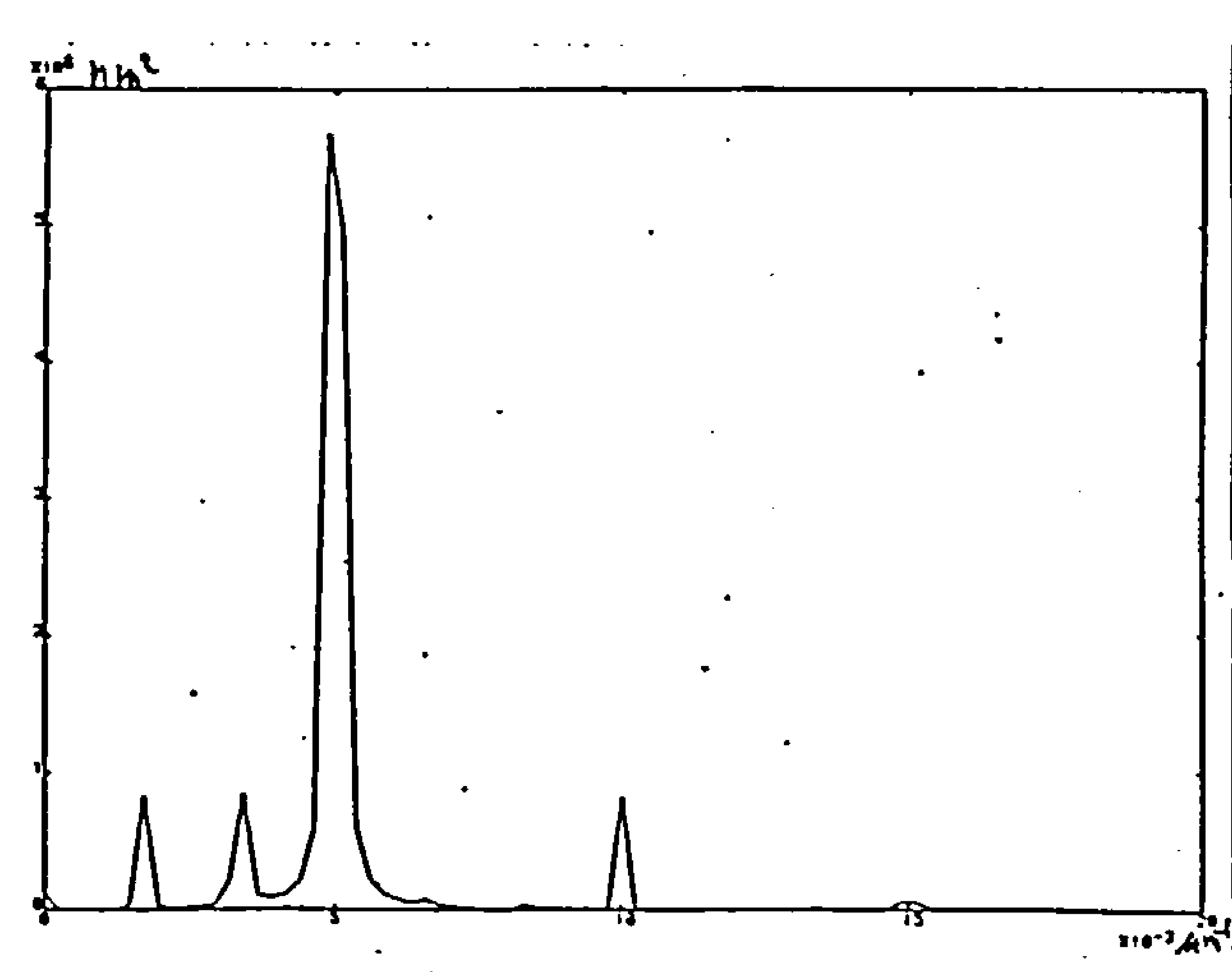


Figure 10.12d.

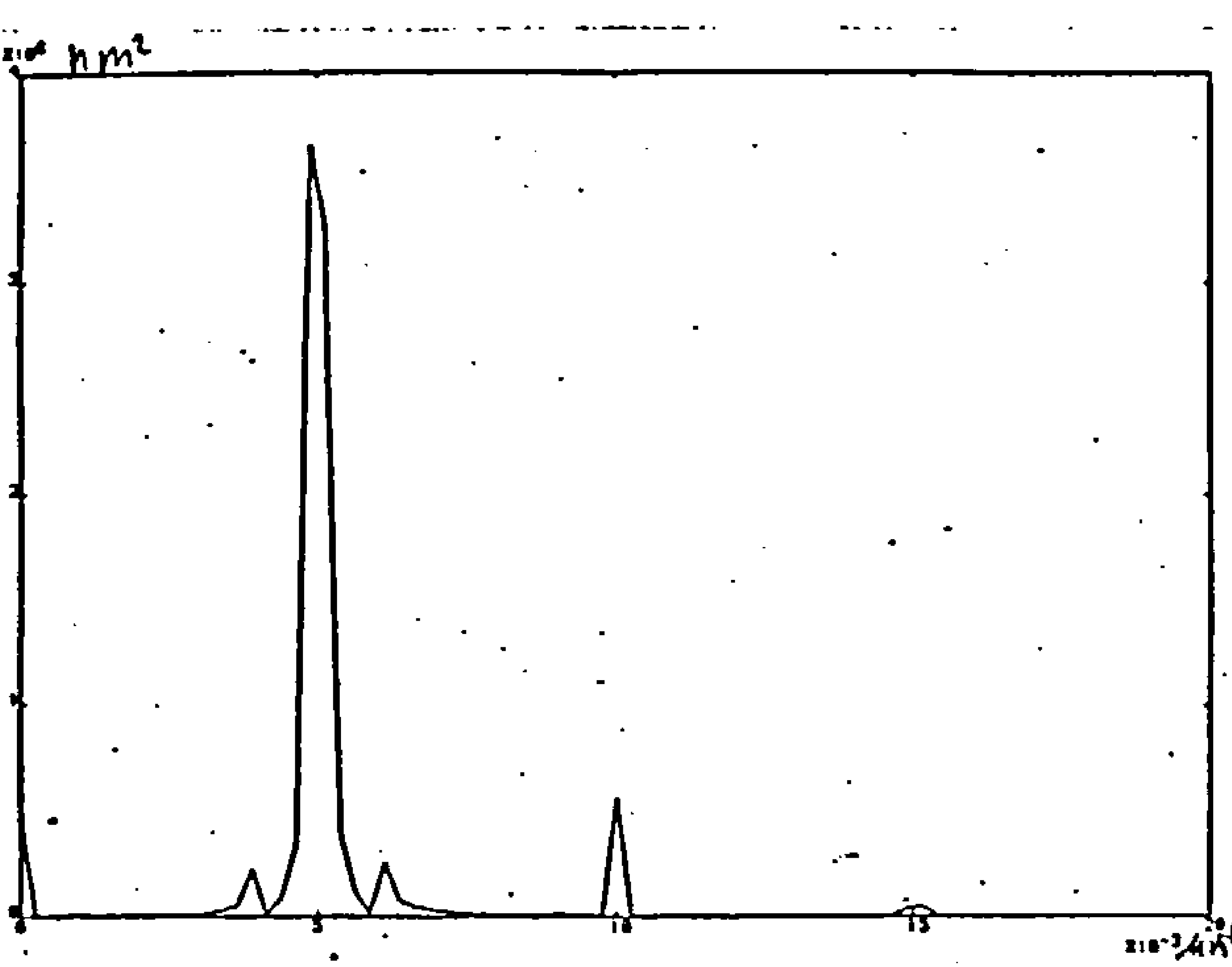


Figure 10.12e.

Fig.10.12 Power spectral density function of surface profiles with certain periodic machining errors
Fig.10.12a Perfect profile.
fig.10.12b Change of depth of cut.
Fig.10.12c Change of traverse.
Fig.10.12d Combination of above.
Fig.10.12e Amplitude modulation.

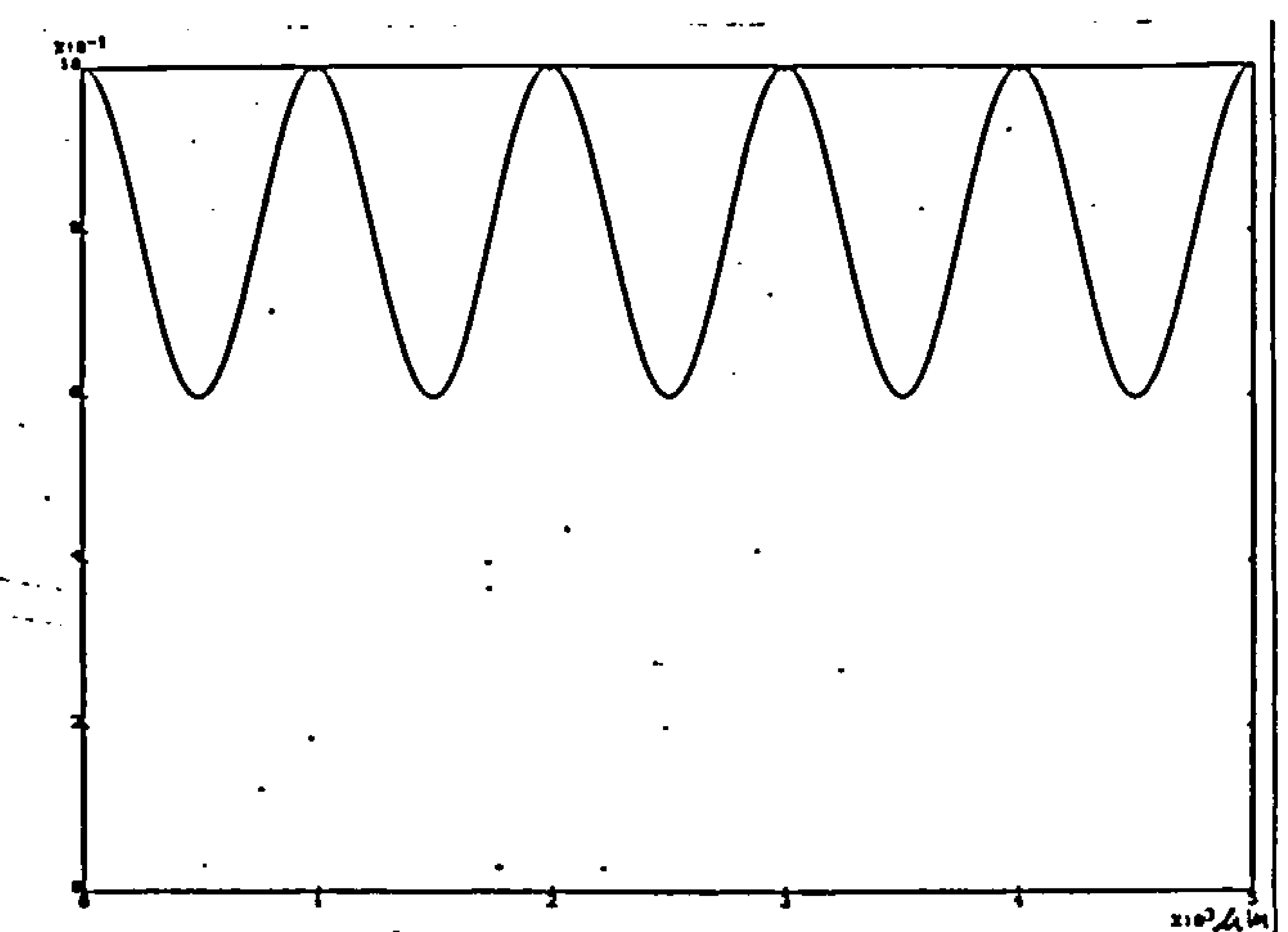


Figure 10.13 Modulation signal with which profile 10.10a has been multiplied to produc e profile 10.10e.

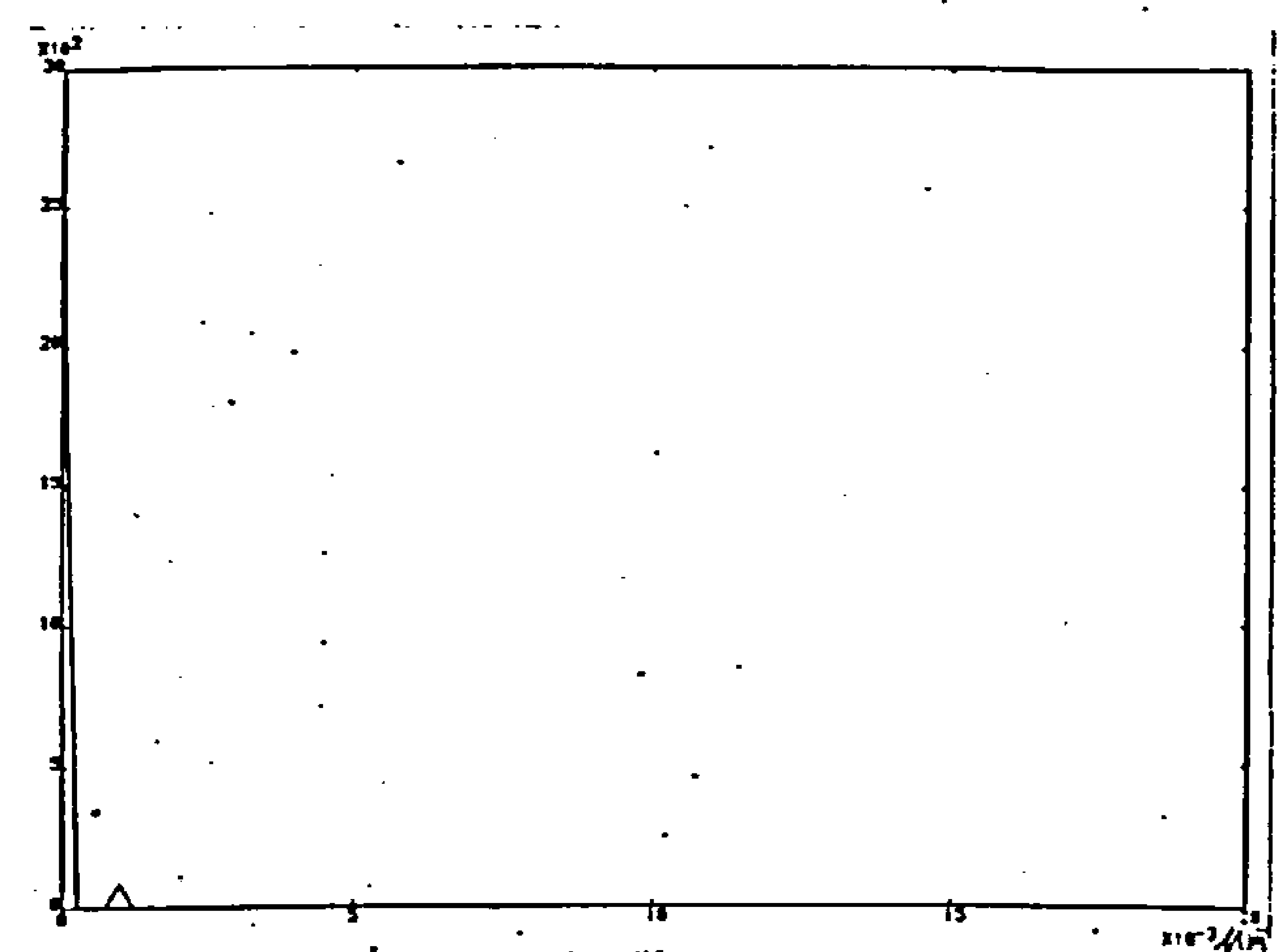


Figure 10.14 Power spectral density function of the modulation siganal 10.13

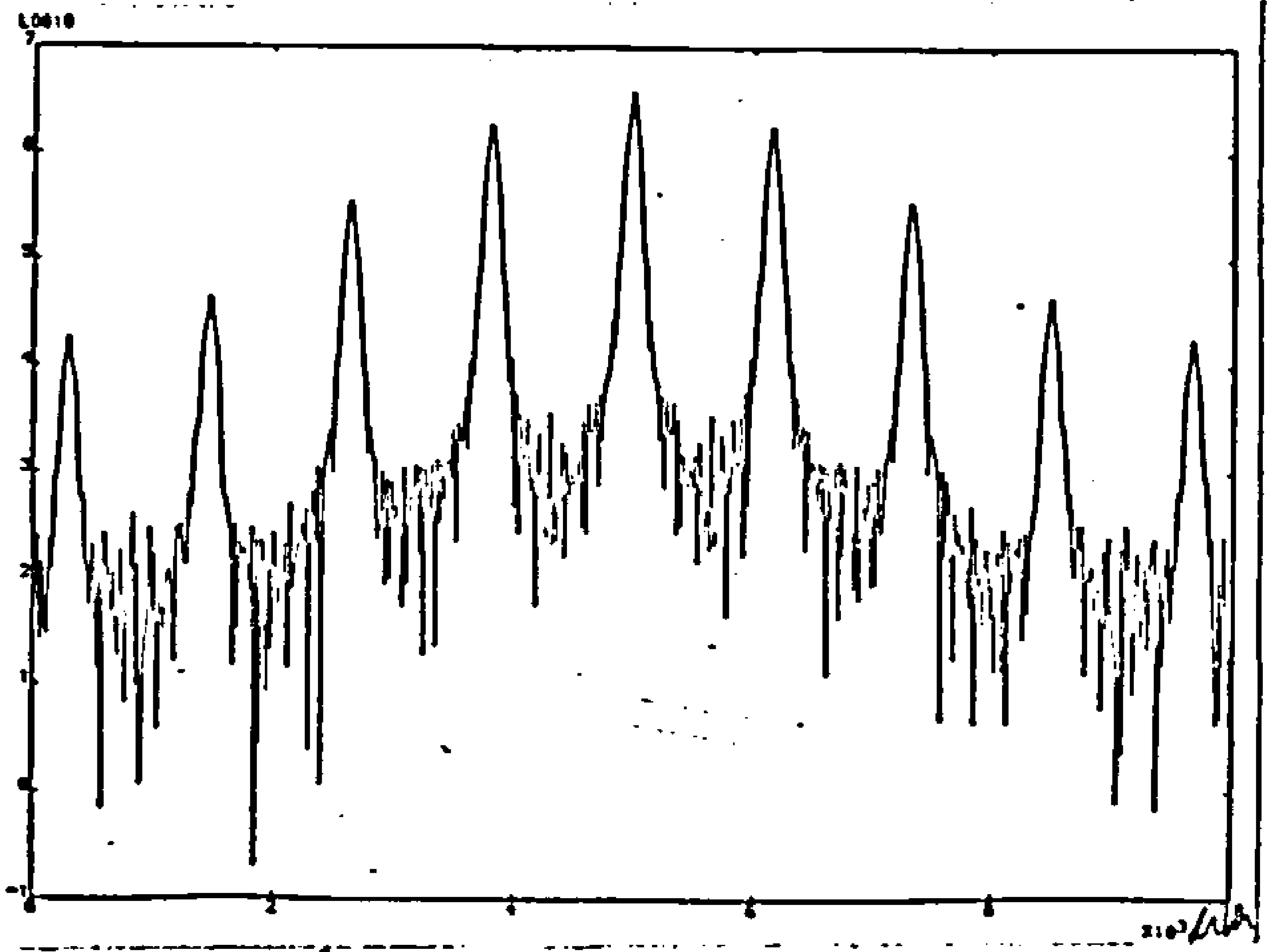


Figure 10.15a.

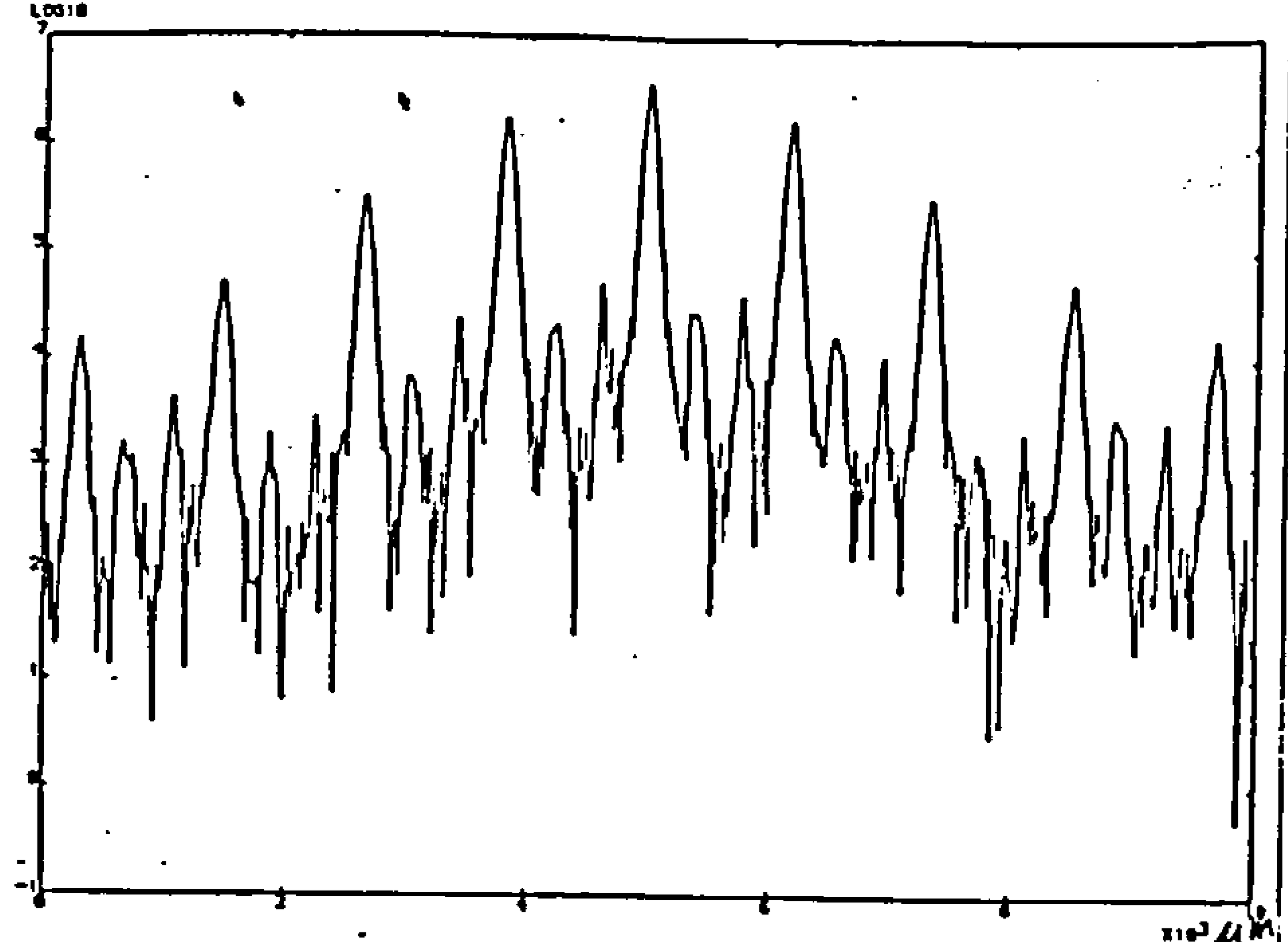


Figure 10.15b.

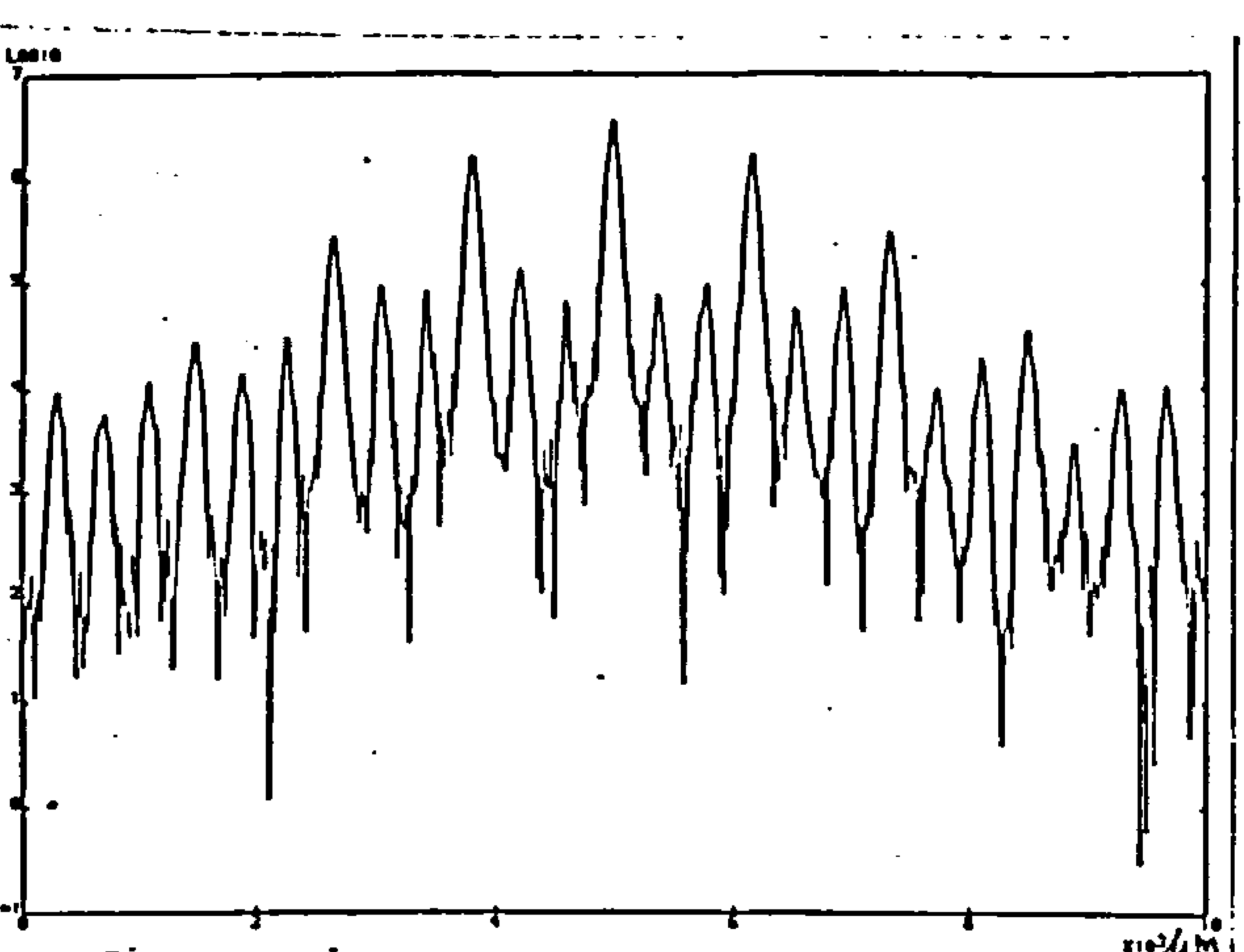


Figure 10.15c.

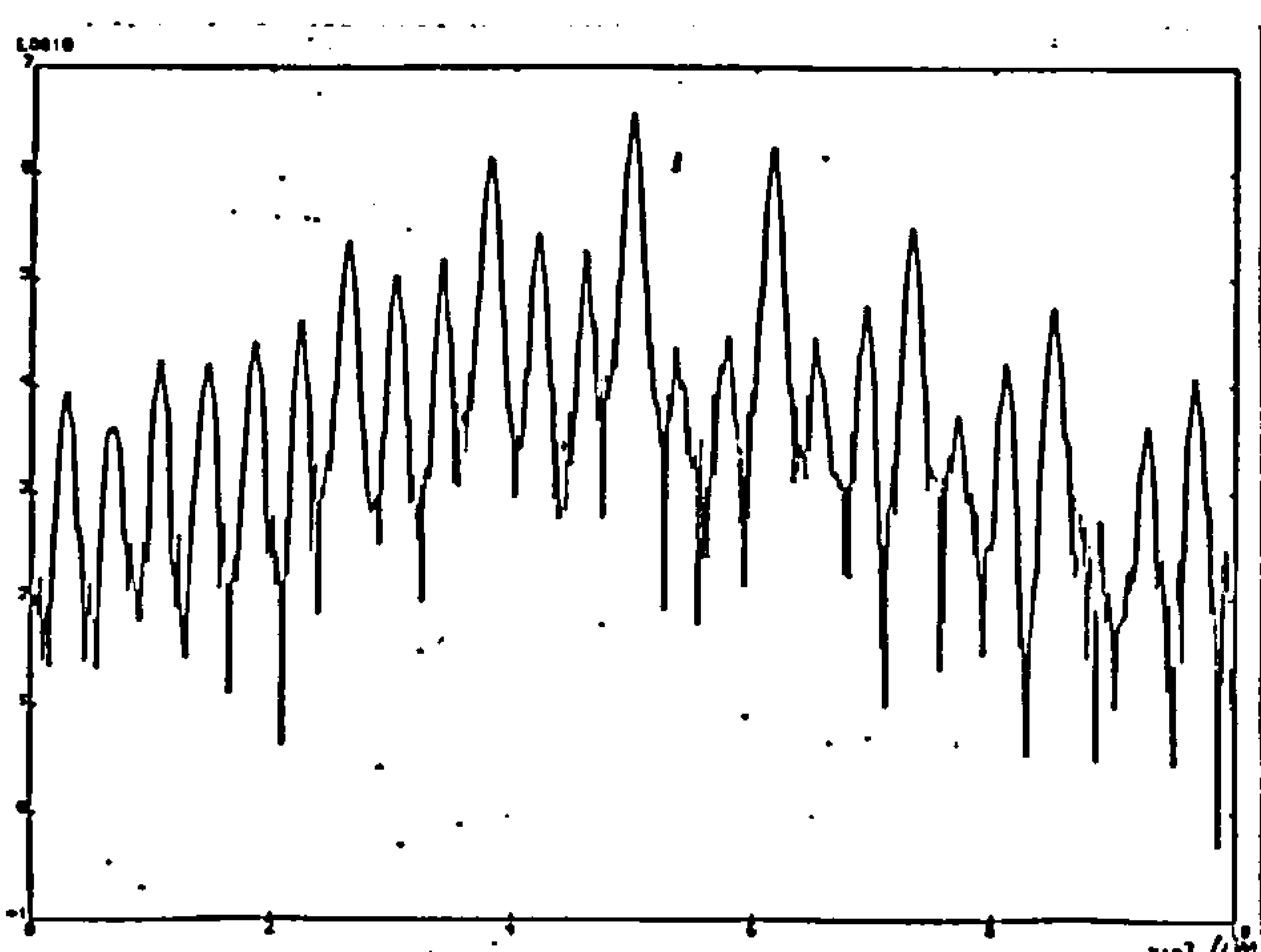


Figure 10.15d.

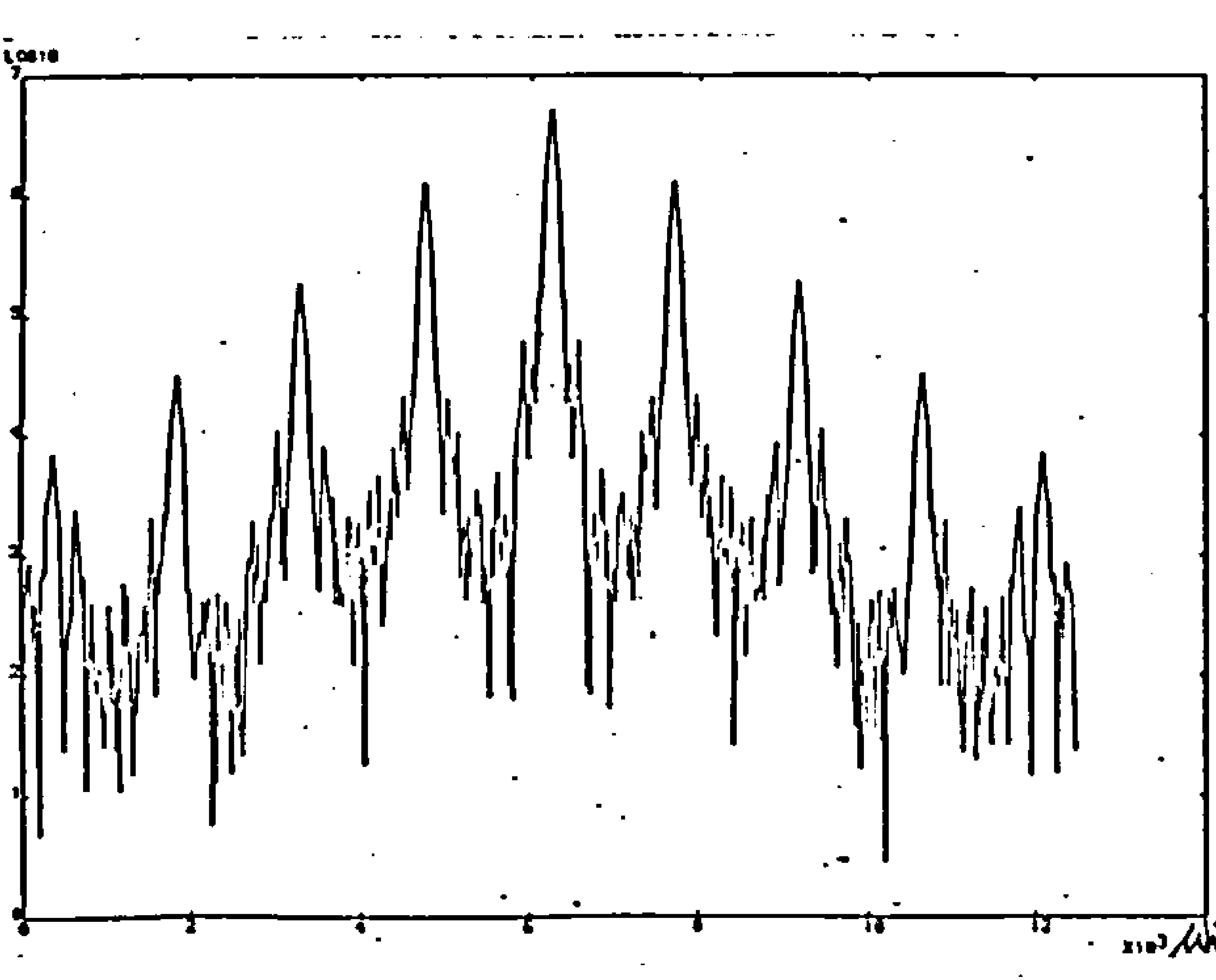


Figure 10.15e.

Fig.10.15 Simulated diffraction patterns of the surface profiles, plotted on a logarithmic scale.

Fig.10.15a Perfect profile.

fig.10.15b Change of depth of cut.

Fig.10.15c Change of traverse.

Fig.10.15d Combination of above.

Fig.10.15e Amplitude modulation.

10.5 Conclusion.

A computer program has been described which simulates successfully an optical sensor for obtaining diffraction patterns of surfaces. The simulation program has been used to produce diffraction patterns of sinusoidal surface profiles and of surface profiles containing certain machining errors. These patterns can be used to analyse real diffraction patterns and to obtain information of the machine tool behaviour. The program can also be utilised to produce an atlas mapping optical diffraction patterns versus machine tool errors [Rakels 1986a,b]. This atlas, when stored in the memory of the control computer of a machining centre, can enable this computer, to make more intelligent decisions than is thusfar possible [Rakels 1987].

From the computer results it is obvious that for machine-tool condition monitoring, a logarithmic light sensitive photo-diode array in the sensor is preferred above a linear sensitive one. The latter one, however, is more useful for the determination of the surface finish, since then the accurate values of the intensity amplitudes are very important.

If the advantages and flexibility offered by this simulation technique over more analytic methods are to be exploited to the full then the simulation needs to be extended to model scattering from a surface area, thus affording a more realistic comparison of simulated and actual diffraction patterns.

11. Deconvolution.

11.1. Introduction.

A surface roughness sensor which is based upon optical diffraction techniques suffers like most instruments that capture signals of physical phenomena, from a finite resolving power caused by abberations of the lenses and the shape and size of the illumination (section 10.3). A high resolution is important for the analysis of surface waviness, since then it is very likely that diffraction peaks merge together creating an apparent single diffraction order. When such contributions add up linearly, the distorted signal may be described by the mathematics of convolution. In optical systems this convolution manifests itself in blurring of the light intensity pattern. The causes of this blurring can be various, like the finite sizes of the optics, quality of the object illumination and the detector which converts the light intensity into an electrical signal. To counteract this blurring it is possible to follow two totally different routes. The obvious one is to improve the optical hardware which causes the blurring. However, this will quite often result in an expensive and sometimes bulky instrument. The other way out is to use mathematical deconvolution techniques, which can be performed by a digital computer. This technique can recover some of the ideal information, which was apparently lost due to the limited resolution of the observing instrument. Jansson [1984] shows various deconvolution techniques available. The most promising

amongst them are the so called non-linear constrained methods. These methods make use of some prior knowledge of the ideal signal. For example intensities cannot be negative or absorption coefficients must be within the range from zero to unity.

11.2. Definition of the convolution problem.

In its simplest form convolution can be considered to be a blurring or smoothing operation. A sharp input function is convolved with a smoothing function to yield a smoothed output function. Typically each output value is identified with a corresponding input value. However, it is obtained by processing its related input value and some of its neighbouring input values. This can be mathematically expressed by the convolution integral:

$$I(X)=\int_{-\infty}^{\infty} S(X-X')O(X')dX'=\int_{-\infty}^{\infty} O(X-X')S(X')dX'=S(X)\otimes O(X) \qquad (11.1)$$

where I(X) represents the known output data, S(X) the blurring phenomena, whose effects have to be eliminated, and O(X) is the ideal data, free of spreading due to S(X).

Now the problem can be defined by: "Is it possible to reconstruct O(X) when I(X) and S(X) are known?".

To find a solution to this problem several deconvolution methods will be discussed. It will also show their limitations.

11.3.1. Linear deconvolution methods.

A linear deconvolution method is one whose output elements (the restored data) can be expressed as linear combinations of the input elements (the captured data). These methods can be developed and analysed in detail by use of long standing mathematical tools. Analysis of linear methods tends to be simpler than that of non-linear methods. Also the computations are shorter. This point is especially important, because deconvolution is inherently computation intensive.

Direct approach.

The direct approach is derived from the discrete counterpart of equation (11.1):

$$I_m = \sum_{l=-L}^{L} S_l O_{m-l} \tag{11.2}$$

whereby it is assumed that the spread function S_l has only nonvanishing values in the range $-L < l < L$. The discrete expression is used to simplify the implementation of these results in digital computers, which have to perform the deconvolution.

Assume that the true object is zero for all subscript values less than 1. When convolved, the function I_m must have zero values of $m < -L+1$. The first nonzero value of I occurs when $m = -L+1$, this is when the end of the spread function encounters the first nonzero

value of O. This first value of I can be expressed by:

$$I_{-L+1}=S_{-L}O_1 \tag{11.3}$$

From this equation an estimate $\hat{O}$ of the first object value can be found

$$\hat{O}_1=I_{-L+1}/S_{-L} \tag{11.4}$$

The next value of the image is expressed by

$$I_{-L+2}=S_{-L+1}O_1+S_{-L}O_2 \tag{11.5}$$

Solving this equation for O_2, yields the estimate $\hat{O}_2$:

$$\hat{O}_2=(I_{-L+2}-S_{-L+1}O_1)/S_{-L} \tag{11.6}$$

Insertion of the estimate for $\hat{O}_1$ yields:

$$\hat{O}_2=(I_{-L+2}-S_{-L+1}I_{-L+1}/S_{-L})/S_{-L} \tag{11.7}$$

In this way one can express the $(k+1)^{th}$ object estimate in terms of the preceeding estimates and the image value I_{k-L+1}. This can be expressed by the recursive relationship given in (11.8).

$$\hat{O}_{k+1}=(I_{-L+1}-\sum_{l=-L+1}^{L}S_L O_{k-L+1-l})/S_{-L} \tag{11.8}$$

This method is ideally suited for implementation in a digital computer. However, there is one problem. The observed input data I_m is subject to high frequency noise, caused by the interference of two merging diffraction orders, and digitizing errors N_m. This means that the first object estimate is:

$$\hat{O}_1=(I_{-L+1}+N_{-L+1})/S_{-L} \tag{11.9}$$

Even for good signal to noise ratios (small N), the estimate $\hat{O}_1$ can suffer considerably, being erroneous by an amount N_{-L+1}/S where S_{-L} is in general small. Consequently there are problems in obtaining the first estimate. This first estimate is subsequently used to obtain the estimate of the second object $\hat{O}_2$. From this it is obvious that this method has its problems.

Van Cittert's method.

Van Cittert (1931) considered the image data I(X) as a first approximation $\hat{O}^{(0)}(X)$ of the object O(X). Now this estimate $\hat{O}^{(0)}(X)$ is blurred to yield an approximation $\hat{I}^{(0)}(X)$ of the image data. This is the form that the data would take if $\hat{O}^{(0)}(X)$ was the true object. Subsequently the difference between the blurred data $\hat{I}^{(0)}(X)$ and the image data I(X) is calculated. This error in the image estimate is related to the error in the object estimate, $O(X)-\hat{O}^{(0)}(X)$. Van Cittert recognised that this image estimate error could be applied as a correction to the object estimate, thus producing a new object

estimate:

$$\hat{O}^{(1)}(X)=\hat{O}^{(0)}(X) + [I(X)-S(X)\otimes\hat{O}^{(0)}(X)] \qquad (11.10)$$

Although $\hat{O}^{(1)}(X)$ is not the final solution, this correction will certainly produce a narrower and taller function than the original data I(X). To obtain better estimates, this process might be continued, which is expressed by:

$$\hat{O}^{(k+1)}(X)=\hat{O}^{(k)}(X)+ [I(X)-S(X)\otimes\hat{O}^{(k)}(X)] \qquad (11.11)$$

A problem with this method is that physically not realizable estimates can be produced such as negative intensities. The method has utility where only modest corrections are required.

Matrix inversion.

As shown earlier, the convolution integral can be written in discrete form:

$$I_n=\sum_m S_{n-m}O_m \qquad (11.12)$$

where I_n are the image values and O_m the object values; the response function is represented by the values S_{n-m}. Equation (11.12) can be expressed in a matrix formulation:

$$I=SO \tag{11.13}$$

In this equation I and O are column vectors such that each element $[I]_n=I_n$ and $[O]_m=O_m$. The elements of S are given by $[S]_{nm}=S_{n-m}$.

It appears straightforward to solve the set of linear equations of equation (11.13), since I and S are known and the only unknowns are O_m. Solving the set would yield the object estimate:

$$O=[S]^{-1}I \tag{11.14}$$

So all that is needed is to invert the matrix S. One computational problem could be that for long runs of information, the matrix S will be too large to handle because of limited computer storage. However, the real problem is the property of S. Typically each row in the matrix is similar to the above, and each of these rows do operate on neighbouring points. Each equation, therefore, brings only a small amount of added information into the system. So the set of equations are not robustly independent, and the matrix S is illconditioned. This can cause large and possibly oscillatory fluctuations in the estimate of O even for small noise perturbations in I. Also round off errors, caused by digitization of the data can cause the previously mentioned problem.

Iterative methods.

An iterative appraoch has the advantage of allowing control of

spurious fluctuations by interaction with the solution as it evolves. This can be done automatically by an algorithm, but then the method is no longer linear. Its use will be discussed in a later section. In this section two iterative methods are described which can be used for nonlinear deconvolution by slight modifications.

Point simultaneous.

A large system of linear equations (11.13) can be solved iteratively provided that the diagonal element is larger than any of its neighbours in the same row. Rewriting equation (11.13) with the diagonal element terms pulled outside:

$$I_n=[S]_{nm}O_n+\sum_{n\neq m}[S]_{nm}O_m \tag{11.15}$$

Solving this set of equations for O_n, gives:

$$O_n=(I_n-\sum_{n\neq m}[S]_{nm}O_m)/[S]_{nn} \tag{11.16}$$

From this it follows that O_n can be determined exactly if all its neighbours O_m are known. Now suppose that estimates of $\hat{O}_m^{(k)}$ of all its neighbours exist, then a refined estimate of O_n could be obtained,

$$O_n^{(k+1)}=(I_n-\sum[S]_{nm}\hat{O}_m^{(k)})/[S]_{nn} \tag{11.17}$$

by cycling through all the values of n until a complete set of $\hat{O}_n^{(k+1)}$ is

obtained.

One could also include the old estimate of $\hat{O}_n^{(k)}$, which eliminates the awkwardness of a summation with a missing element. Now equation (11.17) can be rewritten as:

$$\hat{O}_n^{(k+1)} = \hat{O}_n^{(k)} + (I_n - \sum_m [S]_{nm} \hat{O}_m^{(k)})/[S]_{nn} \tag{11.18}$$

Repeated application of this equation constitutes the point simultaneous relaxation method.

There is no guarantee that the correction term $1/[S]_{nn}$ is scaled correctly to ensure the most rapid convergence. Furthermore, if it is scaled too large, the method might not converge at all. Accordingly, a relaxation factor is introduced so that:

$$\hat{O}_n^{(k+1)} = \hat{O}_n^{(k)} + K(I_n - \sum_m [S]_{nm} \hat{O}_m^{(k)})/[S]_{nn} \tag{11.19}$$

if $K=[S]_{nn}$, the discrete formulation of Van Cittert is obtained.

Point successive.

To compute a single point $\hat{O}_n^{(k+1)}$ in the estimate, one needs the old estimate $\hat{O}_n^{(k)}$ at that point and a number of points to either side as well. Usually the $\hat{O}_n^{(k)}$ are computed in the order of increasing n, thereby using the old values $m<n$, although more recent estimates are available. To avoid this, and thereby using the most recent values available, the summation is to be split into two parts:

$$\hat{O}_n^{(k+1)} = \hat{O}_n^{(k)} + K(I_n - \sum_{m<n} [S]_{nm} \hat{O}_m^{(k+1)} - [S]_{nm} \hat{O}_m^{(k)}) / [S]_{nn} \qquad (11.20)$$

One can expect a faster convergence of this method since use is made of the most recent information. For K=1, this method is called point successive relaxation. However, note that whereas equation (11.19) can be applied with either ascending or descending n, yielding exactly identical results, equation (11.20) can only be applied in the order of ascending n. It is a fundamentally asymmetrical method and can produce slightly asymmetrical artifacts. To overcome this problem binary interleaving has to be employed [Deutsch,1965;Allebach, 1981].

Another advantage, besides the faster convergence, is the fact that this method only needs one array in the computer to store both $\hat{O}^{(k)}$ and $\hat{O}^{(k+1)}$, because the element $\hat{O}^{(k)}$ is never again needed once $\hat{O}^{(k+1)}$ is computed. This is in contrast with the point simultaneous methods, where, for each n, the elements of $\hat{O}^{(k)}$ must be available corresponding to the full range over which $[S]_{nm}$ is non-vanishing.

11.3.2 Constrained non-linear methods.

In the preceeding sections several linear deconvolution methods are discussed. It demonstrated that these methods have problems with noisy signals, and they can produce solutions which are not physical possible. To overcome these shortcomings, some constraints have to be introduced in the deconvolution process. These constraints have to prevent that physically impossible solutions are obtained. In the

case of intensity patterns, they have to make sure that the intensity is never negative, so a lower bound is placed upon the solution. It is obvious that this method will be nonlinear, since the solution O(X) cannot be expressed as a linear function of the measured data I(X). The term "nonlinear" used further on should not be confused with the nonlinearities in the processing of data values, the photometric nonlinearity. It is assumed that these latter effects are compensated for before the deconvolution process is invoked. Some nonlinear constrained methods will be discussed in the next section. They are based on linear methods but are slightly modified to produce sensible results.

Gold's [1964] ratio method.

As noted, some linear deconvolution methods produce solutions that do not make physical sense. Negative values can be obtained for light intensities. To overcome this Gold developed a method of iteration similar to van Cittert's but with multiplicative corrections instead of additive ones.

Given the equation describing the signal distortion

$$I=S\otimes O \tag{11.21}$$

and an estimate $\hat{O}^{(k)}$, this method proposes to estimate the observed data by use of:

$$I^{(k)}=S\otimes\hat{O}^{(k)} \tag{11.22}$$

A ratio containing this quantity and the data may be used to construct a new estimate of the true spectrum,

$$\hat{O}^{(k+1)}=\hat{O}^{(k)}[I/\hat{I}^{(k)}] \tag{11.23}$$

the iteration may be given in a single equation:

$$\hat{O}^{(k+1)}=\hat{O}^{(k)}I/S\otimes\hat{O}^{(k)} \tag{11.24}$$

As with van Cittert's method, $\hat{O}^{(0)}=I$. Obviously, the successive estimates $\hat{O}^{(k)}$ cannot be negative, provided that S and I are everywhere positive. This assumption may be violated in the baseline data, because of instrumentation noise.

Jansson's method.

This nonlinear iterative method is based on the linear relaxation method as described in section 11.3.1. Initially, the correction term K was set equal to zero in regions where $\hat{O}^{(k)}$ was nonphysical. To illustrate this, the point-simultaneous equation can be rewritten with a relaxation parameter $r[\hat{O}^{(k)}]$ that depends on the estimate $\hat{O}^{(k)}$:

$$\hat{O}^{(k+1)}=\hat{O}^{(k)}+r[\hat{O}^{(k)}](I-S\otimes\hat{O}^{(k)}) \tag{11.25}$$

For clipping or simple truncation of the nonphysical part, the relaxation function can be defined by:

$$r[\hat{O}^{(k)}] = r_0 \text{ for } \hat{O}^{(k)} \text{ physical} \qquad (11.26)$$
$$\qquad 0 \text{ elsewhere}$$

It is generally desirable to choose r_0 large so that large corrections will be made to $\hat{O}^{(k)}(X)$, allowing it to approach $O(X)$ with the smallest amount of computation. Assume that r_0 is chosen fairly large, then some initial corrections will be overcorrections. These overcorrections will typically damp out with continued iteration and a solution results. However, if r_0 is chosen too large, this method will not converge.

A problem with the above mentioned defenition of $r[\hat{O}^{(k)}]$ is that an overcorrection, that would normally disappear gradually through ensuing iterations, results a value of O(X) that vanishes for all subsequent iterations. This behaviour occurs since further corrections to that value are prohibited. To use this method, one has to be careful not to use too a large value for $r[\hat{O}^{(k)}]$. Even in this case, erroneously nonphysical values of $\hat{O}^{(k)}(X)$ that have been forced to zero are never allowed to return to the finite range that might better represent the true spectrum of O(X). So this method can yield a solution that, although physical realizable, is not necessarily the best achievable estimate.

Instead of clipping the correction term, a relaxation function can be introduced, that modulates the correction term. It assumes full value in the centre of the physical region and falls to zero at

the amplitude bounds. To refine this idea further, the functional form for $r[\hat{O}^{(k)}(X)]$ can assume negative values for nonphysical solutions:

$$r[\hat{O}^{(k)}(X)]=r_o[1-2|\hat{O}^{(k)}(X)-0.5|] \quad (11.27)$$

Relaxation functions.

The key to the success of constrained deconvolution is the relaxation function $r^{(k)}[y]$, where y is the abbrevation for $\hat{O}^{(k-1)}(X)$. The constrained deconvolution algorithm is then specified by:

$$\hat{O}^{(k)}(X)=\hat{O}^{(k-1)}(X)+r^{(k)}[y](I(X)-y\otimes S(X)) \quad (11.28)$$

The $\hat{O}^{(0)}(X)$ estimate is generally taken to be I(X). Various relaxation functions have been suggested such as:

$$r^{(k)}[y]=r_o[1-2|y-0.5|] \quad (11.29)$$

$$r^{(k)}[y]=r_o[y\cdot(1-y)] \quad (11.30)$$

$$\begin{aligned} r^{(k)}[y]&=r_o 2y && \text{for } y<0.5 \\ &= r_o e^{(0.5-y)} && \text{for } y>0.5 \end{aligned} \quad (11.31)$$

$$r^{(k)}[y]=r_o[e^{[-(y-0.5)^2/0.11]}] \quad (11.32)$$

$$r^{(k)}[y]=r_o \quad \text{for } y>0 \qquad (11.33)$$

$$=0 \quad \text{for } y<0$$

$$r^{(k)}[y]=r_o y e^{(-y)} \qquad (11.34)$$

The above listed functions do not necessarily yield the correct results for one particular problem. The choice of the right function has to be tested by trial and error with simulated data. This data has to be generated by convolving the ideal data with the spread function S(X) in order to obtain the function I(X), which then in turn has to be deconvolved with S(X). Comparison between the deconvolved signal and the original ideal signal must show whether or not the right relaxation function has been decided upon. A least squares method can serve as to this effect.

Deconvolution algorithm.

Step by step, the constrained deconvolution algorithm may be stated as follows:

(1) Set k=1, $\hat{O}^{(0)}(x)=I(x)$

(2) Form $\hat{O}^{(k)}(X)=\hat{O}^{(k-1)}(X)+r^{(k)}[y](I(X)-\sum \hat{O}^{(k-1)}(X)\otimes S(X))$

(3) Invoke termination-convergence tests. If satisfied, the current $\hat{O}^{(k)}(X)$ is the estimate of O(X) and the procedure is completed. If not satisfied, set k=k+1 and return to step (2)

Convergence may be monitored by forming the root-mean-square error of the current estimate as:

$$e= \sqrt{\sum_{X}[I(X)-\hat{O}^{(k-1)}(X) \circ S(X)]^2/N} \qquad (11.35)$$

Smoothing and reblurring.

In Jansson's original work, smoothing of the data was found to be essential for acceptable restoration. Smoothing, however, potentially limits the resolution achievable. When applied to the data only before deconvolution, the smoothing function enters as additional and uncompensated spreading. This undesirable effect may be alleviated by including the additional spreading in S(X) too. This is accomplished by smoothing S(X) before it is used for deconvolution.

11.4. Tests on the non-linear deconvolution algorithm.

The constrained deconvolution algorithm, as described in section 11.3.2. has been applied to the artificially created data shown in figure 11.1. This data has been produced by convolving the "ideal" spectrum shown in figure 11.2. with the blurring function depicted in figure 11.3. This blurring function has been obtained by a diffractometric device. A flat mirror was used as an object. Ideally, all the light reflected from the mirror should collect in the focal point of the optical diffraction system, thereby producing an infinitely narrow zero order diffraction peak. The ideal pattern

consists of two groups of "diffraction peaks" at decreasing mutual distances. The amplitudes of the peaks are constant in the first group and unequal in the second group. This enables the testing of the algorithm not only for spatial resolution, but for amplitude restoration as well. The positions and the relative amplitudes of the ideal diffraction pattern are listed in table 11.1. Table 11.2 lists the relative intensity distribution of the blurring function. The spatial unit is set equal to the distance between two adjacent photo diodes (25.4um).

Table 11.1.

Ideal diffraction pattern.

position	amplitude %
12	100
35	100
47	100
57	100
65	100
71	100
75	100
77	100
95	100
107	50
117	25
125	75
131	100
135	50

Table 11.2

Blurring function

position	amplitude %
1, 23	1
2, 22	2
3, 21	3
4, 20	4
5, 19	6
6, 18	9
7, 17	13
8, 16	17
9, 15	26
10, 14	57
11, 13	87
12	100

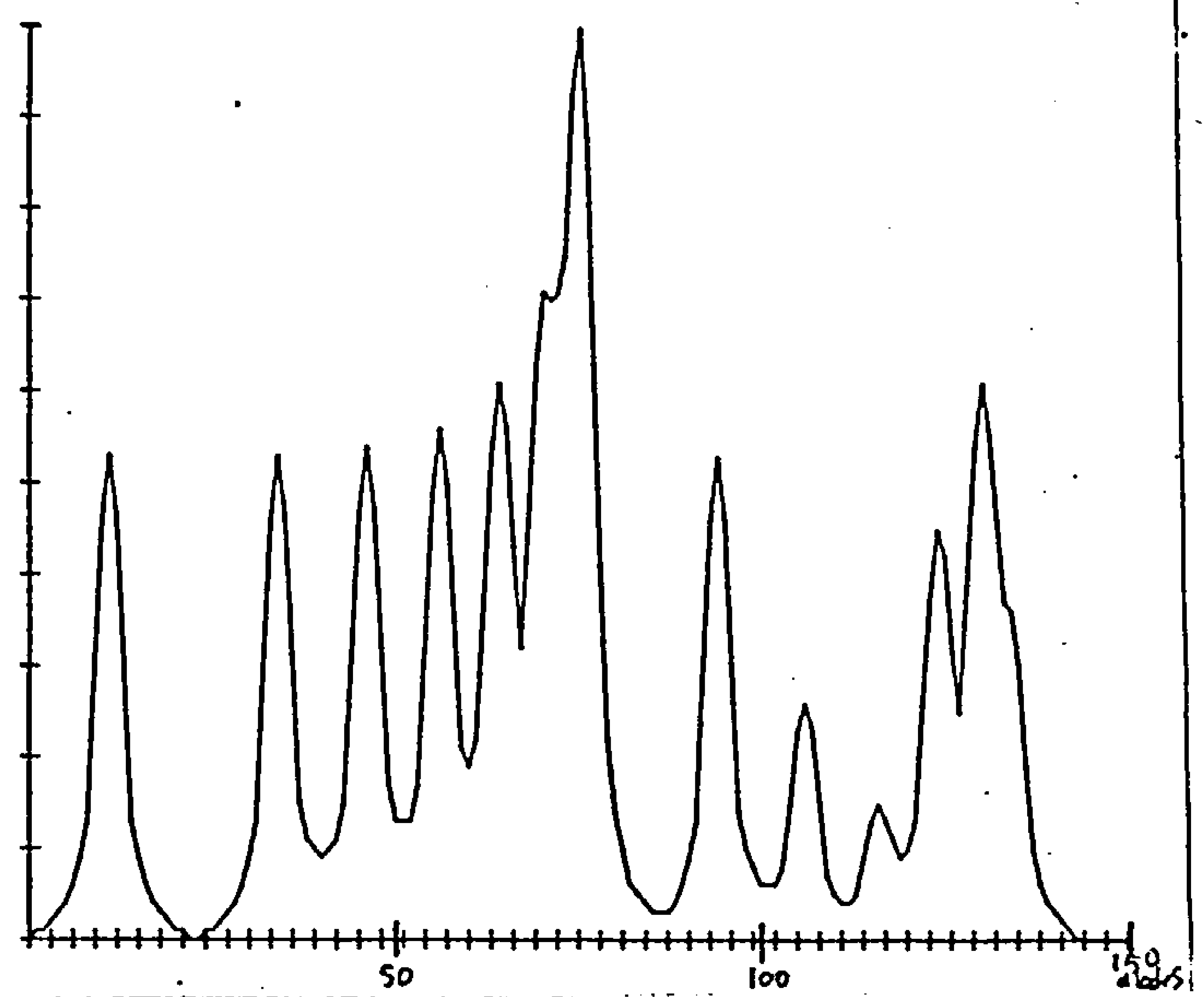


Figure 11.1. Convoluted input signal.

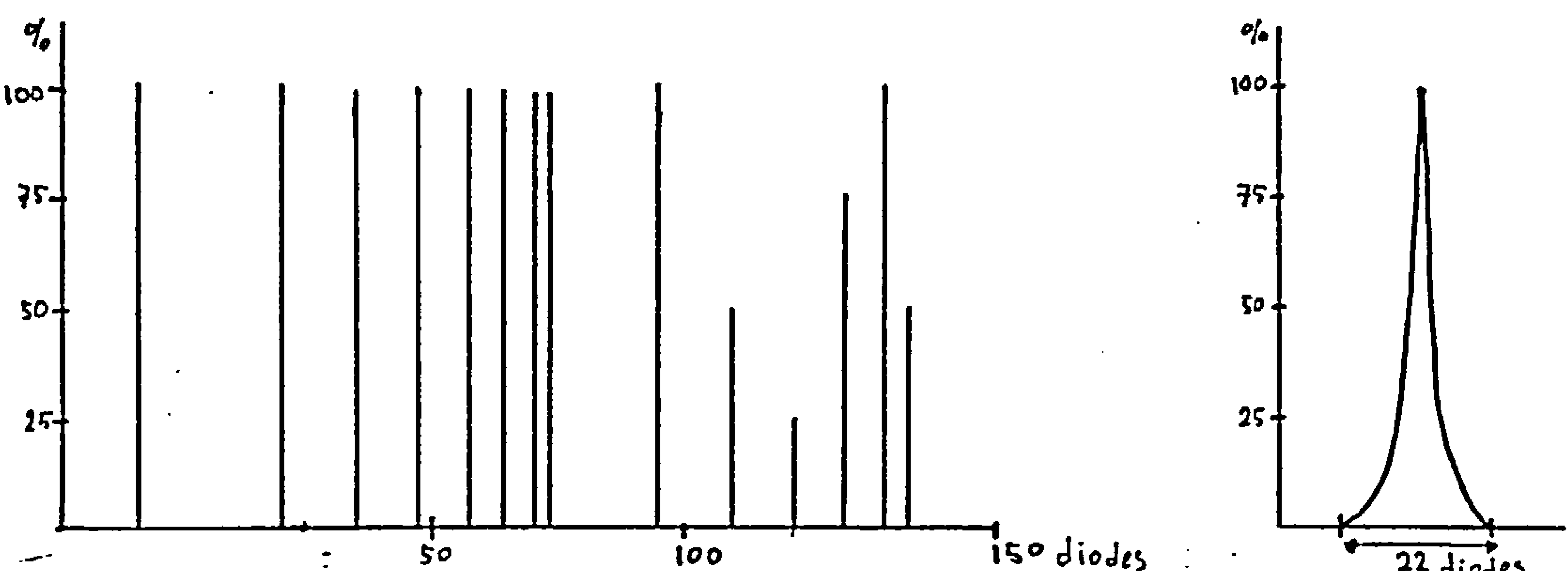


Figure 11.2. "Ideal" diffraction pattern Figure 11.3 Blur-function.

The six relaxation functions mentioned in section 11.3.2. and an extra one have been used in the deconvolution algorithm. They are:

A $r^{(k)}[\hat{y}]=1-2|\hat{y}-0.5|$

B $r^{(k)}[\hat{y}]=4\cdot\hat{y}\cdot(1-\hat{y})$

C $r^{(k)}[\hat{y}]=e^{-(\hat{y}-0.5)^2/.11}$

D $r^{(k)}[\hat{y}]=\hat{y}\cdot 2$ $\qquad \hat{y}<0.5$

$\qquad =e^{+0.5-\hat{y}}$ $\qquad \hat{y}>0.5$

E $\quad r^{(k)}[\hat{y}]=8\cdot\hat{y}\ e^{-\hat{y}/3}$

F $\quad r^{(k)}[\hat{y}]=1 \qquad \hat{y}>0$

$\qquad\quad =0 \qquad \hat{y}<0$

G $\quad r^{(k)}[\hat{y}]=1-\hat{y} \qquad \hat{y}>0$

$\qquad\quad =0 \qquad \hat{y}<0$

where $\hat{y}=\hat{y}/y_{max}$. Figure 11.4. shows the shape of these relaxation functions.

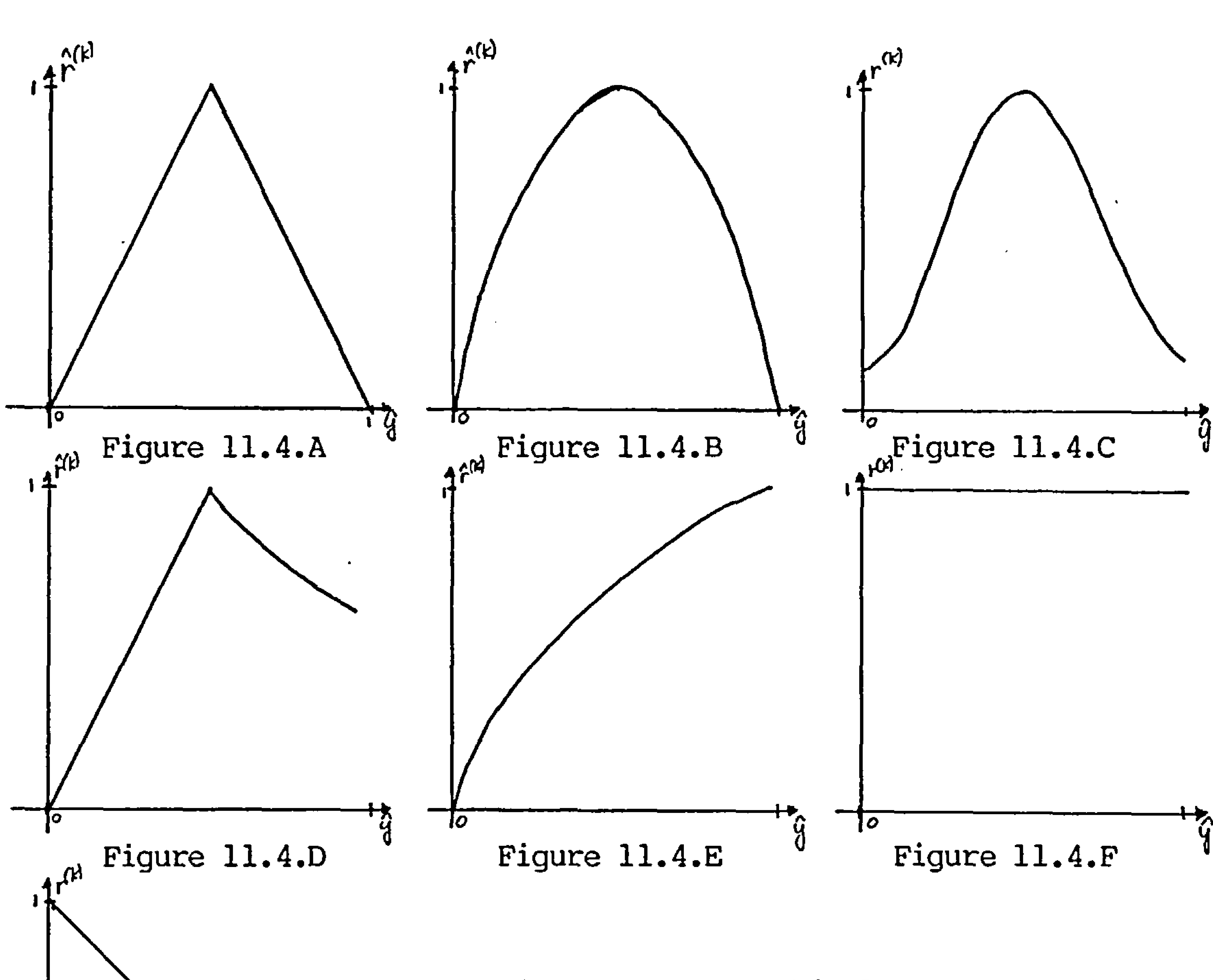


Figure 11.4.G

Figure 11.4. Graphical representation of the relaxation functions A-G.

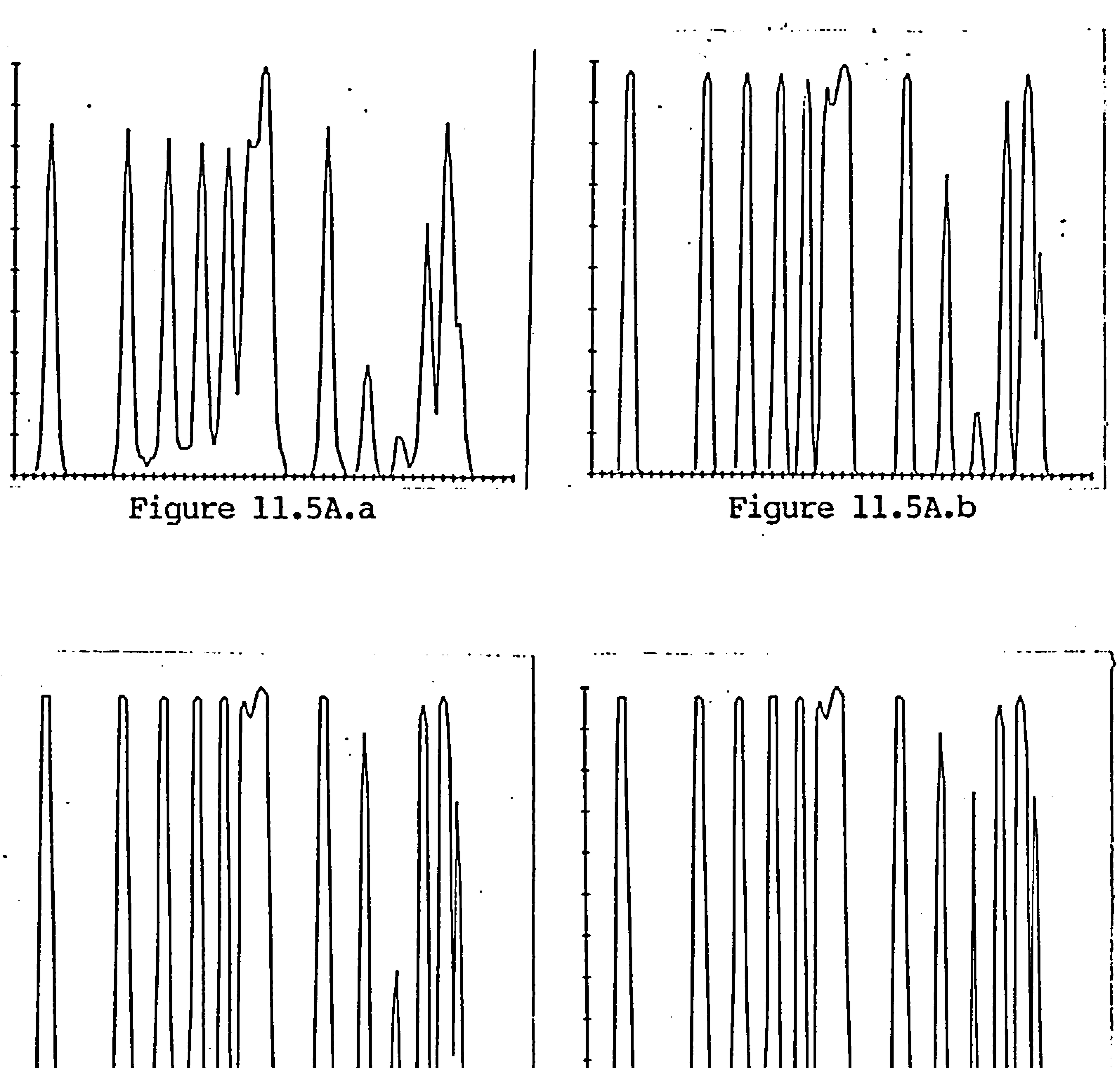


Figure 11.5A.a

Figure 11.5A.b

Figure 11.5A.c

Figure 11.5A.d

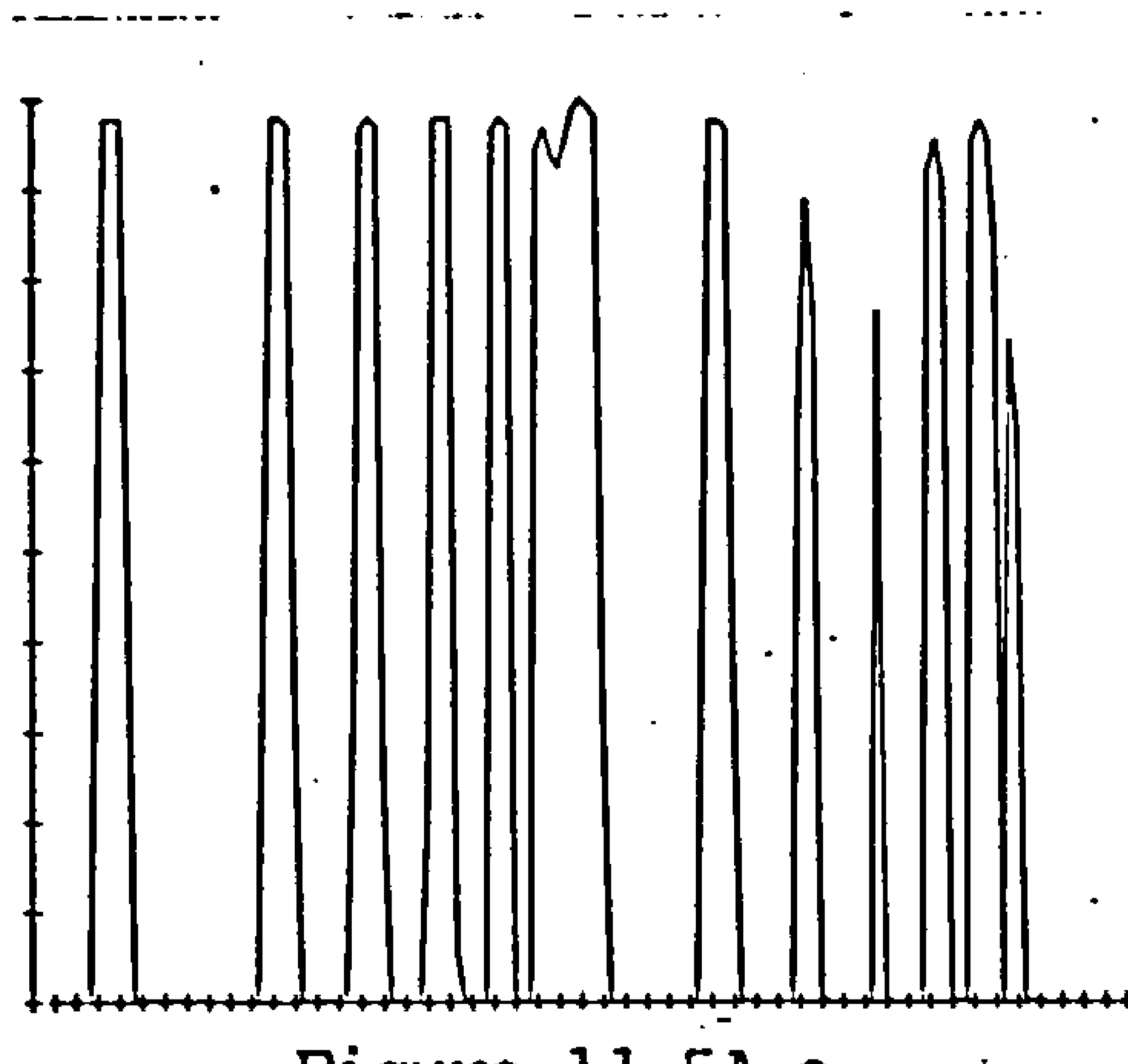


Figure 11.5A.e

Fig.11.5Ae After 1 iteration.

Fig.11.5Ab After 5 iterations.

Fig.11.5Ac After 10 iterations.

Fig.11.5Ad After 50 iterations.

Fig.11.5Ae After 100 iterations.

Figure 11.5A Results obtained by the relaxation function A.

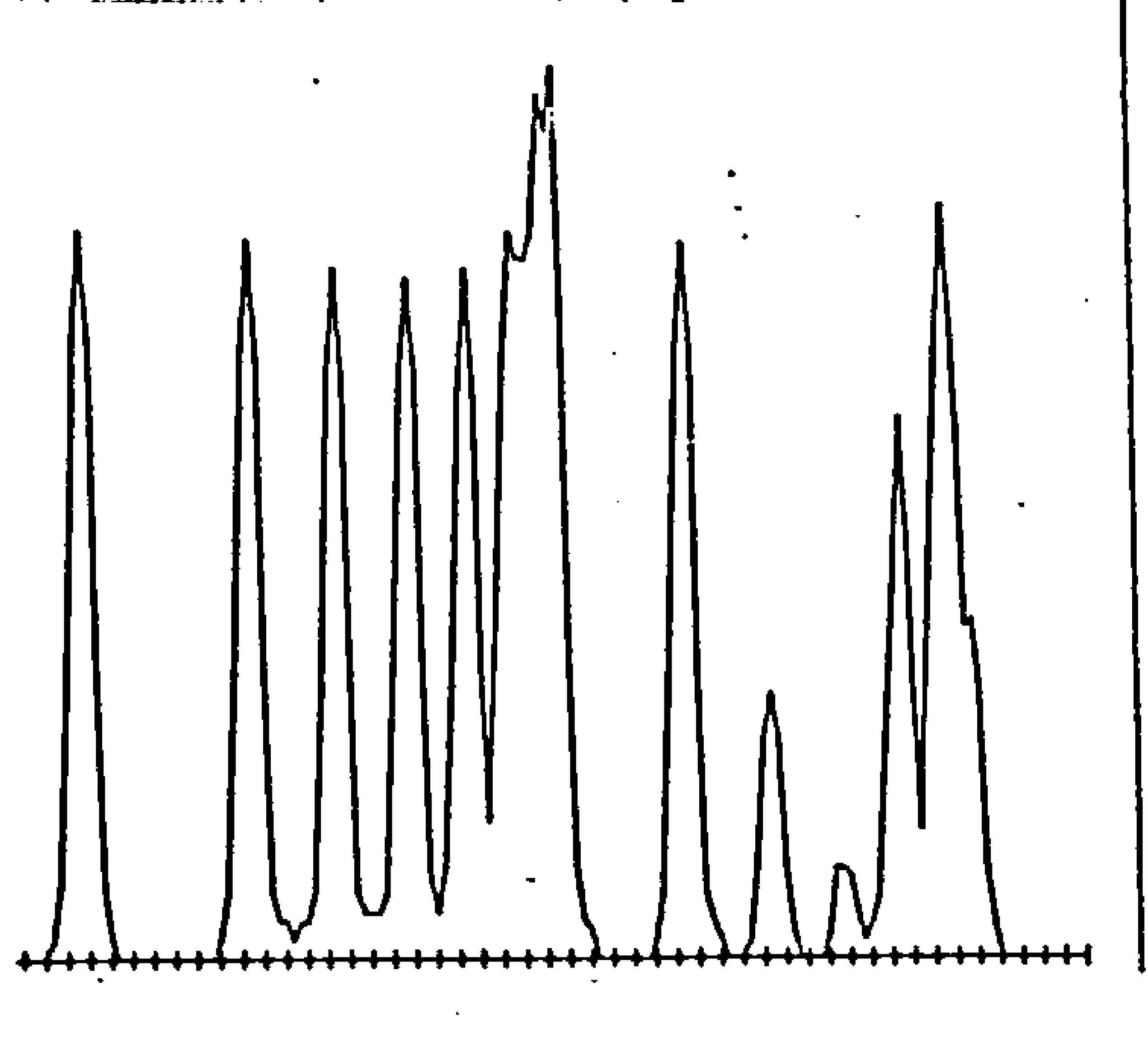


Figure 11.5B.a

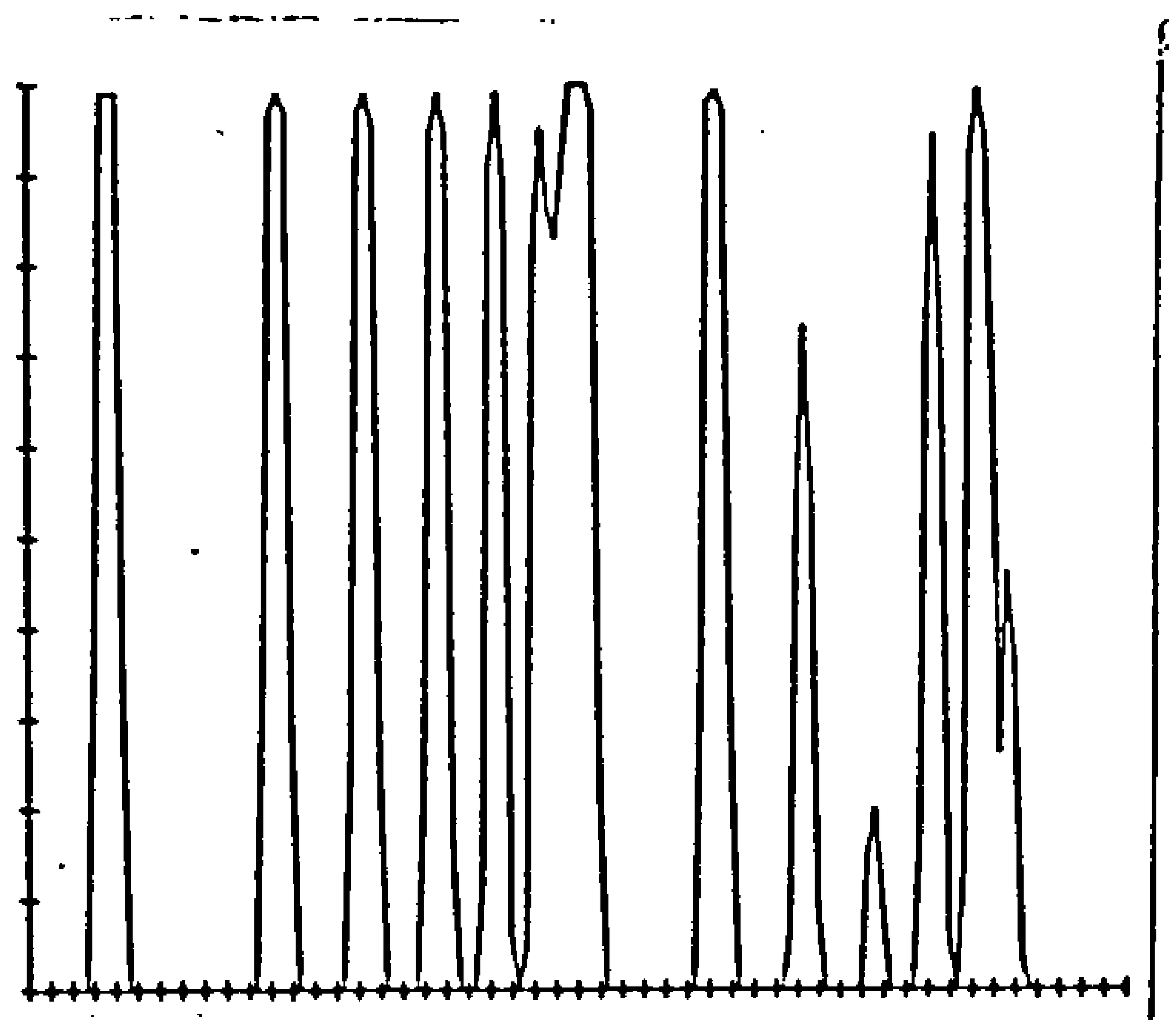


Figure 11.5B.b

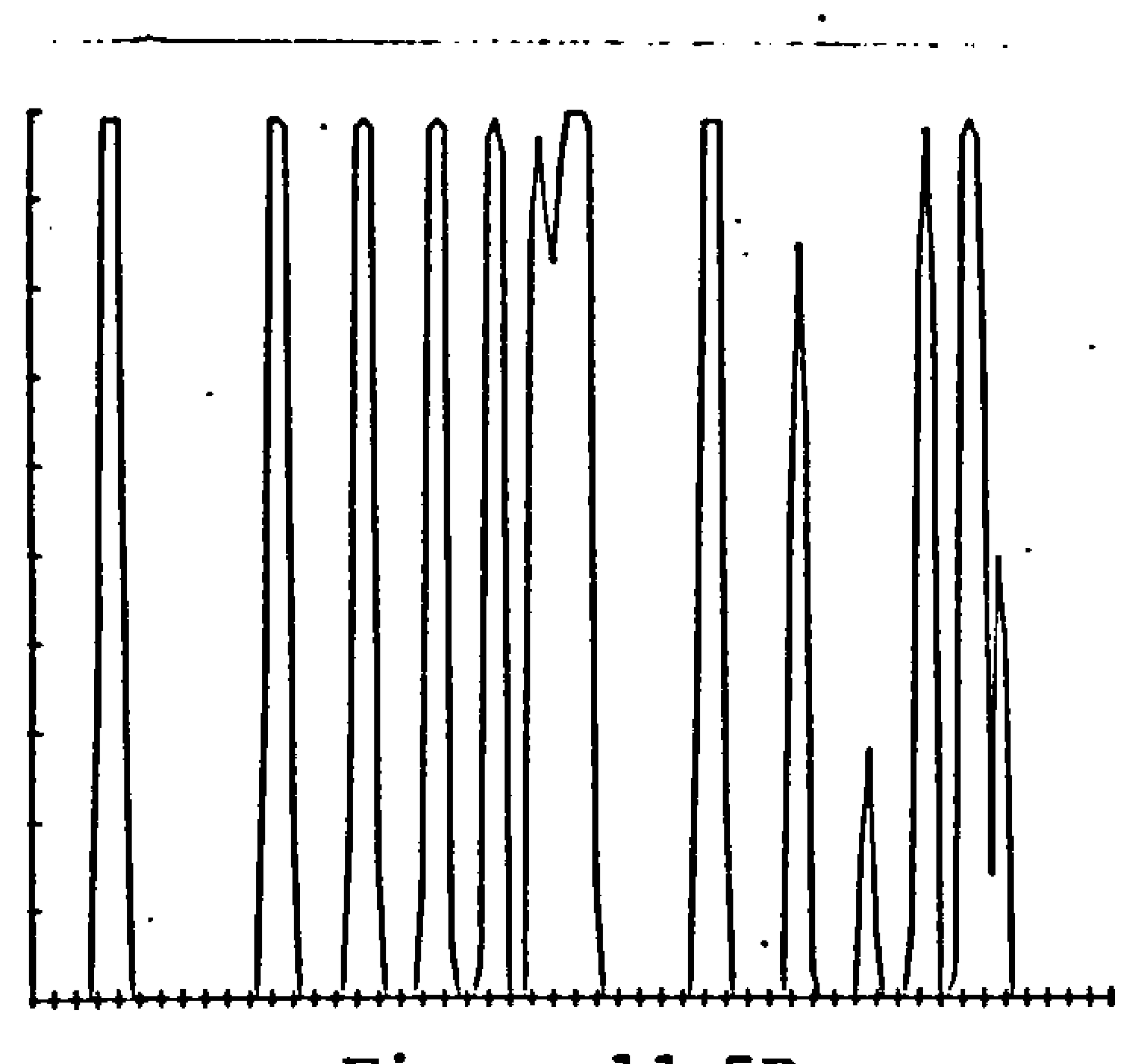


Figure 11.5B.c

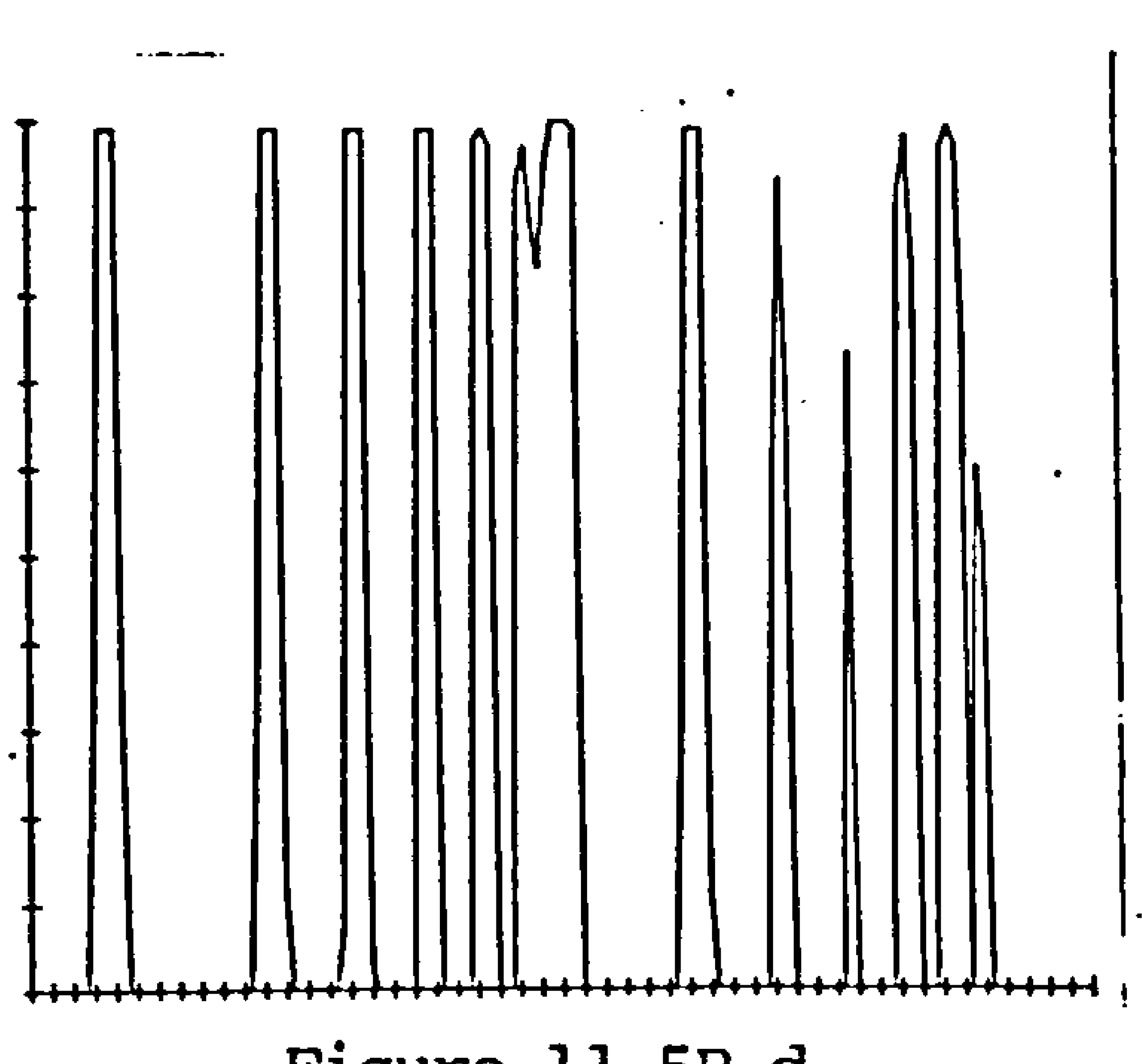


Figure 11.5B.d

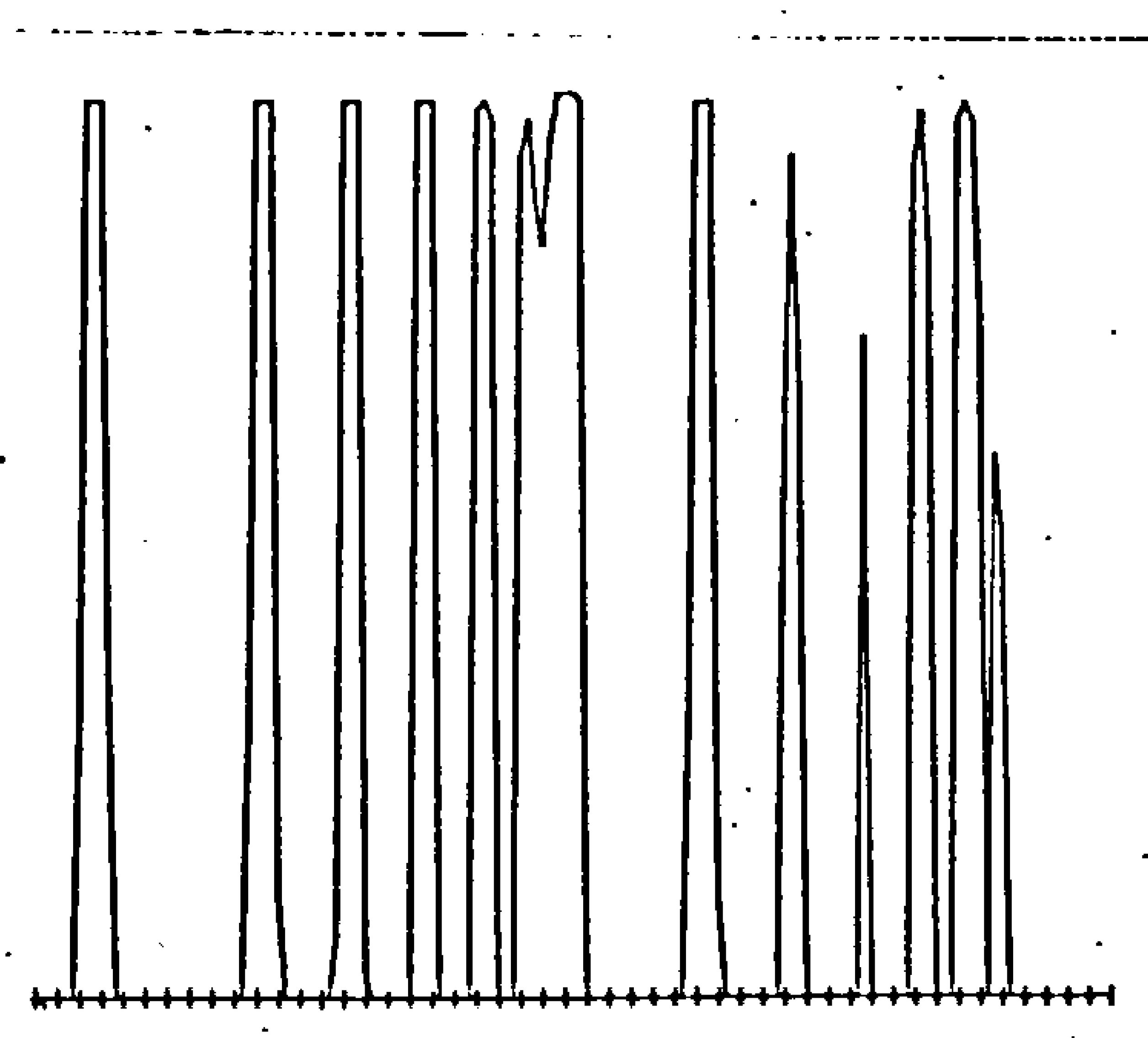


Figure 11.5B.e

Fig.11.5Be After 1 iteration.
Fig.11.5Bb After 5 iterations.
Fig.11.5Bc After 10 iterations.
Fig.11.5Bd After 50 iterations.
Fig.11.5Be After 100 iterations.

Figure 11.5B Results obtained by the relaxation function B.

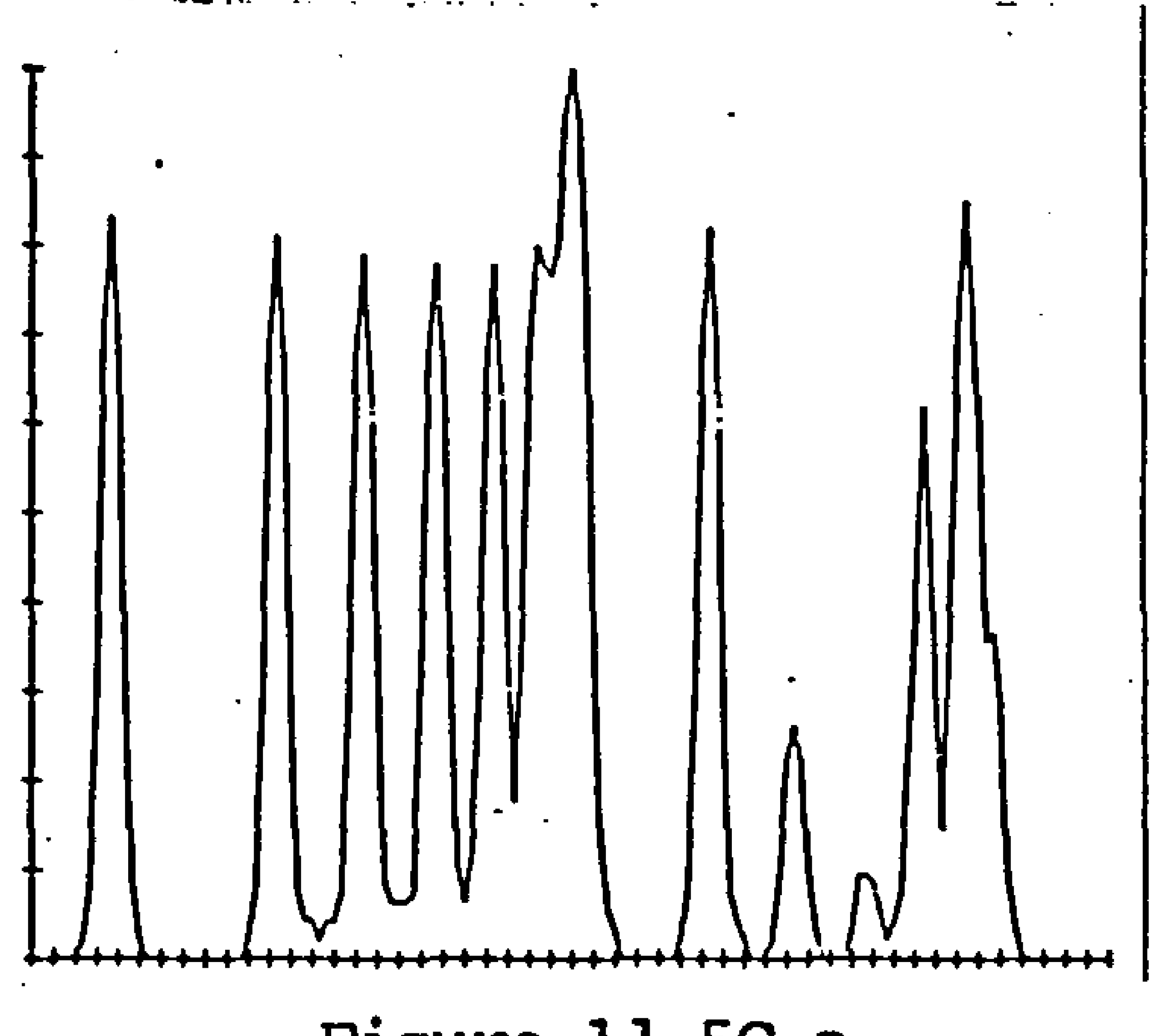


Figure 11.5C.a

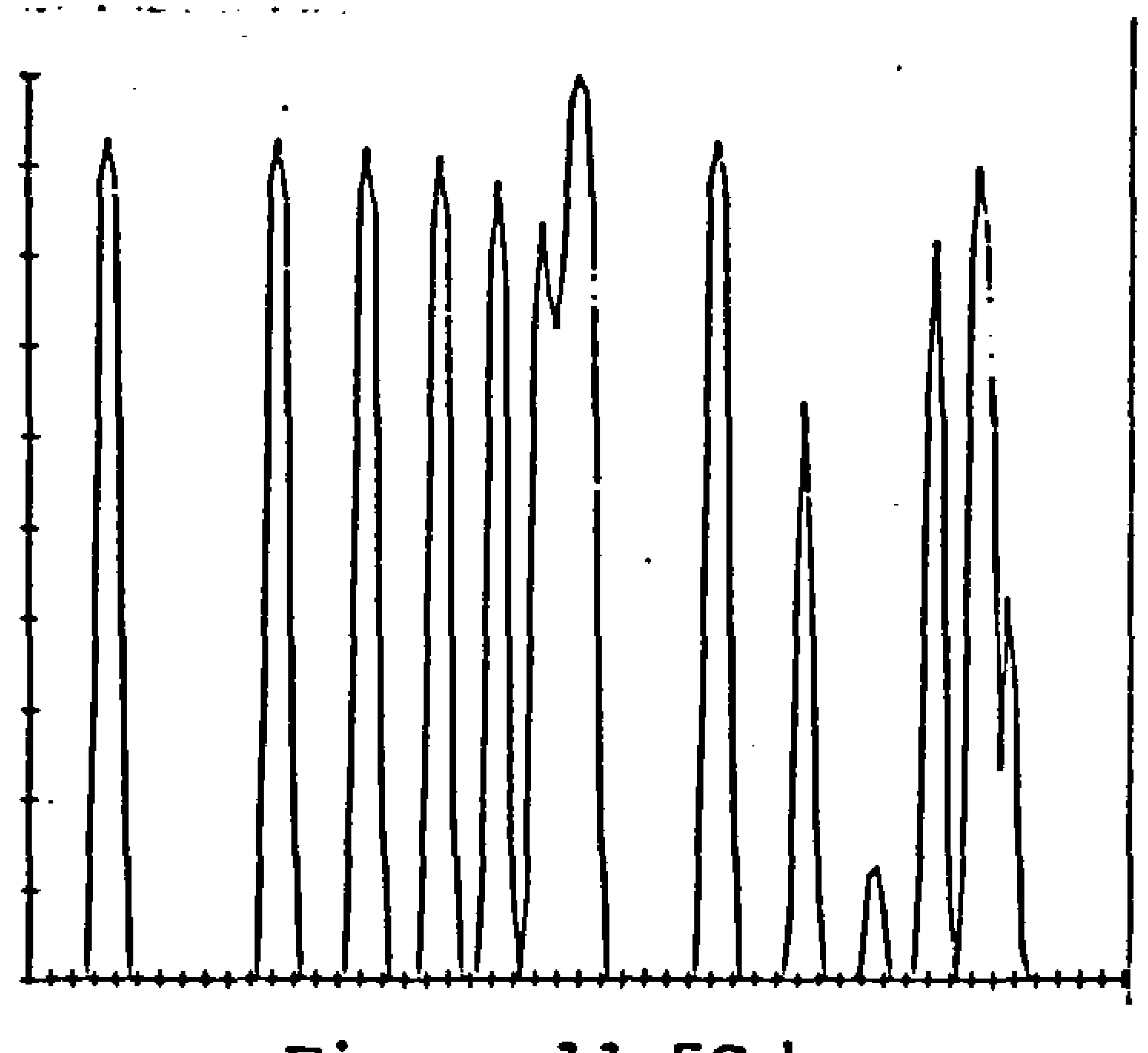


Figure 11.5C.b

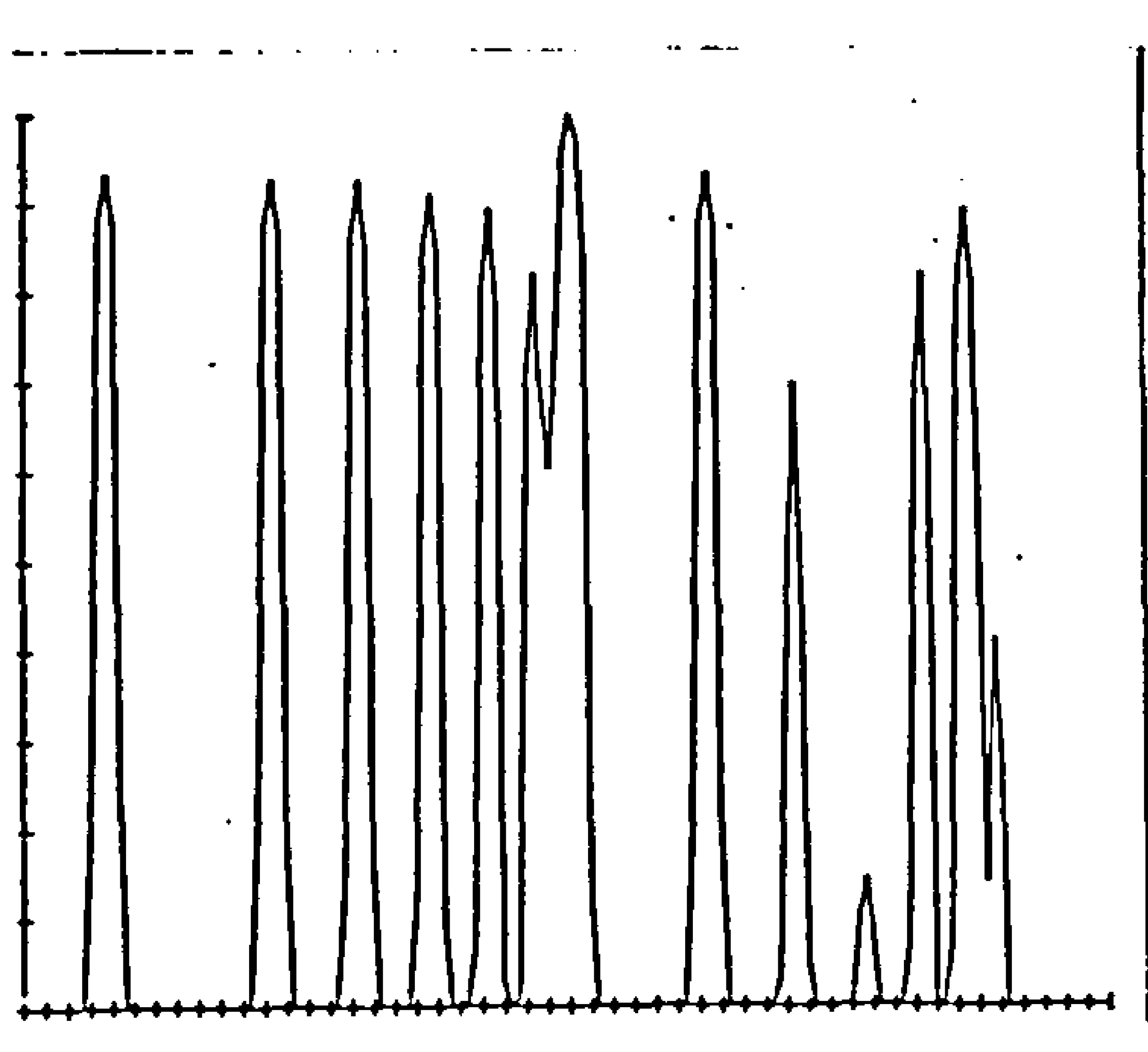


Figure 11.5C.c

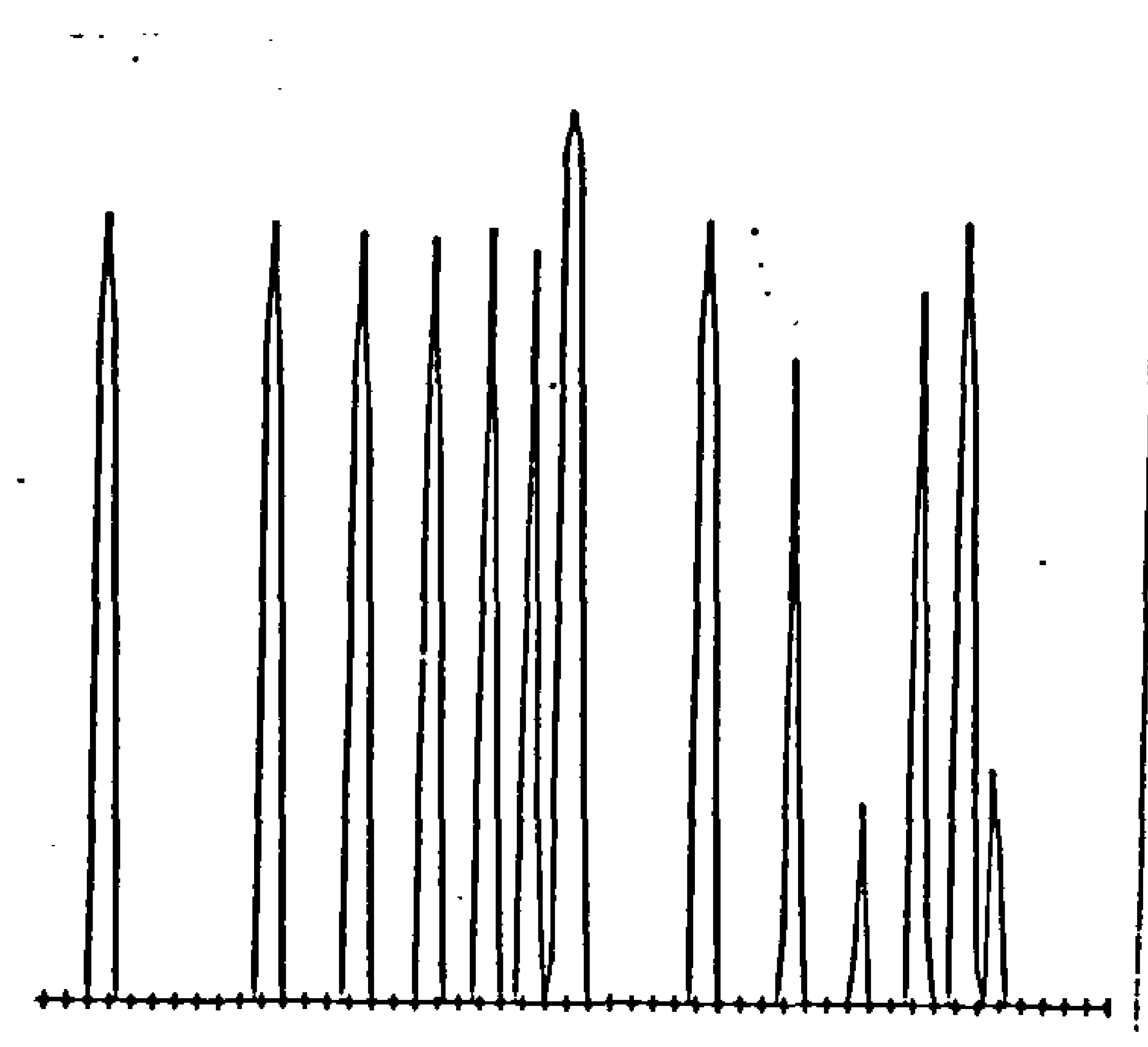


Figure 11.5C.d

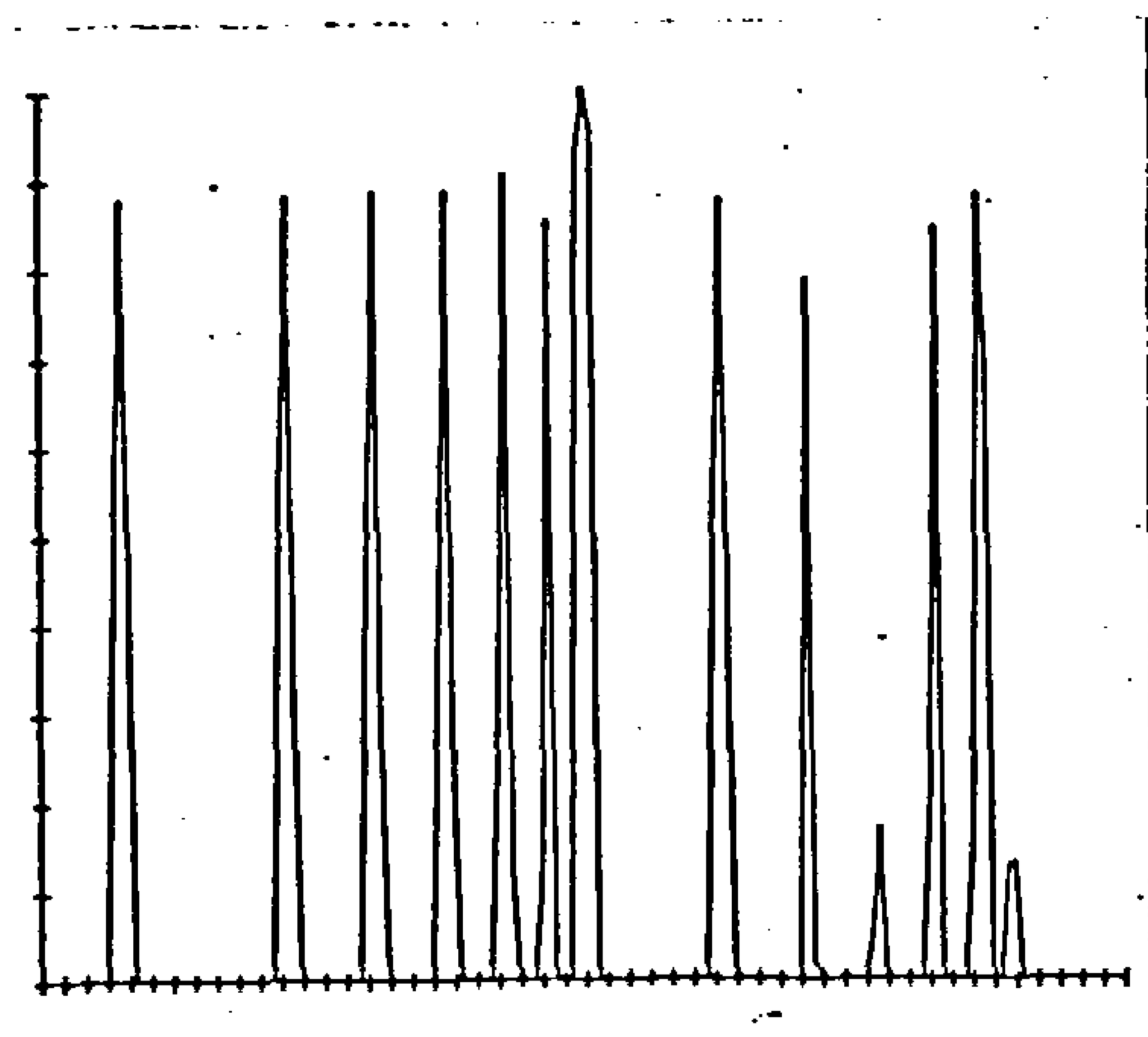


Figure 11.5C.e

Fig.11.5Ce	After	1 iteration.
Fig.11.5Cb	After	5 iterations.
Fig.11.5Cc	After	10 iterations.
Fig.11.5Cd	After	50 iterations.
Fig.11.5Ce	After	100 iterations.

Figure 11.5C Results obtained by the relaxation function C.

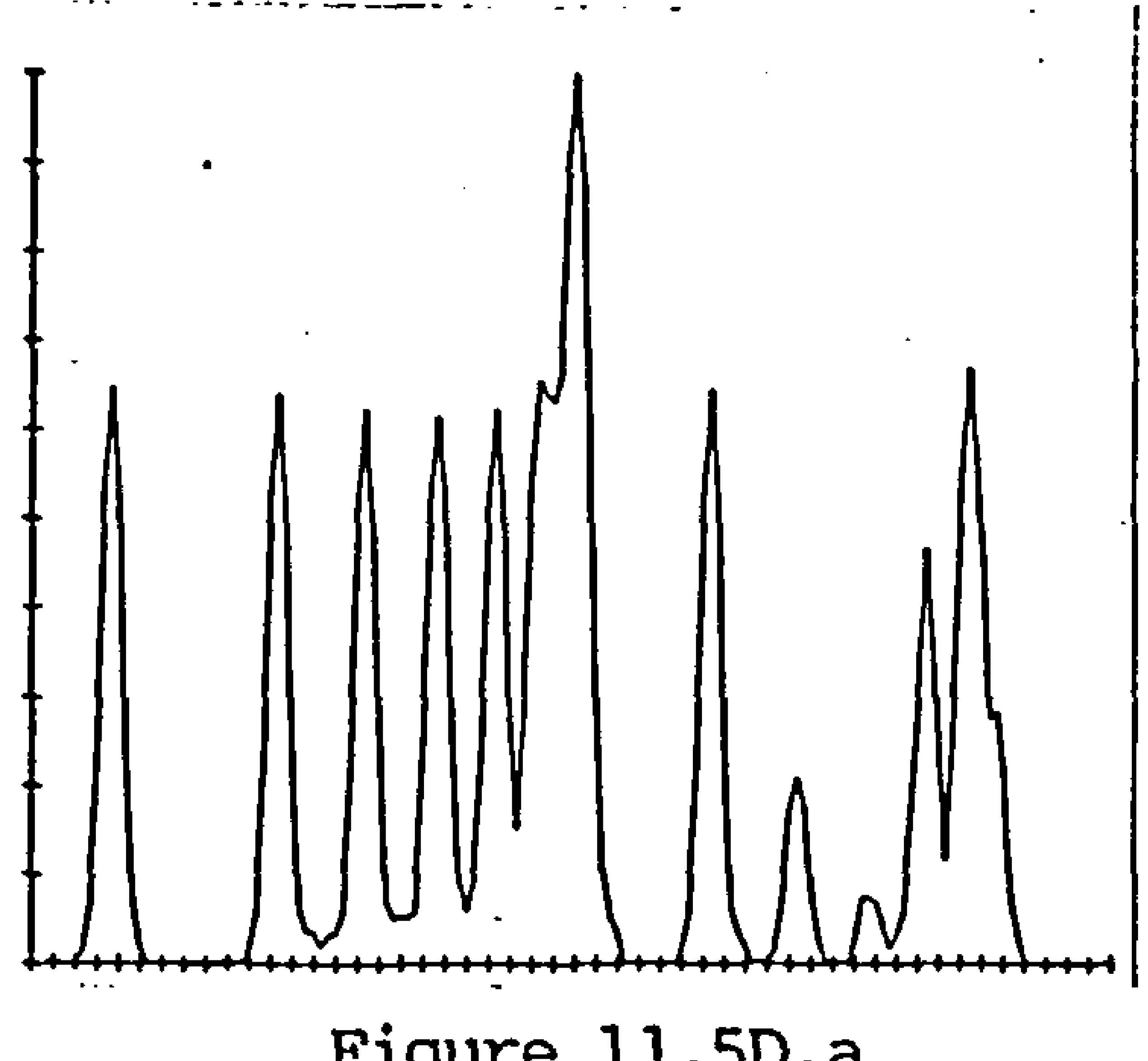


Figure 11.5D.a

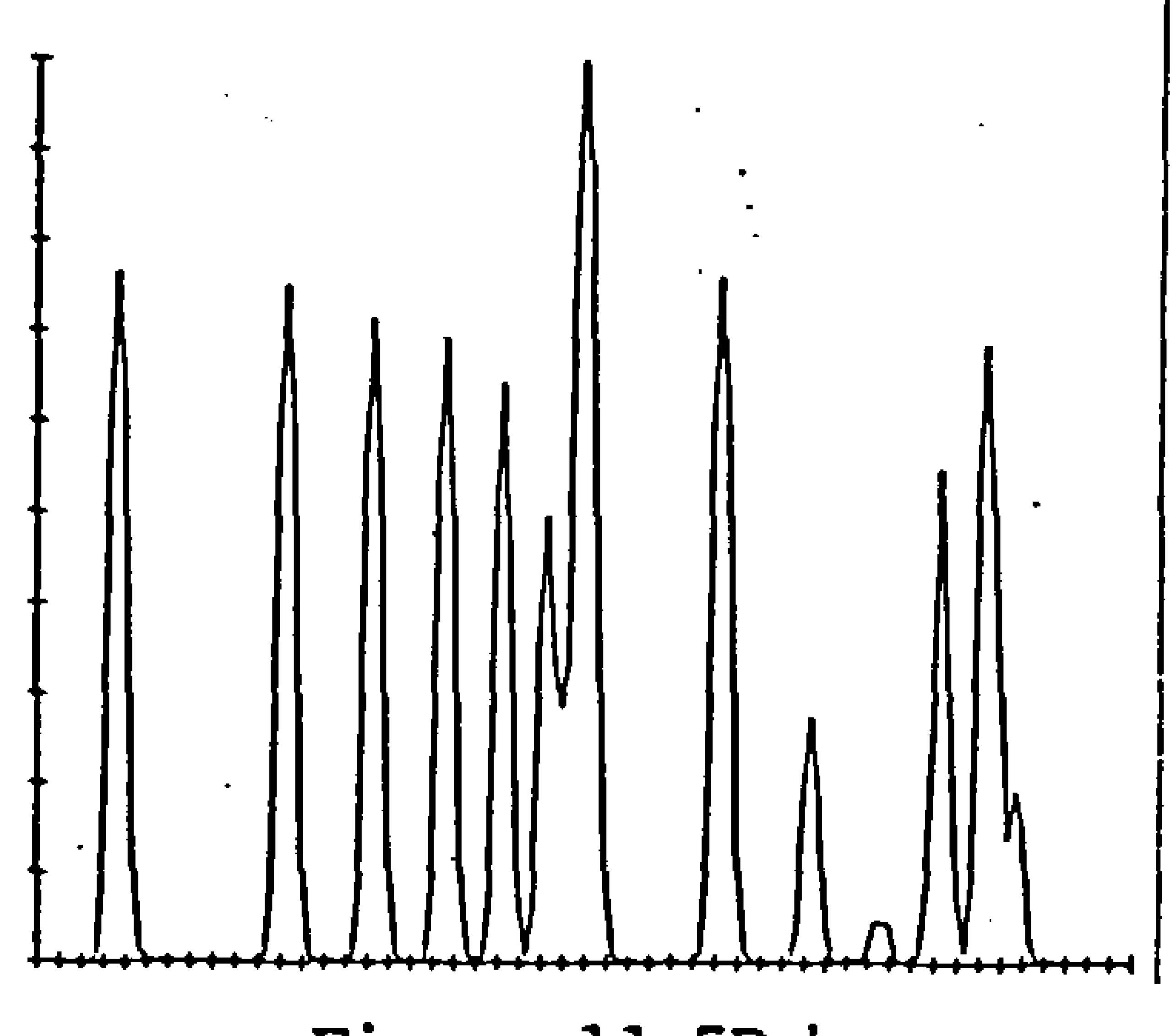


Figure 11.5D.b

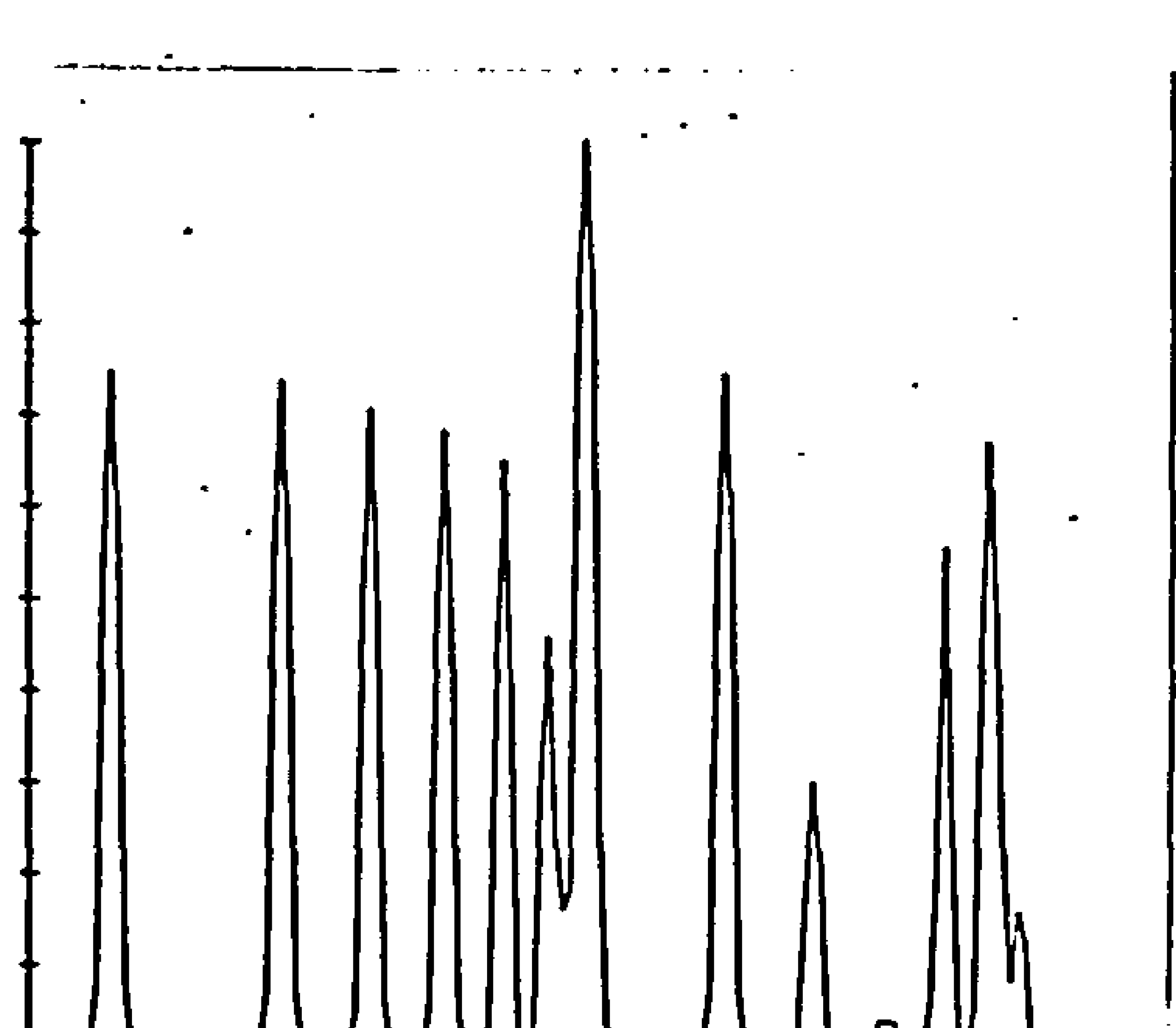


Figure 11.5D.c

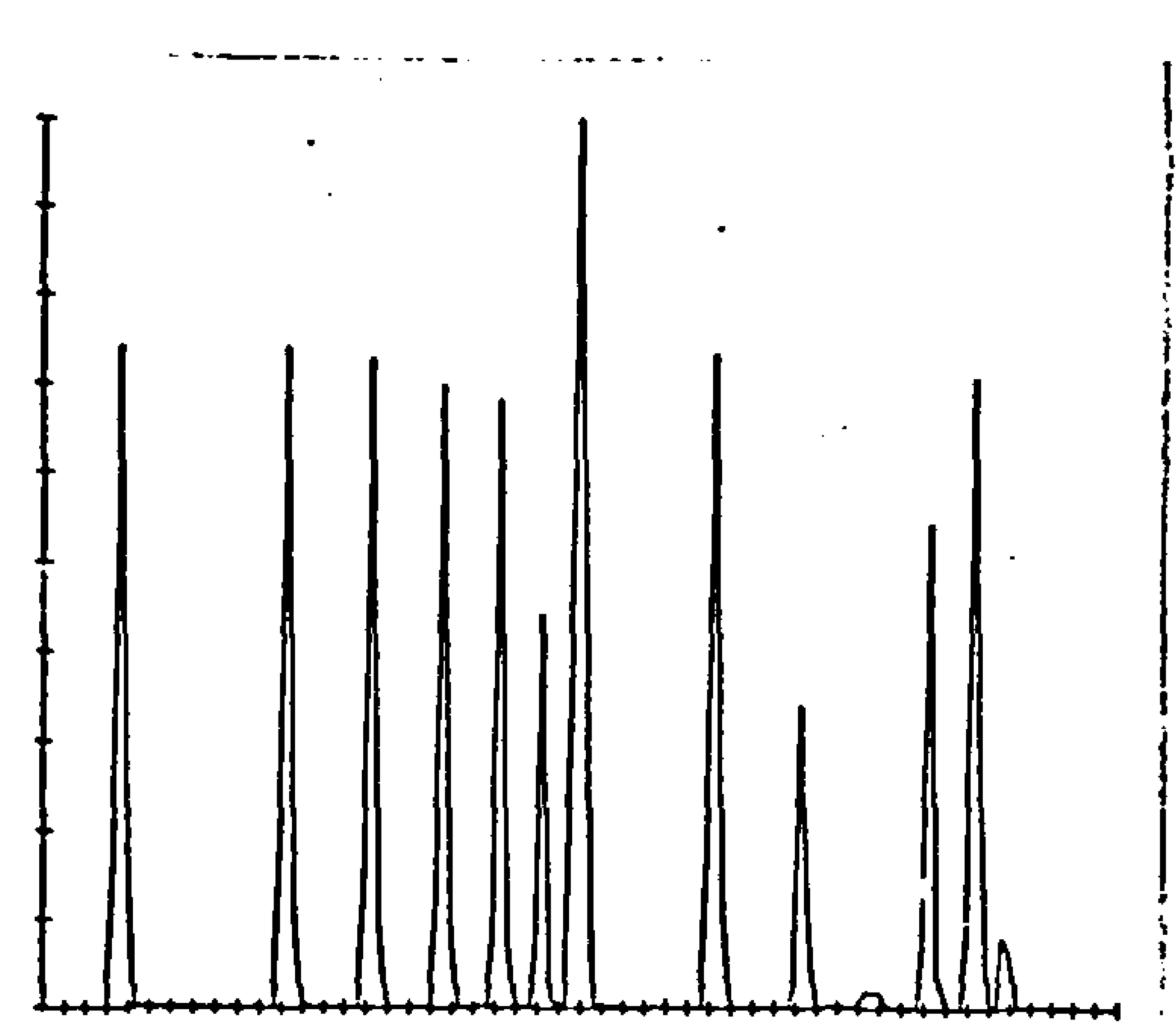


Figure 11.5D.d

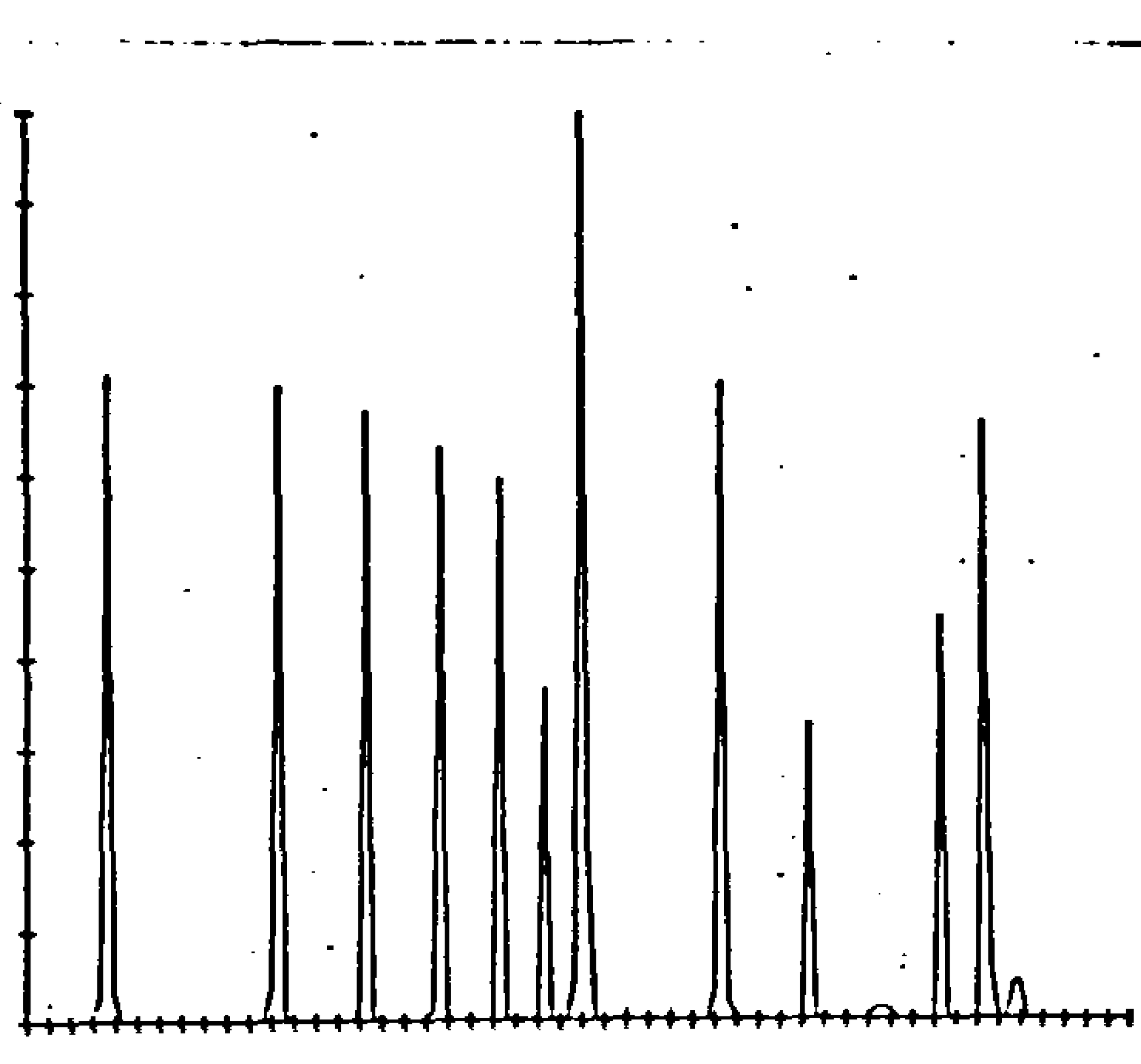


Figure 11.5D.e

Fig.11.5De After 1 iteration.

Fig.11.5Db After 5 iterations.

Fig.11.5Dc After 10 iterations.

Fig.11.5Dd After 50 iterations.

Fig.11.5De After 100 iterations.

Figure 11.5D Results obtained by the relaxation function D.

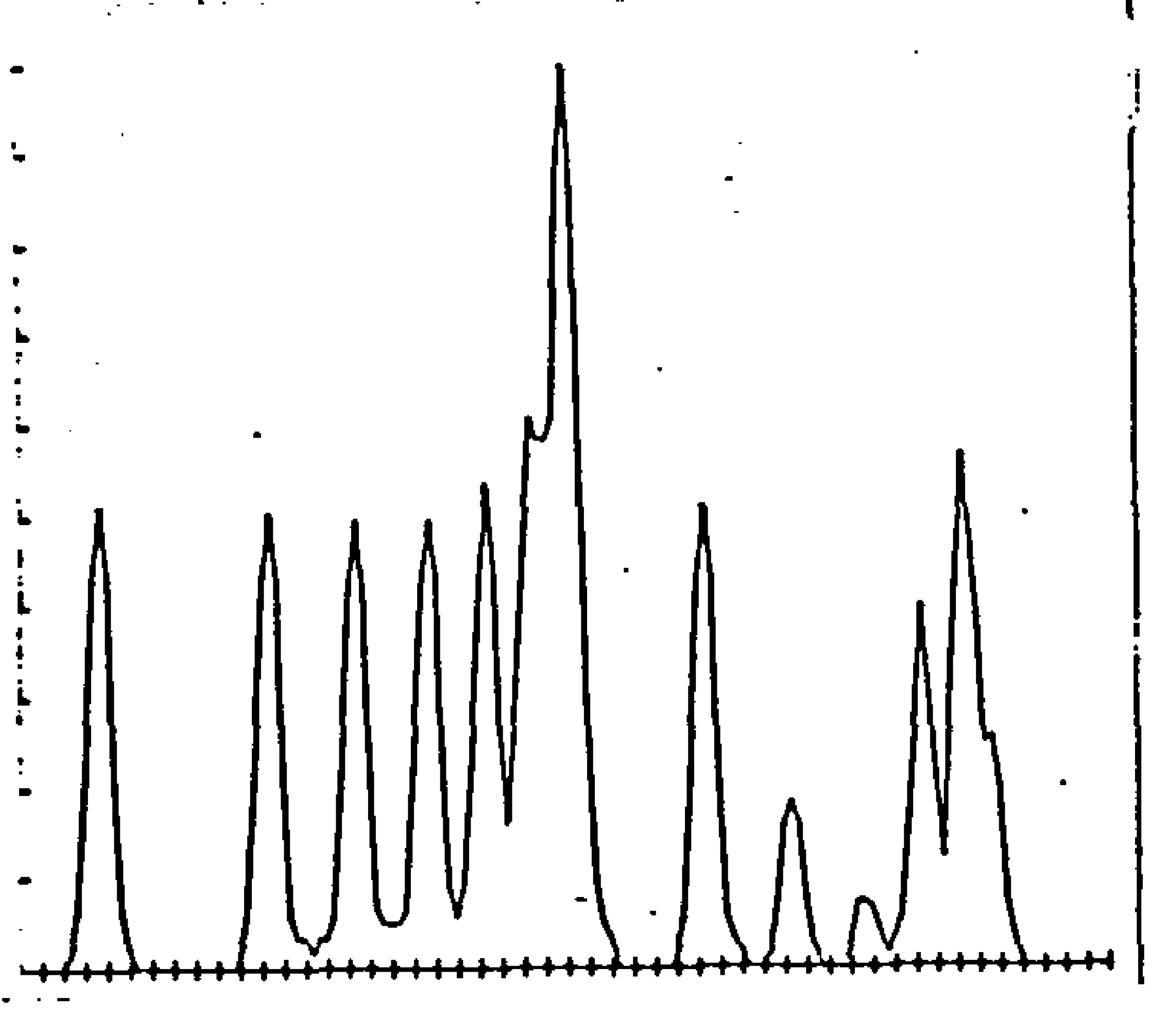


Figure 11.5E.a

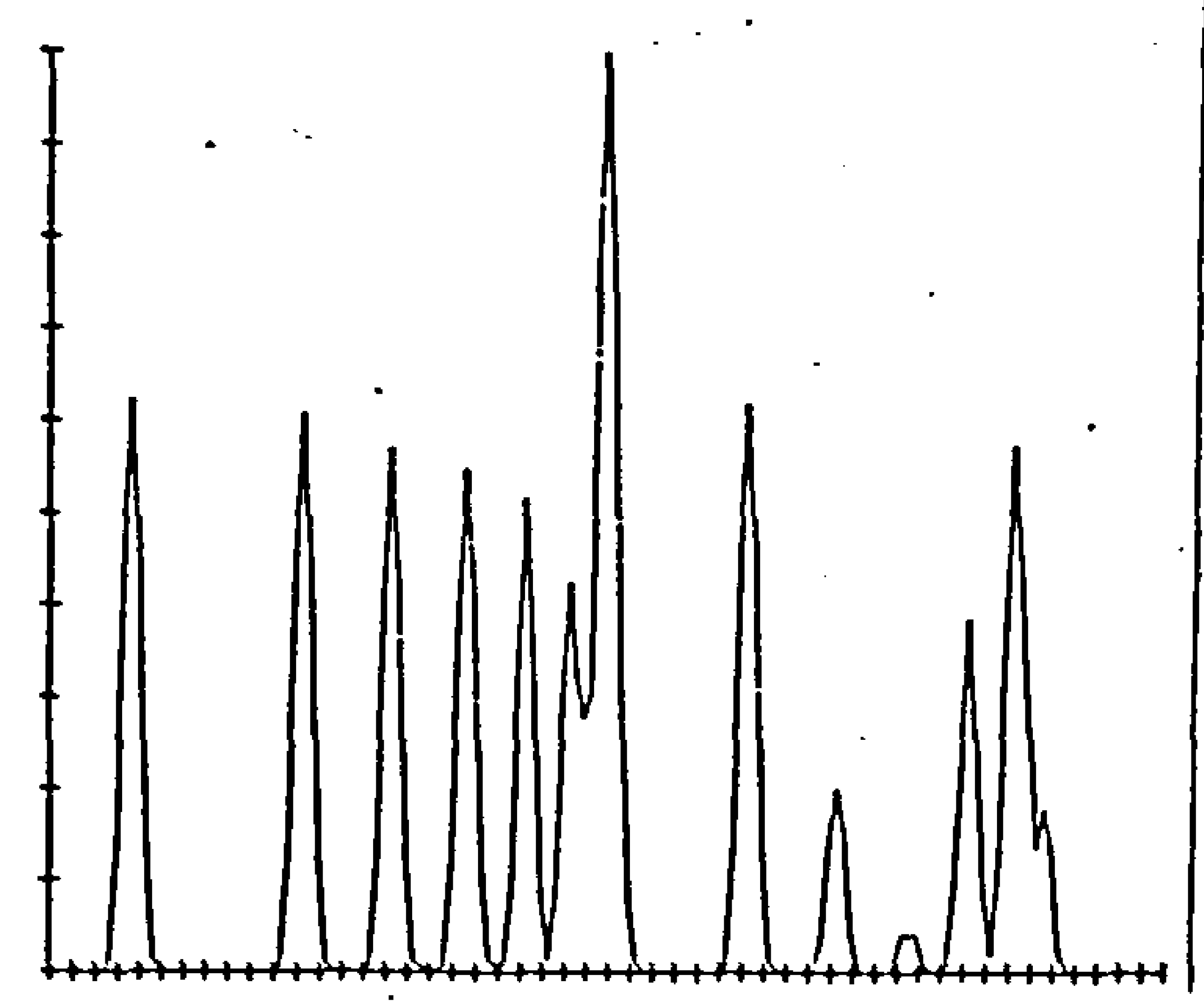


Figure 11.5E.b

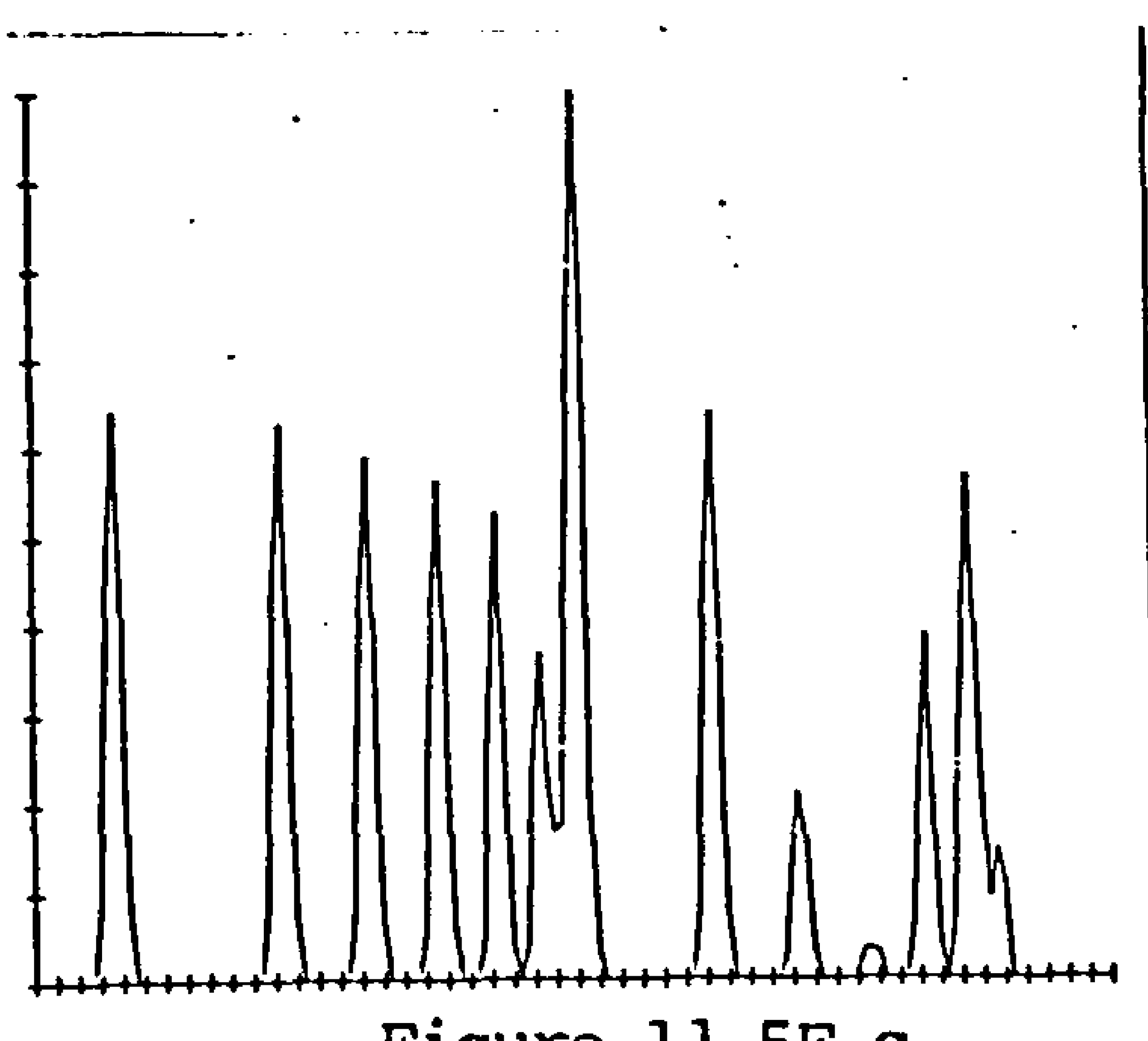


Figure 11.5E.c

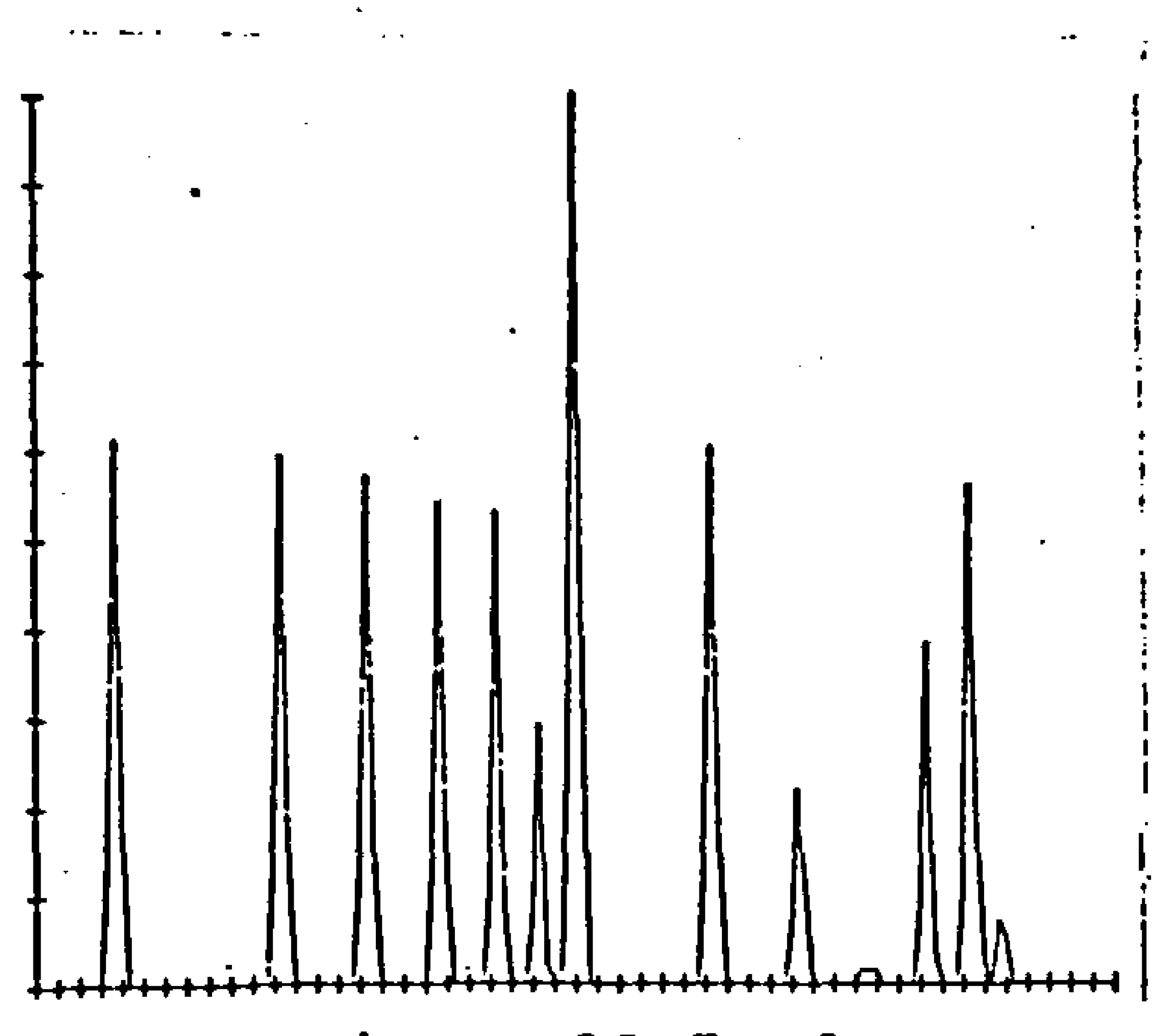


Figure 11.5E.d

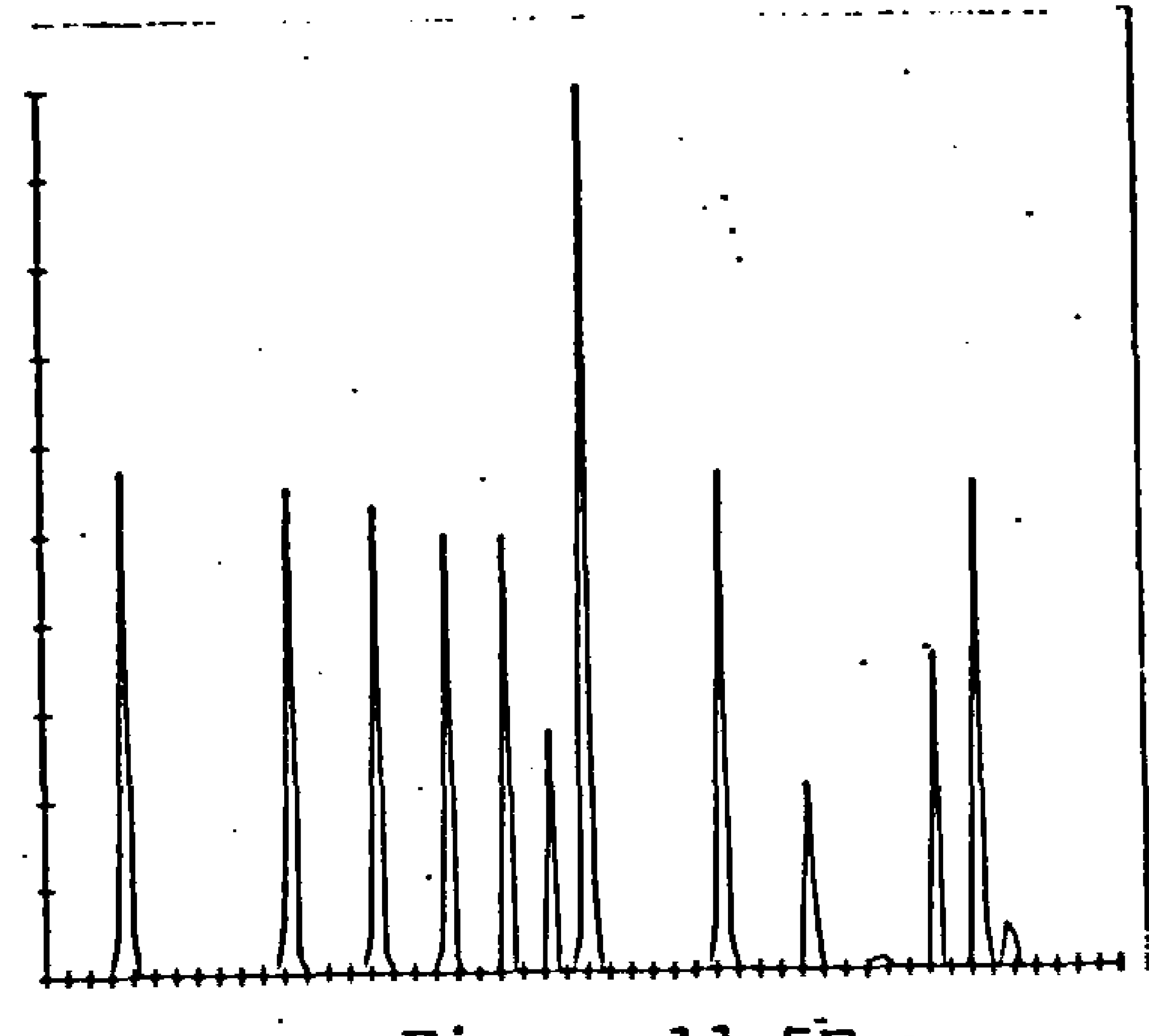


Figure 11.5E.e

Fig.11.5Ee After 1 iteration.

Fig.11.5Eb After 5 iterations.

Fig.11.5Ec After 10 iterations.

Fig.11.5Ed After 50 iterations.

Fig.11.5Ee After 100 iterations.

Figure 11.5E Results obtained by the relaxation function E.

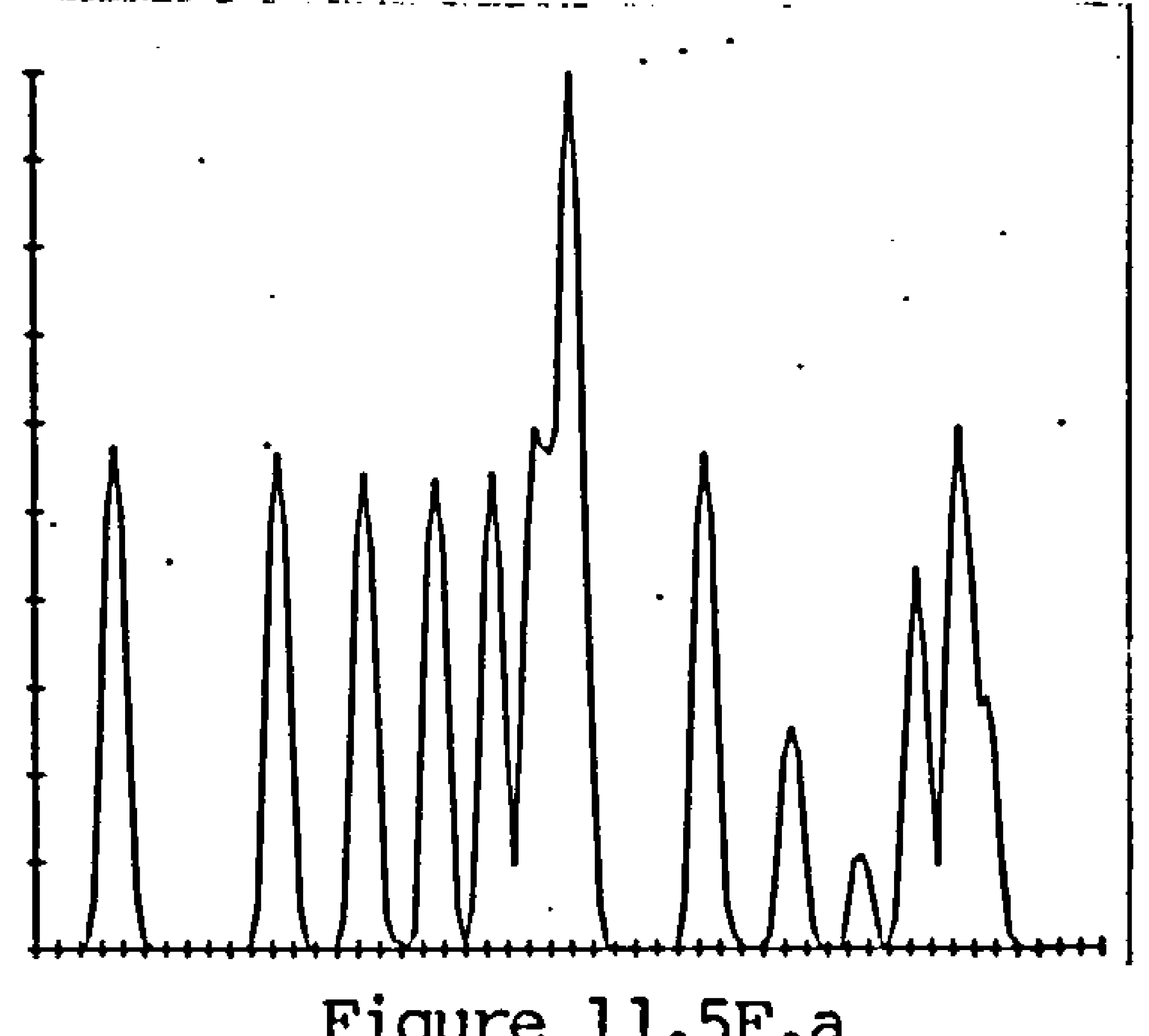


Figure 11.5F.a

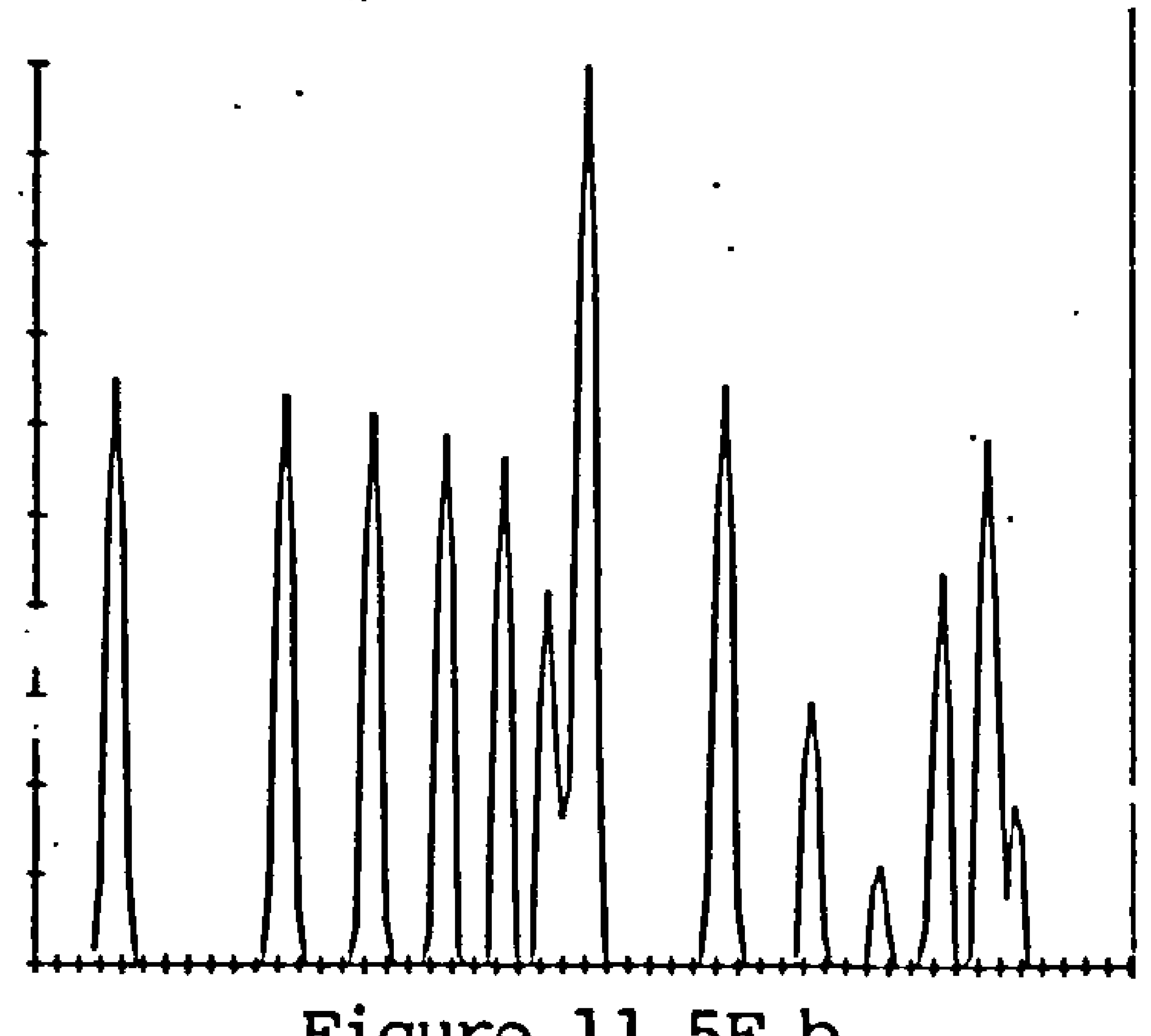


Figure 11.5F.b

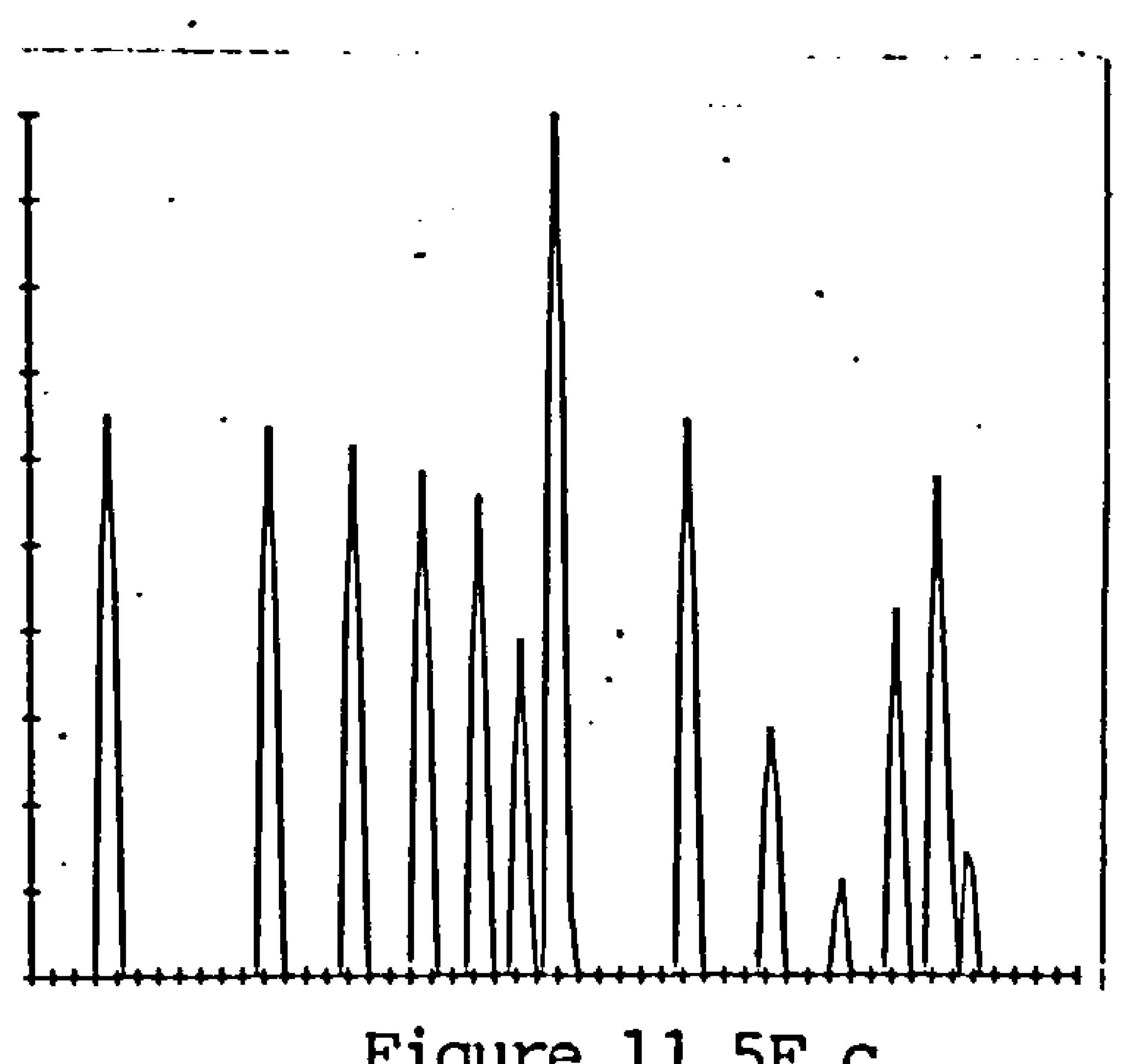


Figure 11.5F.c

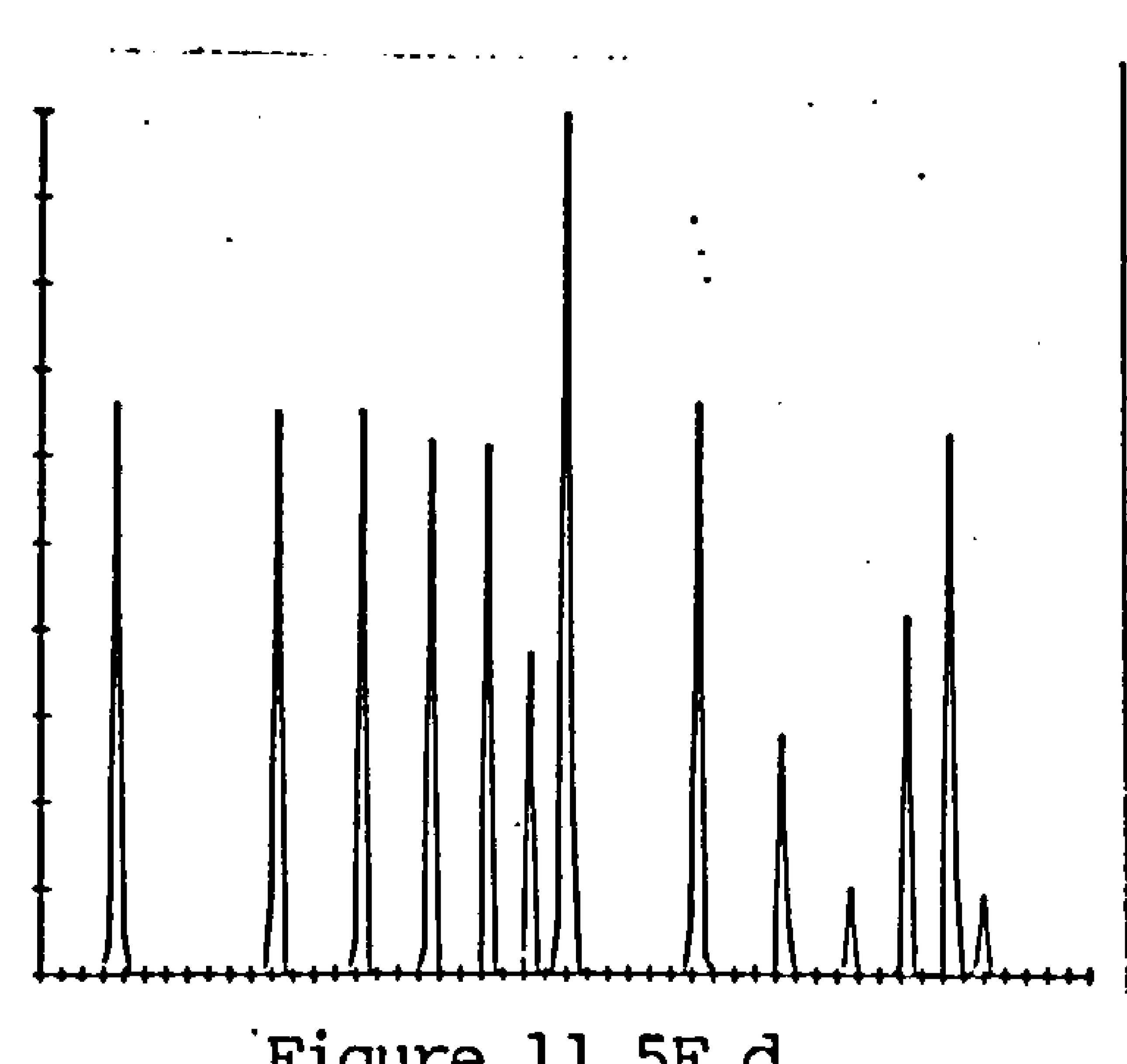


Figure 11.5F.d

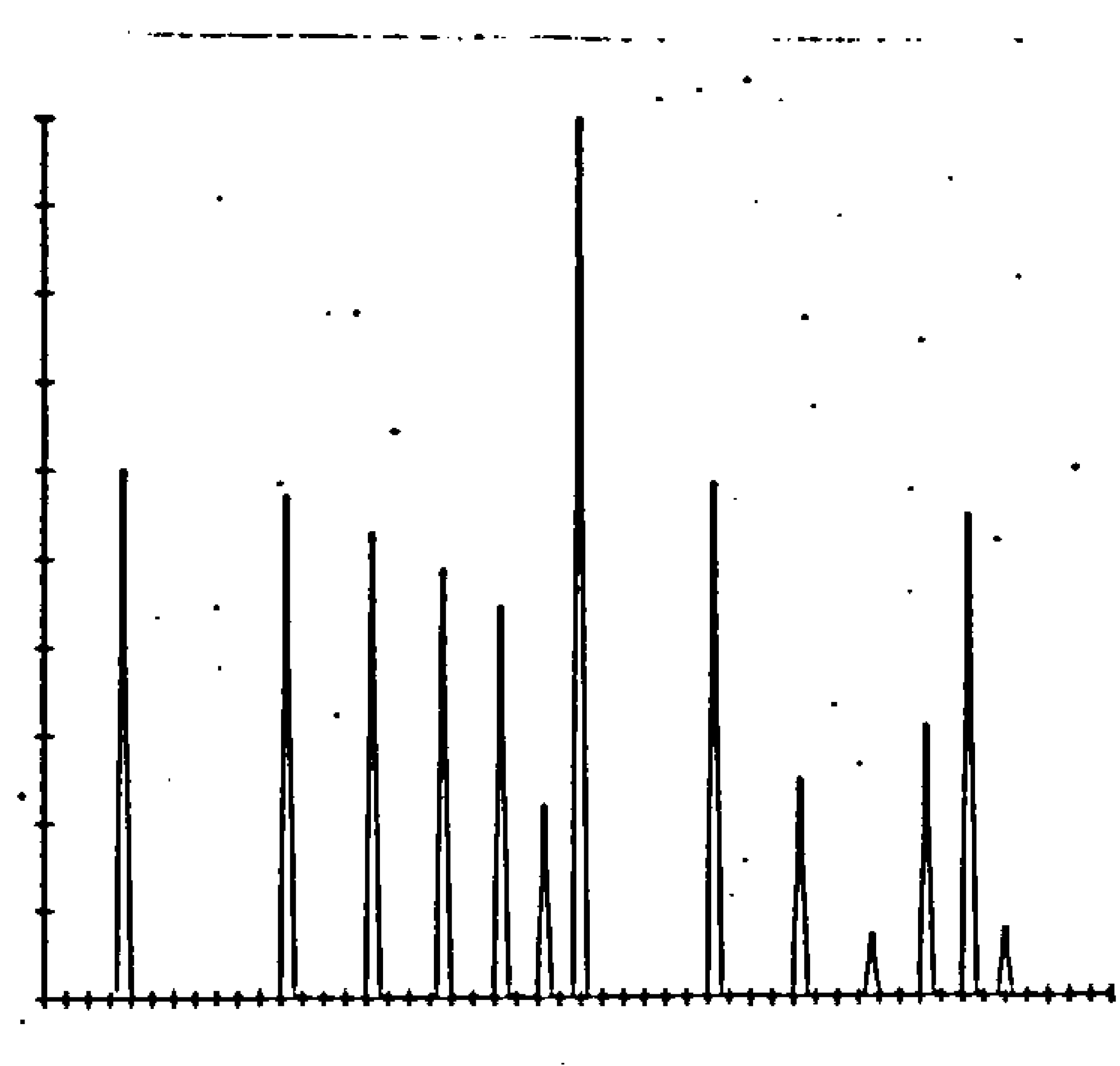


Figure 11.5F.e

Fig.11.5Fe After 1 iteration.
Fig.11.5Fb After 5 iterations.
Fig.11.5Fc After 10 iterations.
Fig.11.5Fd After 50 iterations.
Fig.11.5Fe After 100 iterations.

Figure 11.5F Results obtained by the relaxation function F.

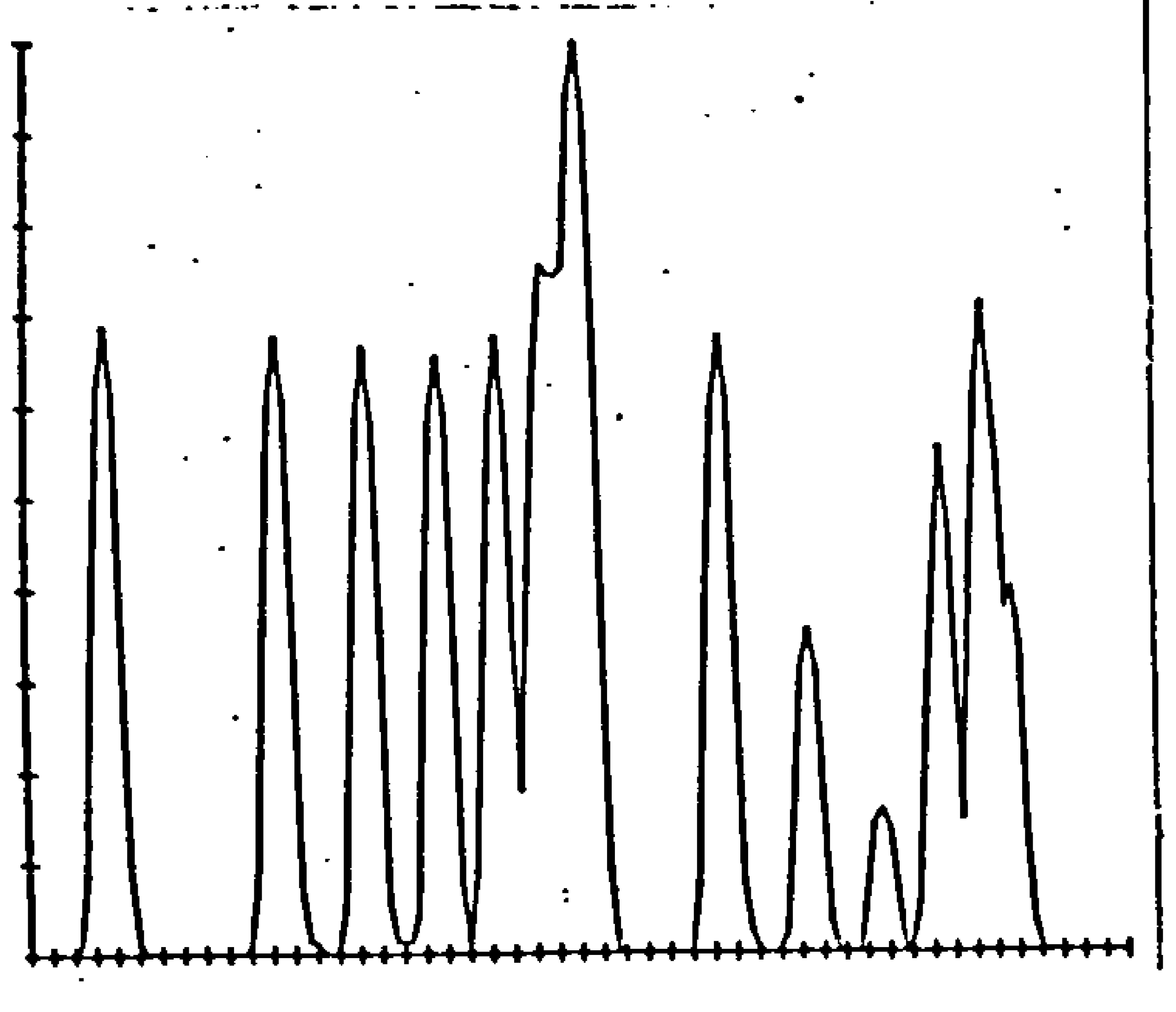


Figure 11.5G.a

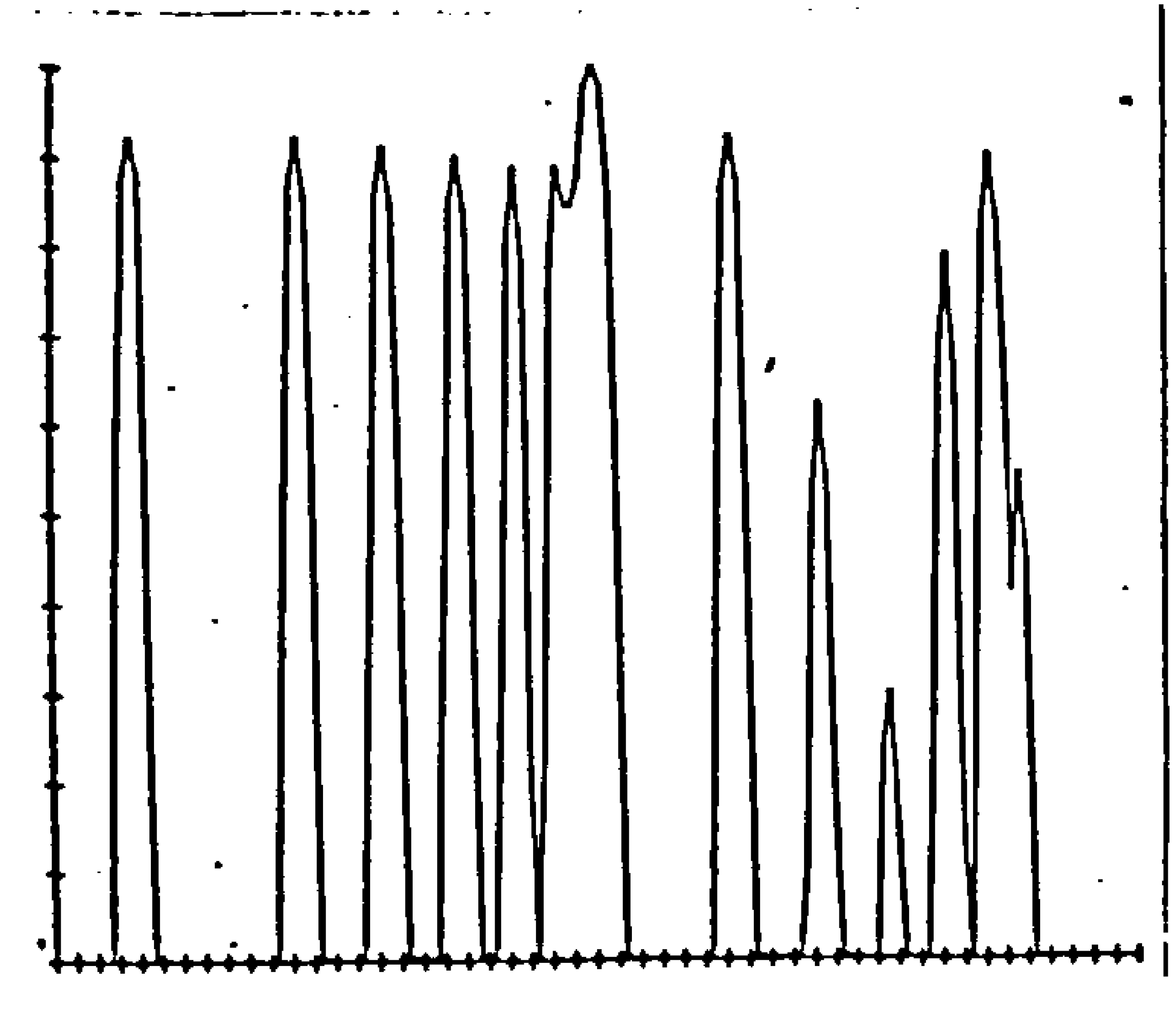


Figure 11.5G.b

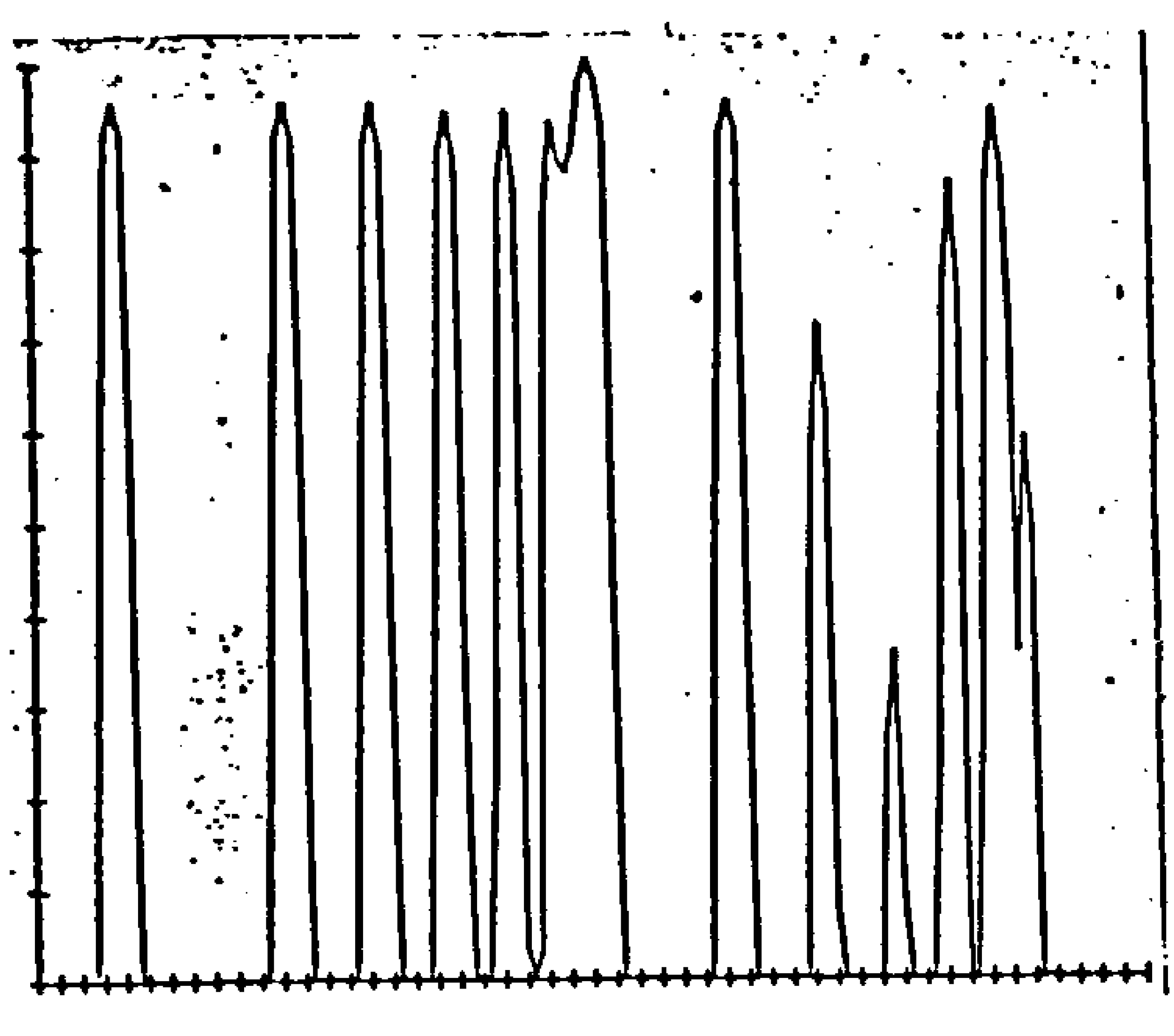


Figure 11.5G.c

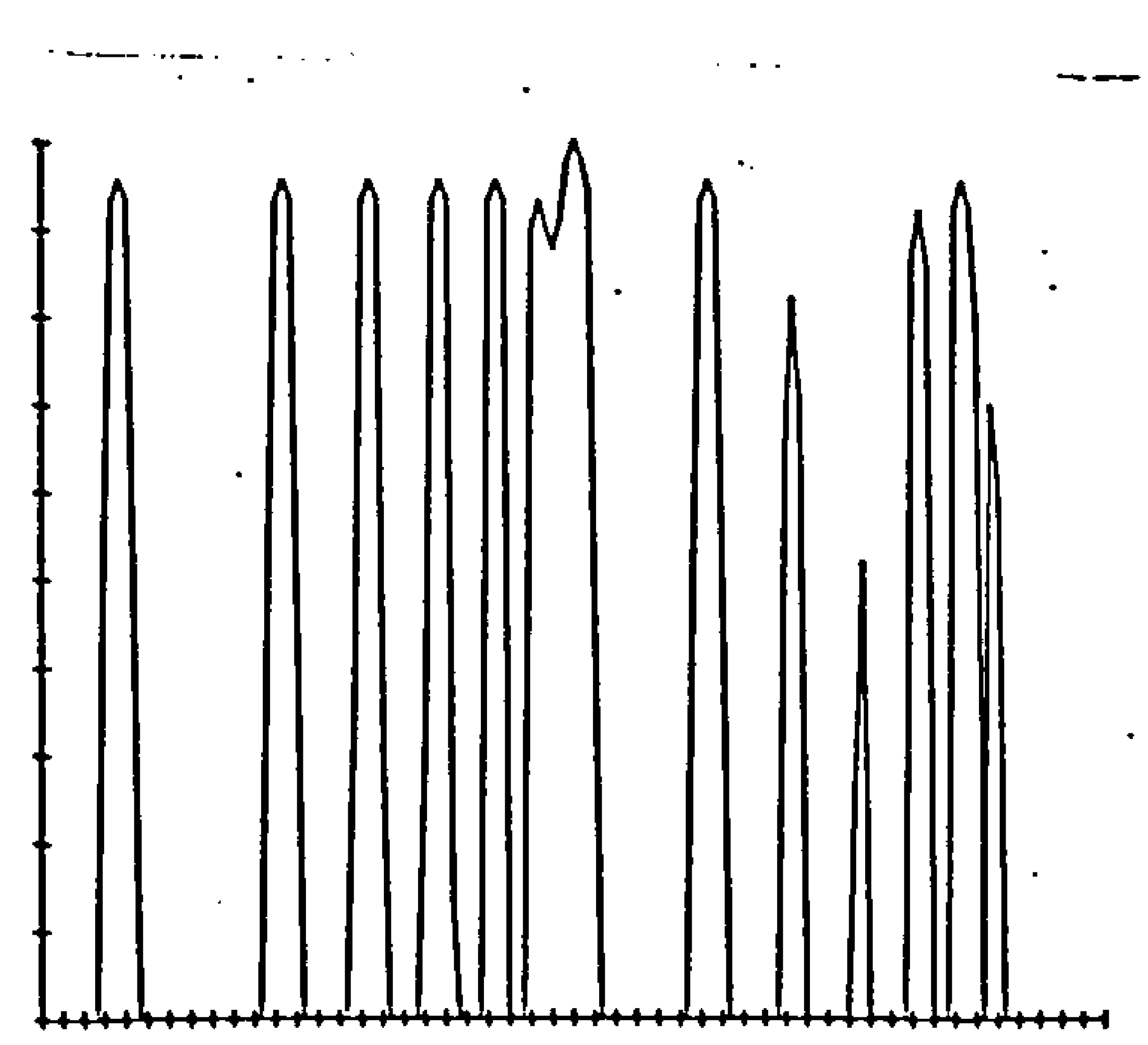


Figure 11.5G.d

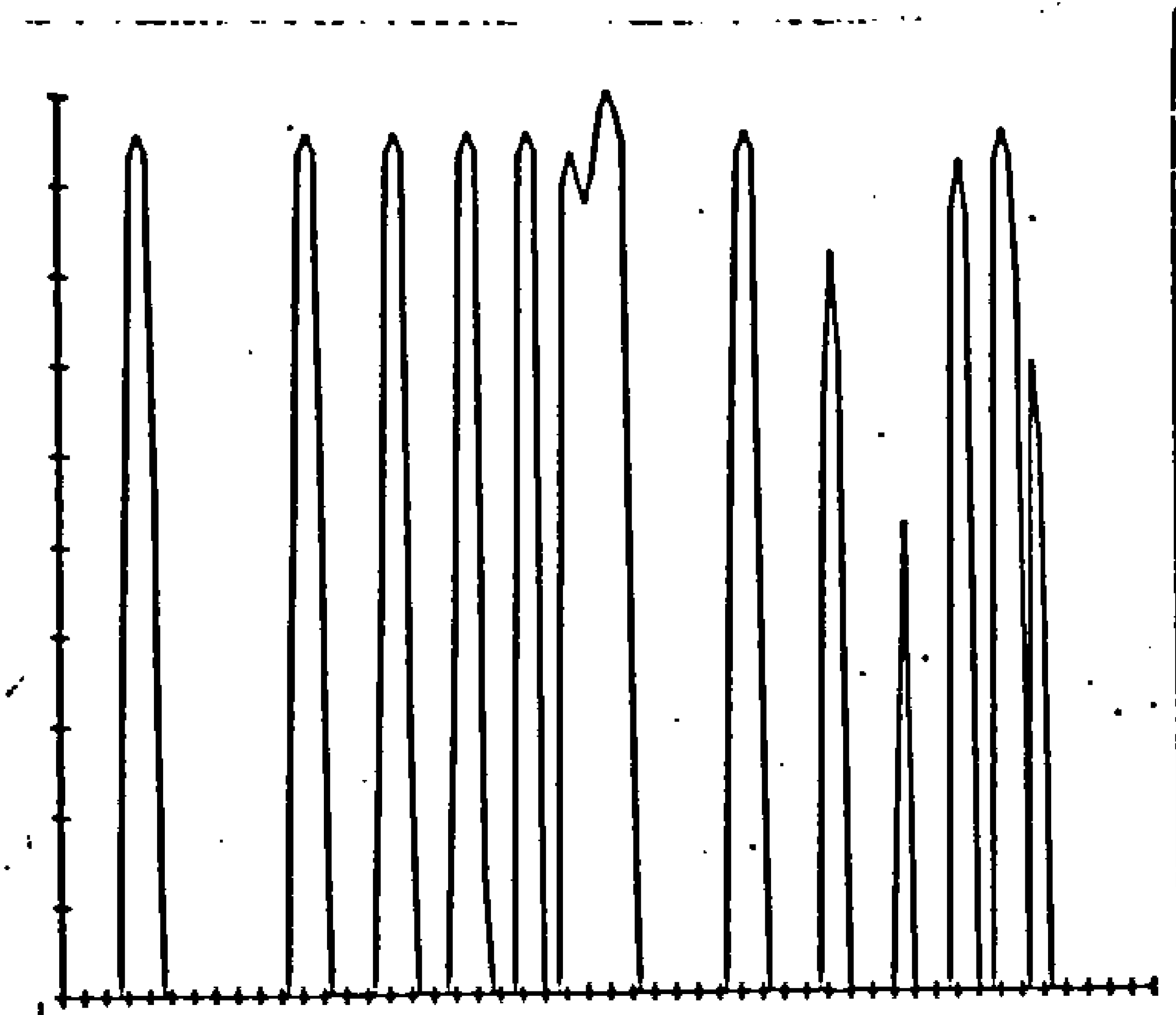


Figure 11.5G.e

Fig.11.5Ge After 1 iteration.

Fig.11.5Gb After 5 iterations.

Fig.11.5Gc After 10 iterations.

Fig.11.5Gd After 50 iterations.

Fig.11.5Ge After 100 iterations.

Figure 11.5G Results obtained by the relaxation function G.

11.5. Discussion of the results.

The results obtained by the deconvolution algorithms are presented in the figures 11.5A-G. Both the blurring function and the convolved input signal were once more convolved with the blurring function prior to applying the deconvolution algorithm. They show the results after respectively 1,5,10,50 and 100 iterations.

In general, not much improvement is achieved after 50 iterations. Actually a distortion of the relative amplitudes can be the result of not terminating the deconvolution in time. So it is essential to monitor the deconvolution closely by checking the convergence by say the root-mean-square error. Also a strong intensity peak seems to weaken a smaller neighbouring peak.

The results obtained by the relaxation functions A, B and G are very similar. All three fail to resolve the sixth peak (position 71). That the results produced by A and B do not differ much is not too surprising, since their relaxation functions are very similar. On the other hand, they also resemble very closely the function C, which actually resolves the sixth peak.

From the results obtained by the relaxation functions C, D, E and F it can be concluded that F is the most efficient one. All four are capable of resolving the sixth peak, but F resolves it completely after 9 iterations. By completely resolved is meant that the neighbouring peaks are completely detached. Furthermore, it restores the relative amplitudes better than the other relaxation functions.

So two diffraction peaks can be completely resolved when their mutual distance is at least equal to the distance between the centres

of four consecutive photo-diodes.

According to Lord Rayleigh's criterion for resolving equal irradiant overlapping fringes, they are just resolvable when the combined irradiance of both fringes at the centre, or saddle point, of the resultant broad fringe is $8/\pi^2$ times the maximum irradiance. Application of this criterion to the blurring function (figure 11.3) results in a minimum resolvable separation equal to the distance between the centres of five consecutive photo-diodes.

11.6. Conclusion.

Deconvolution techniques can improve the resolution by at least 20%. A greater improvement might be feasible by changing interactively the value of r_o during the deconvolution process.

It is important to monitor the convergence of the deconvolution processs, otherwise gross errors can occur in the estimation of the amplitudes of the diffraction peaks.

The deconvolution algorithm based upon the relaxation function F has not only the advantage that it is fast, but it is also relatively easy to implement in machine code into digital computers.

12. Practical embodiment.

A system has been designed and constructed for the acquisition and further processing of optical spectra from turned surfaces. The system consists of two main units.

i a sensor-head, containing a collimated laser diode, Fourier optics, a linear photo-diode array and associated drive electronics.

ii a BBC computer, containing a purpose built 7-bit A/D converter and an EPROM containing the software for controlling the sensor-head and the A/D converter and to analyse the optical information from the sensor-head.

The computer and sensor-head are interfaced by a purpose built A/D converter, which is connected to the parallel input/output port of the BBC computer. The BBC's own A/D converter has not been used since it is rather slow (4 msec per 8 bit conversion) and it suffers from an inaccurate conversion due to an unstable internal voltage reference.

12.1. The sensor-head.

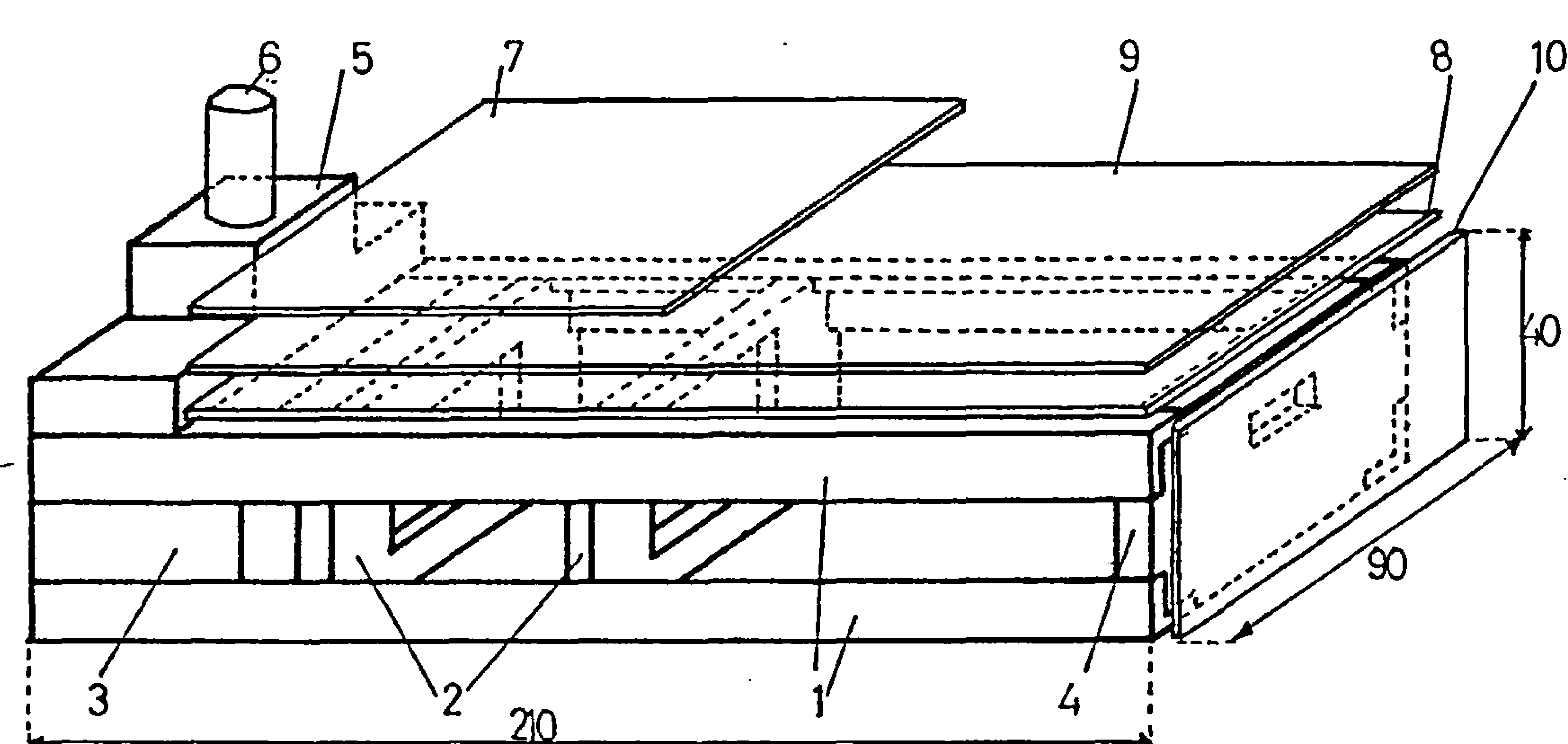
Apart from two power supply units, the three main components in the sensor-head are:

i a collimated laser diode.

ii the optics, consisting of a beam-splitter plate and two cylindrical lenses.

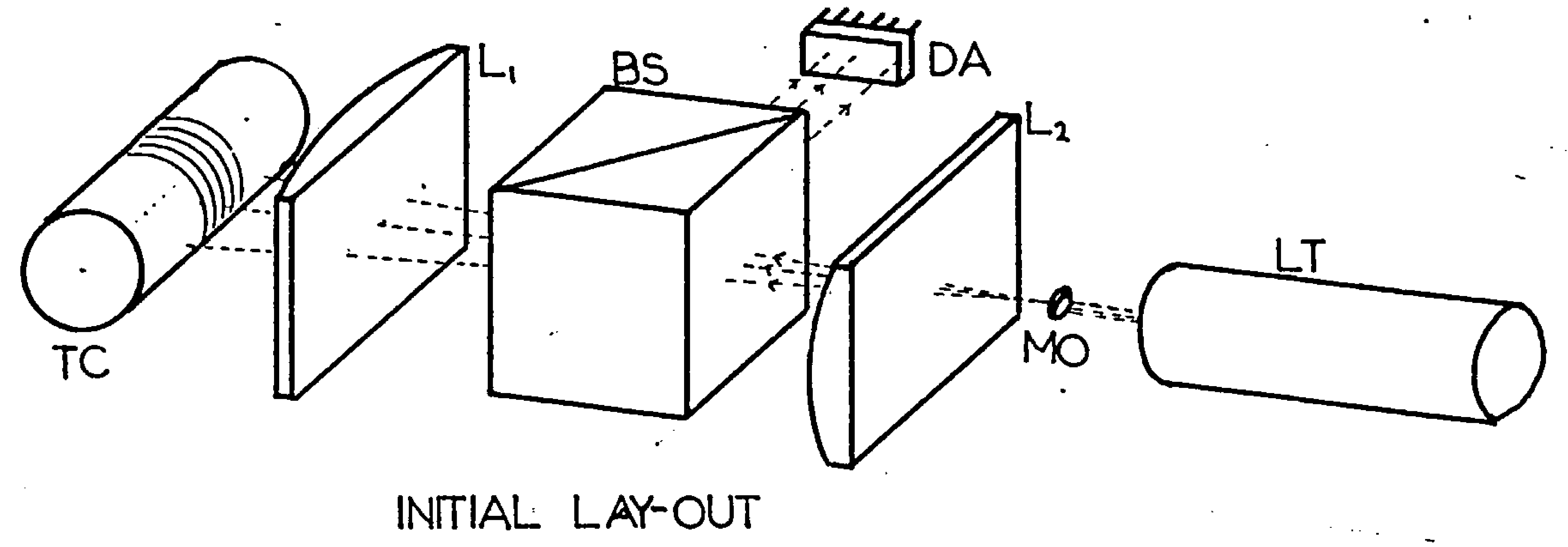
iii a linear photo-diode array, with drive electronics.

These components are mounted on and in a rigid aluminium frame (figure 12.1), which is contained in an ABS box together with two power supplies and a fan. The purpose of the fan is to create an overpressure inside the box so as to keep the optics free from dust. All the components, except for the frame, are commercially available items.



Mechanical components.	Opto-electronic components.
1 aluminium rails	6 laser diode
2 lens mounts	7 laser diode supply board
3 beamsplitter plate mount	8 head board with diode array
4 diode array holder	9 diode array driver board
5 laser diode holder	10 diode array processor board

Figure 12.1. Physical lay-out of the sensor head.



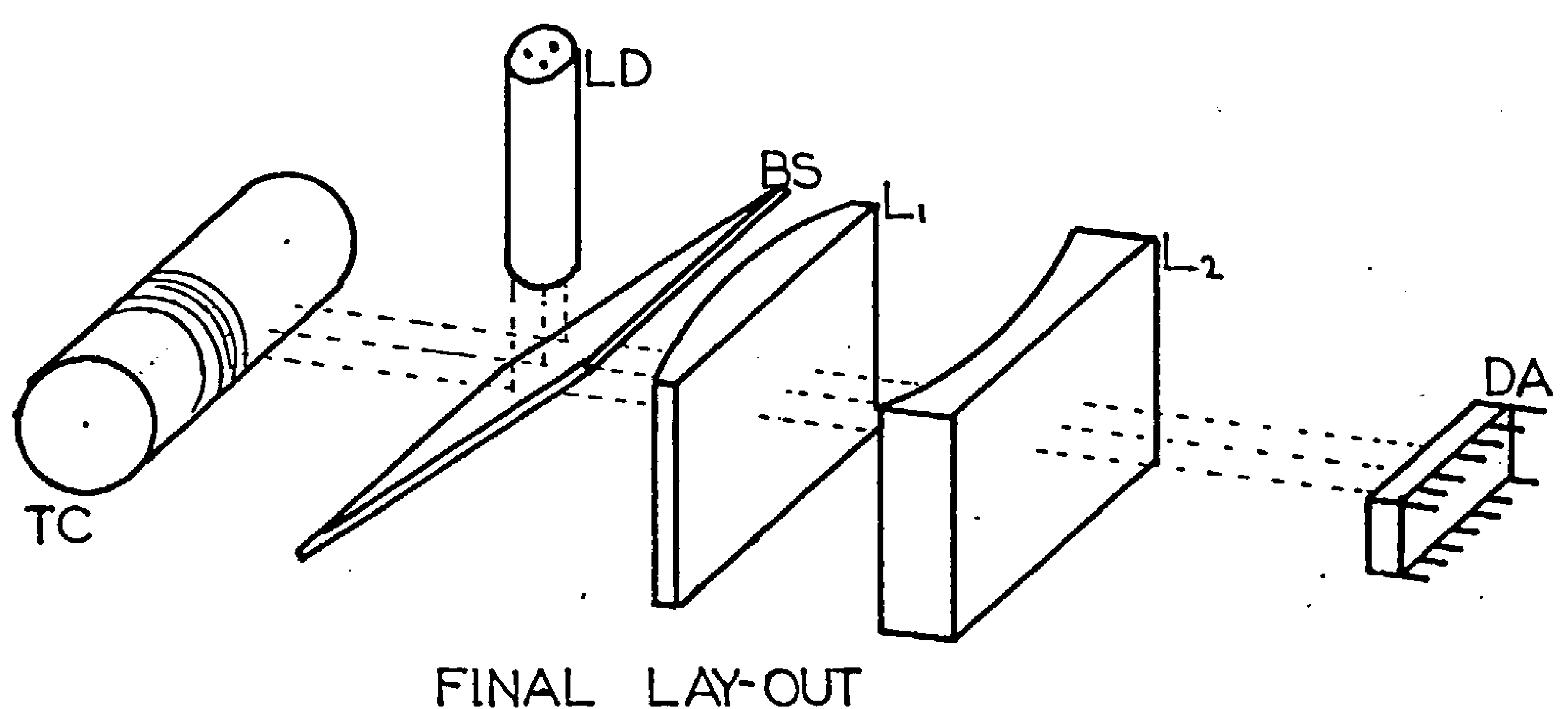


Figure 12.2. Top: Initial lay-out of the transform system with a laser tube (LT).

Bottom: Final lay-out with a collimated laser diode (LD).

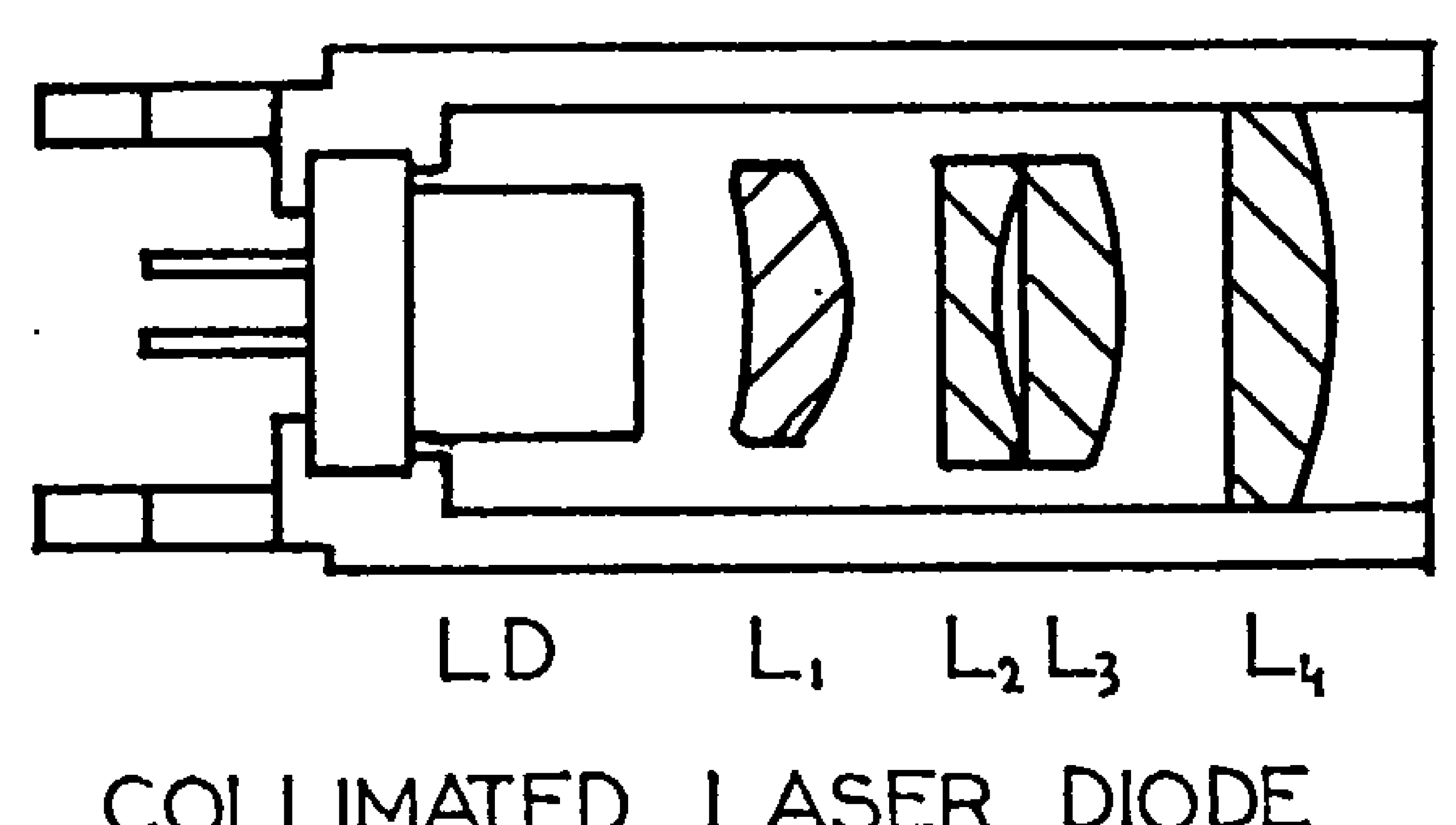


Figure 12.3. Collimated laser diode. L1 is a cylindrical lens to correct the astigmatism of laser diode LD. L2-4 are spherical lenses.

12.1.1 The laser diode.

Initially a Helium Neon laser tube was employed to provide the coherent optical wavefront. In the final design this was substituted by a collimated laser diode (MULLARD CQL13A) with a continuous output, see figure 12.2.

Since the output beam of a laser diode is both highly divergent and elliptical in cross-section, special collimating optics are necessary to produce a collimated output beam. This collimation is achieved by four lenses fitted on a cylindrical housing, which also contains the laser diode, see figure 12.3. The lens (L_1) nearest to the to the laser diode (LD) is a cylindrical lens, correcting the astigmatic beam produced by the laser diode. The lenses L_{2-4} are spherical lenses for collimation.

Although the laser diode can electrically withstand current surges of several amperes, optically the laser is more susceptible to damage because an extremely high optical flux density passes through both facets while in operation. The optical flux density can rise to unacceptable values (10 to 100MW/cm^2), if the laser diode is overdriven or subjected to transients. This will result in gradual or catastrophic degradation of the laser facets. So close control of the optical output is necessary. To this purpose an optoelectronic feedback system is used to control the laser output. See appendix A6 for the control system.

Compared with a He-Ne laser, the laser diode has the following advantages:

i The laser diode, together with its collimating optics, measures 34X11ø mm. A He-Ne laser with comparable output (2mW) measures 300X42ø mm.

ii The diode laser power supply is smaller in size and there are no high voltages in the system.

iii The diode laser is very rugged, because it is a solid state device.

iv The laser diode is supplied with its own collimating optics, (figure 12.3) delivering a beam with the correct diameter (5.4mm). A He-Ne laser would need an external beam expander.

v The light emitted by the diode laser has a longer wavelength (870nm) than the He-Ne laser (633nm). This increases the roughness range of the surfaces which can be measured.

The collimated laser diode has the following characteristics:

output power	3 mW
wavelength	875±10 nm
bandwidth	4 nm
wavelength temperature coeff.	.25 nm/K
beam diameter	5.4mm
non convergent plane wavefront	
with a maximum divergence	0.3 mrad

Further information on the laser diode can be found in the data sheets in appendix A7.

12.1.2 The photo-diode array.

The linear photo-diode array (IPL 512M5) is manufactured by Integrated Photomatrix Ltd. It contains 512 photo-diodes in a row at 0.001 inch centre to centre spacing. They are self scanned with integrated scanning circuitry included on the same chip. The peak spectral response occurs near 820nm, which makes the photo-diode array well suited to operation in conjunction with semi-conductor lasers. Further details of these chips can be found in appendix A7.

The photo-diode is operated by a M series Processor Driver Assembly (M PDA). It comprises of three printed circuit boards. The two main boards contain the logic circuitry and the video processing circuitry respectively. The smallest p.c. board is the array carrier board, which contains high level clock drivers and the video pre-amplifier. An oscillator is located on the logic board, providing the main system clock. In the actual system this clock can be disconnected and the system clock pulses have to be provided by the BBC-computer. The logic board and the processor board are connected to each other by a fixed DIL plug and socket. The array carrier board is connected to the main p.c. boards by means of a flexible ribbon cable. The physical positions of these boards can be seen in figure 12.1. Further details of the M PDA circuits can be found in appendix A8.

Figure 12.4 shows the electrical connections for interfacing the boards to the computer or an oscilloscope. To operate the sensor-head with the computer, the video output (BNC3) has to be connected to the A/D converter, the external oscillator (BNC1) to the buffered output

of the parallel I/O port of the BBC-computer and the switch has to be in the left position, thereby disconnecting the internal oscillator.

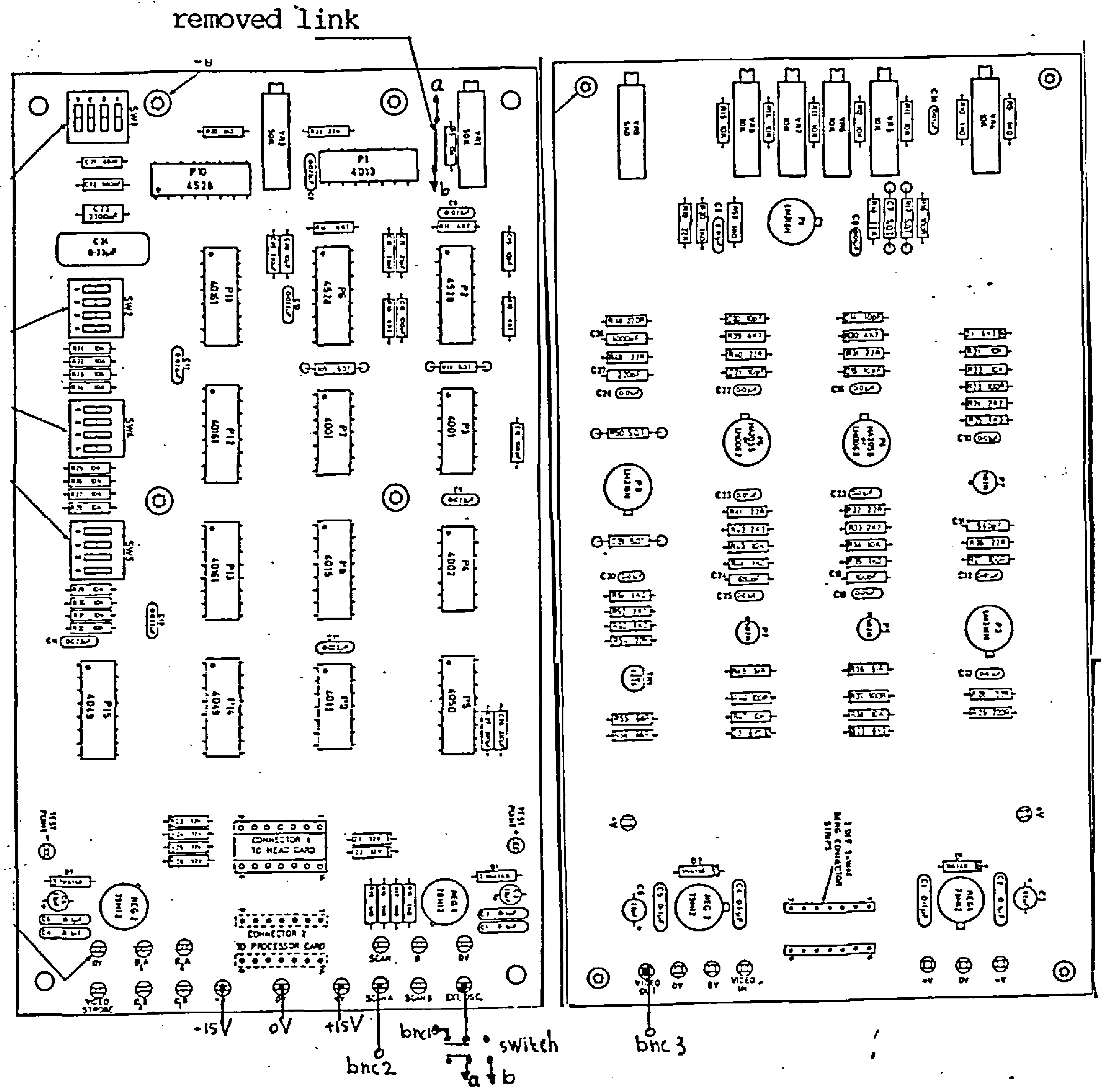


Figure 12.4. Electrical connections of the diode array electronics.

For operation with an oscilloscope, the switch has to be in the right position and the video output (BNC3) has to be connected to an input channel on the oscillscope and the trigger signal (BNC2) to the external triggering input of the oscilloscope.

12.1.3 The optics.

To enable the acquisition of optical spectra of turned surfaces, two sensor-heads have been constructed with different optics. The differences are caused by the light sources used, respectively a He-Ne laser tube and a collimated laser diode. In the head with the laser tube a beam expansion system is necessary to expand the original beam of .63mm to about 5mm. The two basic lay-outs are shown in figure 12.2. The system with the laser tube contains two extra mirrors in order to achieve a more compact system by folding the optical path. The two mirrors are located between the laser tube and the microscope objective MO.

The system with the He-Ne laser.

Here (figure 12.2 top), the combination of the laser LT and the microscope objective MO (focal length 24.8mm) produces a pointsource in the focal planes of the two cylindrical lenses L_1 and L_2 whose optical axes are perpendicular. The lens L_1 (focal length 22mm) collimates the light in the vertical plane and the lens L_2 (focal length 200mm) in the horizontal plane. This produces a long narrow collimated beam with an elliptical cross-section. The major axis equals 5.08mm and the minor axis 0.56mm. The major axis is parallel to the axis of the cylindrical, turned component TC. The light, reflected by TC, is collected by the lens L_1. This lens acts now as an one-dimensional Fourier transform lens. The reflected light

is directed towards the photo-diode array DA via the beamsplitter cube BS. The photo-diode array is located in the focal plane of L_1.

The system with the laser diode.

Since the laser diode emits a collimated beam of suitable diameter (5.4mm), no further beam expansion optics are necessary. The lay-out for this system is shown in figure 12.2 bottom. Figure 12.1 shows the mechanical lay-out of this system. The laser output is directed via a beamsplitter plate BS towards the turned surface TC. The reflected light passes through the beamsplitter plate and the two cylindrical lenses L_1 and L_2 towards the photo-diode array DA. Here the optical axes of the two lenses are parallel to one another and perpendicular to the axis of the cylindrical, turned component and the length axis of the photo-diode array. The lens L_1 is plano-convex and L_2 is plano-concave. This arrangement enables to obtain a long effective focal length with a reasonable short distance between the lens L_1 and the back focal plane of the lens combination, where the photo-diode array is situated. This contributes to the compactness of the sensor. In order to vary the focal length of the system, the lens L_1 can be positioned anywhere at a distance between 90-170mm, and the lens L_2 anywhere between 80-160mm away from the photo-diode array, depending upon the position of L_1. Figure 12.5 shows the range of focal lengths which can be achieved by this system. The curves are calculated by the following formulae for compound lens systems:

The effective focal length efl is

$$efl=f_1f_2/(f_1+f_2-d) \qquad (12.1)$$

and the back focal length bfl is

$$bfl=(f_2d-f_2f_1)/(d-f_1-f_2) \qquad (12.2)$$

which results in a total length bfld between the first lens L_1 and the the back focal plane

$$bfld=bfl+d=efl+d+f_2d/(d-f_1-f_2) \qquad (12.3)$$

where f_1 is the focal length of L_1, f_2 the focal length of L_2 and d the distance between L_1 and L_2. The maximum value of bfld is 170mm.

The curves are calculated for focal lengths of respectively f_2=-60mm and f_1=60,80 and 100mm. These values are choosen because they are normally available from various manufacturers. As can be deduced from the graphs, an effective focal length rangeing from 150 to 295mm can be achieved by combining the above mentioned lenses. Obviously, single lenses can be used to achieve a focal length between 70 to 170mm. The minimum distance of 70mm is due to a power supply unit which is situated between the photo-diode array and L_2. This unit is not shown in the drawings. So the focal length can be varied between 70 to 295mm. This facility is useful for matching the surface wavelength to the length of the photo-diode array in order to obtain an optimum spatial resolution.

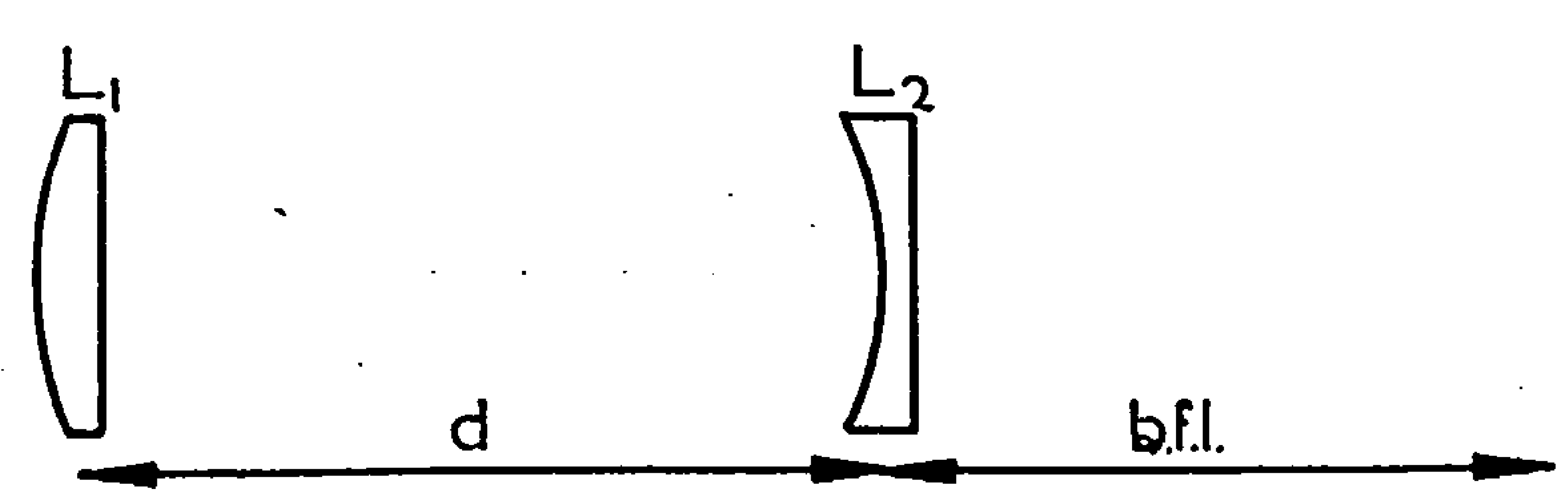


Figure 12.5. Range of focal lengths obtainable with various combinations of lenses indicated.

12.1.4 Difference between the theoretical and practical system.

In chapter 9 the following assumptions have been made for the derivation of the Kirchhoff difraction integral upon which the operation of the sensor-head is based:

- i The illuminated area is infinitely long and uniformly illuminated.
- ii All the light is reflected without secondary reflection or shadowing effects caused by the surface slopes.
- iii No polarization effects are present.
- iv The transform lens has an infinitely wide aperture in order to process all the surface data uniformly, in other words, the system is space-invariant.

In practice these conditions cannot be fully met. The influence of the first deviation of the practical system from the theoretical model has been discussed in section 10.3.

For the estimation of the importance of the various deviations, it is assumed that the surface is subjected to normal incidence illumination and that a double-sided optical spectrum is obtained.

Shadowing effect.

The shadowing effect occurs when the reflected light, in the

non-specular direction, is obstructed by a surface asperity. An estimate of the importance of this can be obtained by assuming a square profile (see figure 12.6). This is the most severe condition.

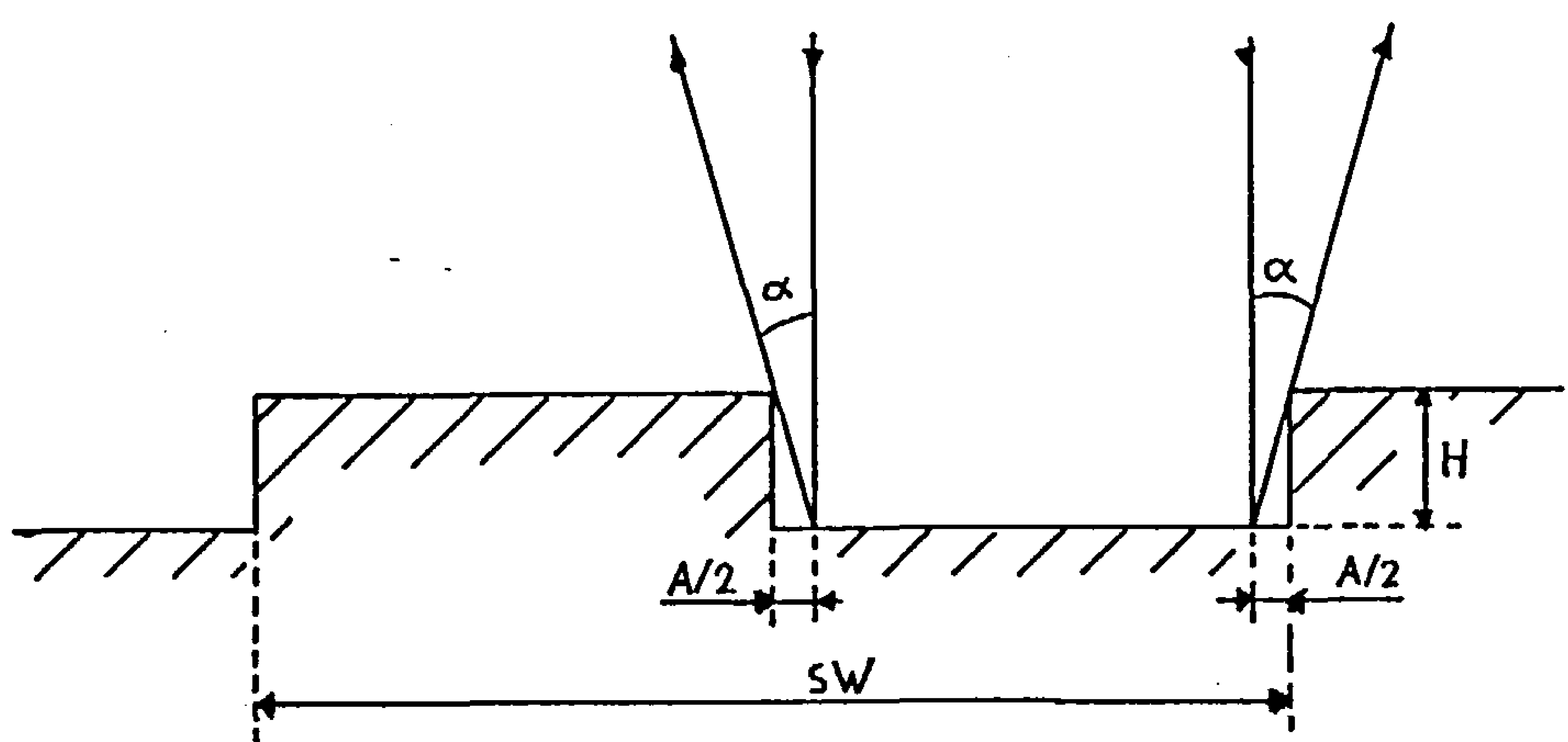


Figure 12.6. Influence of the shadowing effect.

The step height H is 175nm, which results in R_q/λ =0.1 when λ is 870 nm. The maximum diffraction angles under consideration are 1.23° and 5.18° for focal lengths of the transform system of respectively 295mm and 70 mm. The angle is calculated by α=arctan(L/f), where L is the half length of the photo-diode array (6.35mm) and f the focal length of the optical system. This results in a length A of the profile SW which is obstructed:

$$A=H/\tan(90-\alpha)= 3.7nm \quad (f=295mm)$$

$$A=H/\tan(90-\alpha)=15.8nm \quad (f= 70mm)$$

If 1% of the surface profile is allowed to be obstructed, then the minimum spatial wavelength of the square surface profile equals respectively 0.74um and 3.16um. This condition is generally met in

precision turning where a minimum spatial wavelength of 20um is common. Furthermore, sudden step heights are very rare and the surface has normally very shallow slopes. From this it can be concluded that shadowing effects can be ignored for normal incident illumination. However, for non-normal incidence illumination this effect can become important.

Polarisation.

The maximum diffraction angles under consideration vary between 1.23° and 5.18°. Referring to the graphs of figure 7.3 on page 133, it is obvious that the polarisation effects are nil for diffraction angles smaller than 10°. So for the system under discussion the polarisation effects can be safely ignored.

Effect of the finite size of the aperture.

Another constraint which has to be satisfied by the optical system is that it is space-invariant. An optical system is space-invariant when it acts uniformly on every part of the input data. This depends on:

i the aperture of the optical system, 2A=60mm

ii the distance D between the lens and the object.

iii the focal length f of the system.

iv the length of the photo-diode array, 2L=12.7mm

v the length of the illuminated surface profile, 2b=5mm.

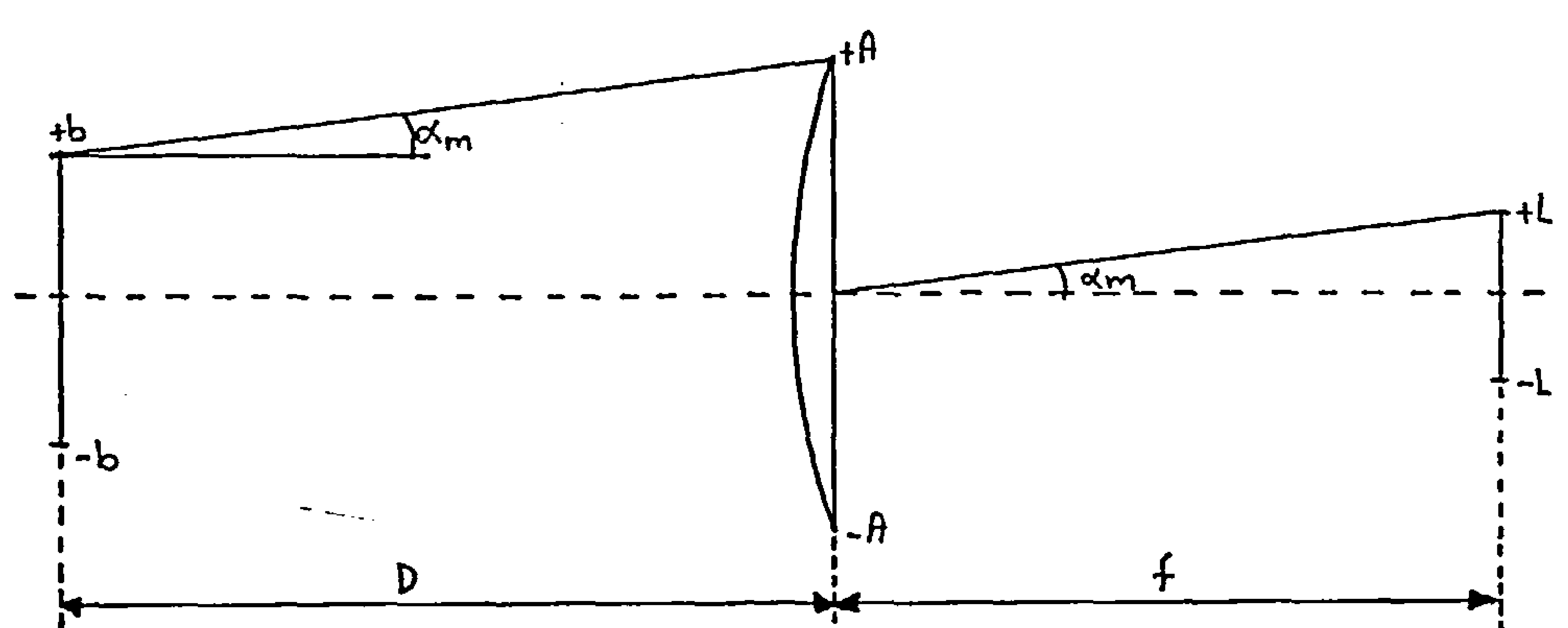


Figure 12.7. Conditions for a space-invariant transform system.

Figure 12.7 shows the various parameters involved. From this geometry it follows that the maximum angle α_m is given by:

$$\alpha_m = \arctan(L/f) \tag{12.4}$$

In order for the system to be space-invariant, all the rays leaving the surface aperture for the transform plane, must reach the transform lens aperture. This can be expresed by:

$$b + D \cdot \tan(\alpha_m) = A$$

or

$$D = (A-b) \cdot f/L \tag{12.5}$$

where D denotes the maximum distance between the lens and the surface. Depending on the focal length of the system, the maximum stand-off distance D equals 1277mm or 303mm for focal length of respectively 295mm and 70mm.

Maximum and minimum detectable surface wavelength.

The relationship between the surface wavelength SW and its spectral position d on the photo diode array, is given by:

$$SW=f\,\lambda/d \tag{12.6}$$

From this follows that the minimum detectable surface wavelength SW_{min} is related to the maximum value possible for d, i.e. half the length of the photo diode array. On the other hand, the maximum detectable surface wavelength SW_{max} is related to the resolution limit, set by the blurring effect of the optical system. This limit (see page 253) equals 4uinches=0.102mm. This leads to the following results:

f(mm)	SW_{max}(um)	SW_{min}(um)
70	597	9.4
295	2516	39.5

SW_{max}, however, is also limited by the illuminated aperture. From computer simulation results this limit seems to be one fifth of the illuminated aperture, in this case SW_{max} 1000um (see figure 12.8).

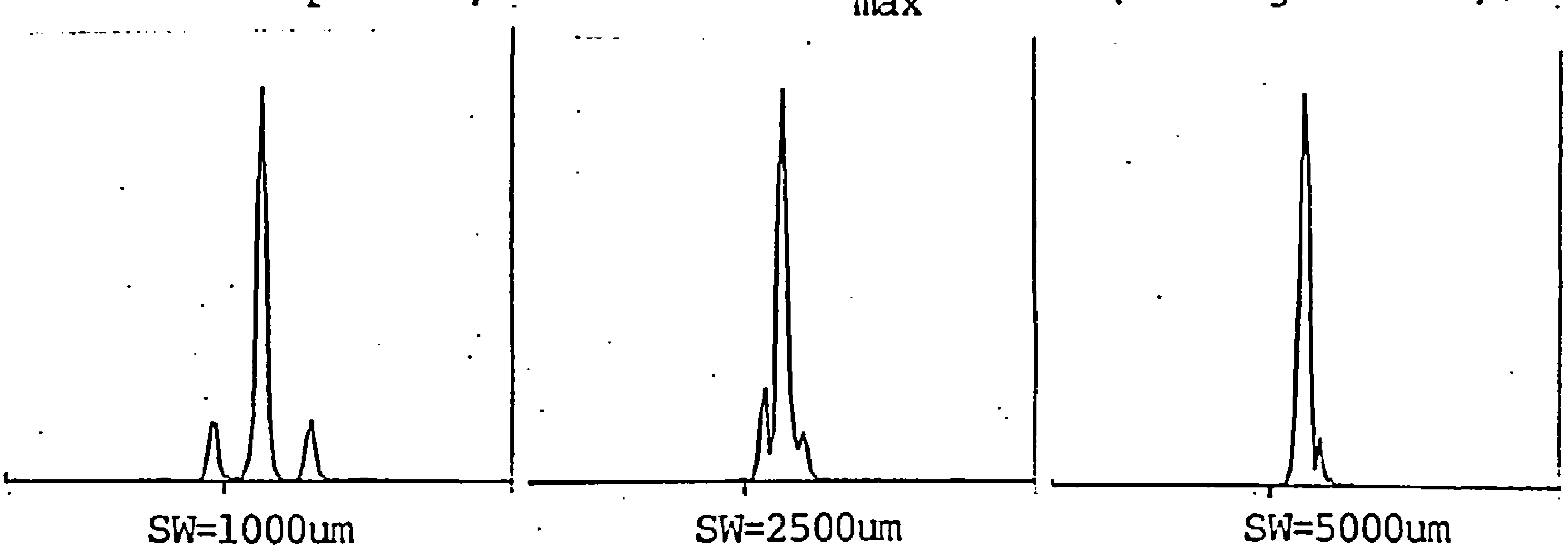


Figure 12.8. Determination of SW_{max}, caused by the illumination.

12.2. The computer.

A BBC-computer is used for the control of the sensor-head and the analysis of the optical diffraction pattern.

12.2.1. Hardware.

A computer program, stored in an EPROM, produces clock pulses on pin PB7 of the 8 pins of the parallel input-output (I/O) port of the computer for the control of the scanning rate of the photo diode array in the head. The frequency of these clock pulses is software controlled, and depends on the amount of light received by the diode array. It acts as an automatic exposure control. The output voltage of the I/O port is 5V, which is not sufficient to drive the electronics of the photo-diode array, which require a voltage in the range of 12V. Therefore the computer signal is buffered by using two stages of a Hex inverter 7406 (figure 12.9a) whose output is amplified by a single stage transistor amplifier (BFY50), see figure 12.11, to 12V. The computer output signal is also used for provision of the start-conversion pulse for the 8-bit A/D converter ZN427 (figure 12.9b). The computer sigal is first fed through one stage of an inverting Hex Schmitt trigger 7414 (figure 12.9c). The Schmitt trigger cleans up the signal and an inverting output is choosen since it produces an effective phase shift between the start-conversion pulses and the clock-pulses to the sensor. This ensures that the data

from the sensor is ready for conversion by the time that the A/D converter starts the conversion, thereby simplifying the software by eliminating the need of a handshaking protocol.

The analogue output of the sensor-head is fed into the purpose built A/D converter. The digital output of this converter is connected to the pins PB0-6 of the parallel I/O port. The A/D converter provided by the BBC computer is not used for two reasons, firstly it is rather slow (4msec for an 8 bit conversion) and secondly the conversion is inaccurate because of an unstable reference voltage supplied to this converter. Due to the lack of the availabilty of 8 digital inputs to the I/O port (one pin is used for output), only a 7-bit conversion is performed.

The ZN427 is an 8-bit bipolar successive approximation A to D converter i.c. This device incorporates tri-state buffers to permit bussing on common data lines, a voltage switching D to A converter, a 2.5V precision reference, a fast comparator and succesive approximation logic. Operation is from a +5V and -3 to -30V supplies and data outputs are TTL compatable. A 12us conversion time can be obtained with a 750KHz clock signal. Figure 12.9c shows the pin connections of the ZN427. A maximum analogue input voltage of 10V is obtained by the connection of the simple resistor network (figure 12.10) between the pins 6,7 and 9 of the i.c. The clock pulses are generated by i.c. 7432 (figure 12.9d), which can provide three output frequencies namely f, f/2 and f/4, where f is the frequency of the crystal (3MHz) connected across the pins 14 and 15 of this i.c. The f/4 output at pin 6 is chosen, thereby providing the A/D converter with a 750KHz clock which falls within its range from 600-900KHz.

The complete circuit diagram for interfacing the BBC computer and the sensor-head is shown in figure 12.11. The +5,12 and-5V supplies are provided by the internal BBC computer power supply.

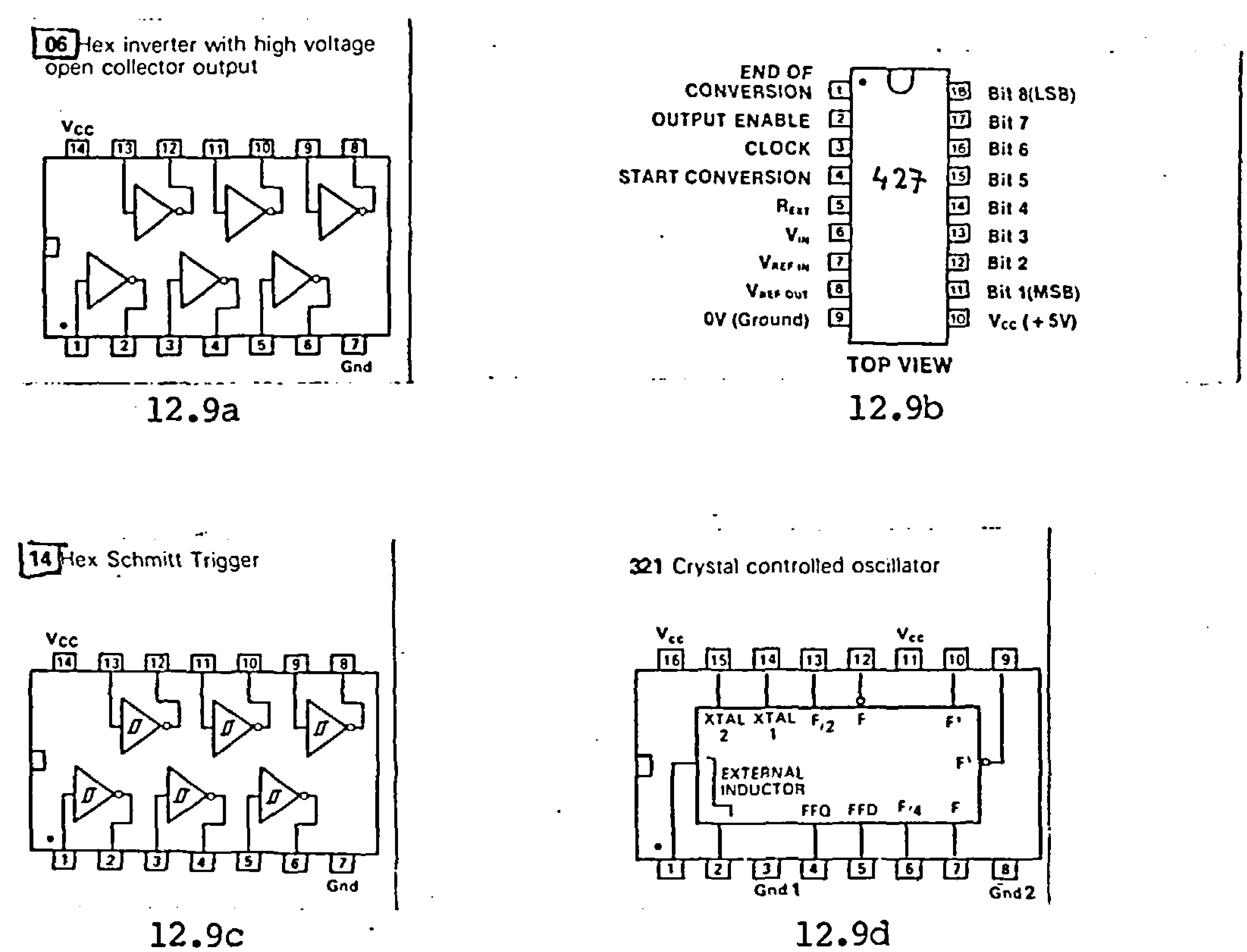


Figure 12.9. Pin connections of the i.c.'s used in the interface.

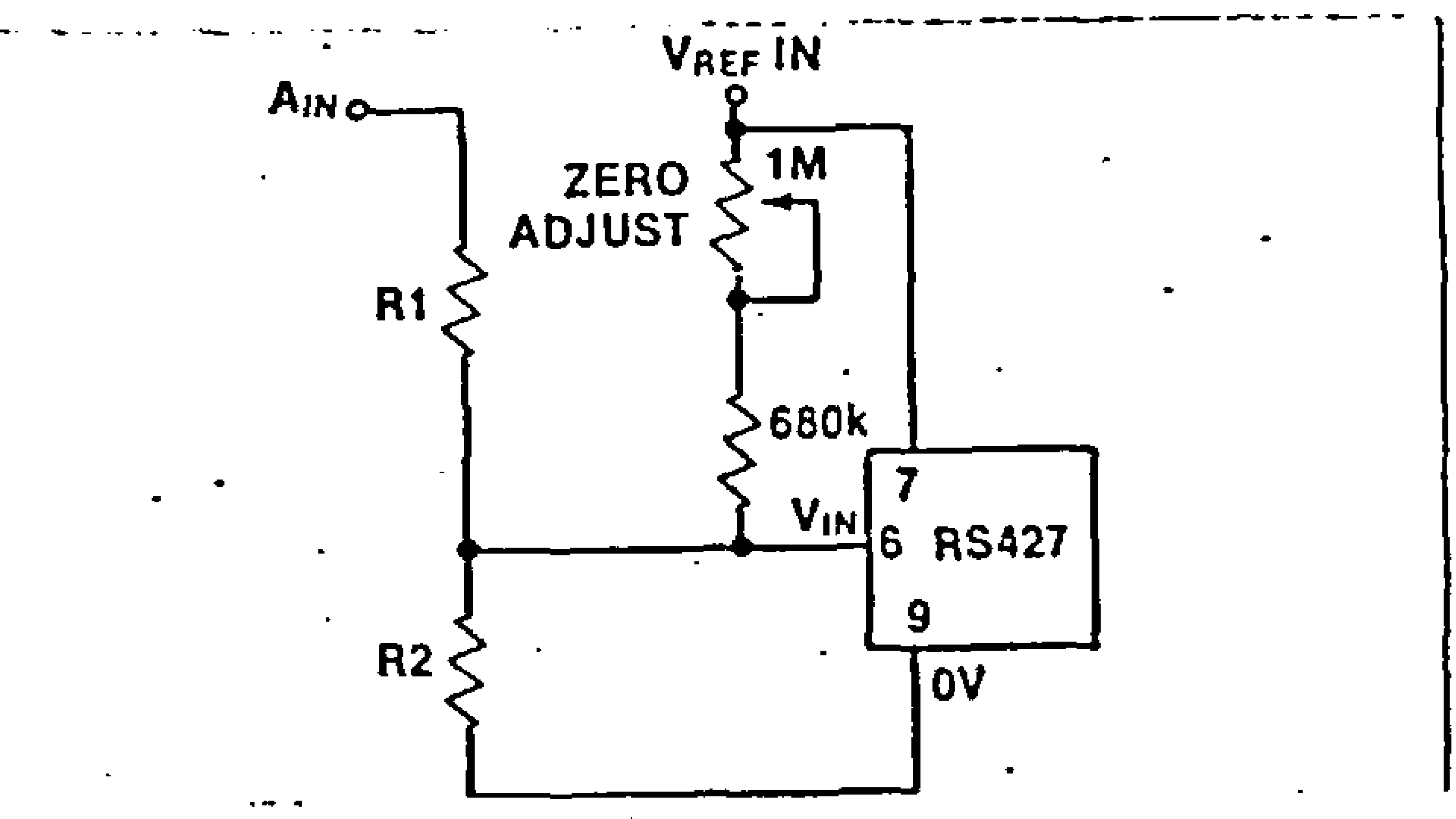


Figure 12.10. General resistor network for unipolar operation of the A/D converter i.c. ZN427.

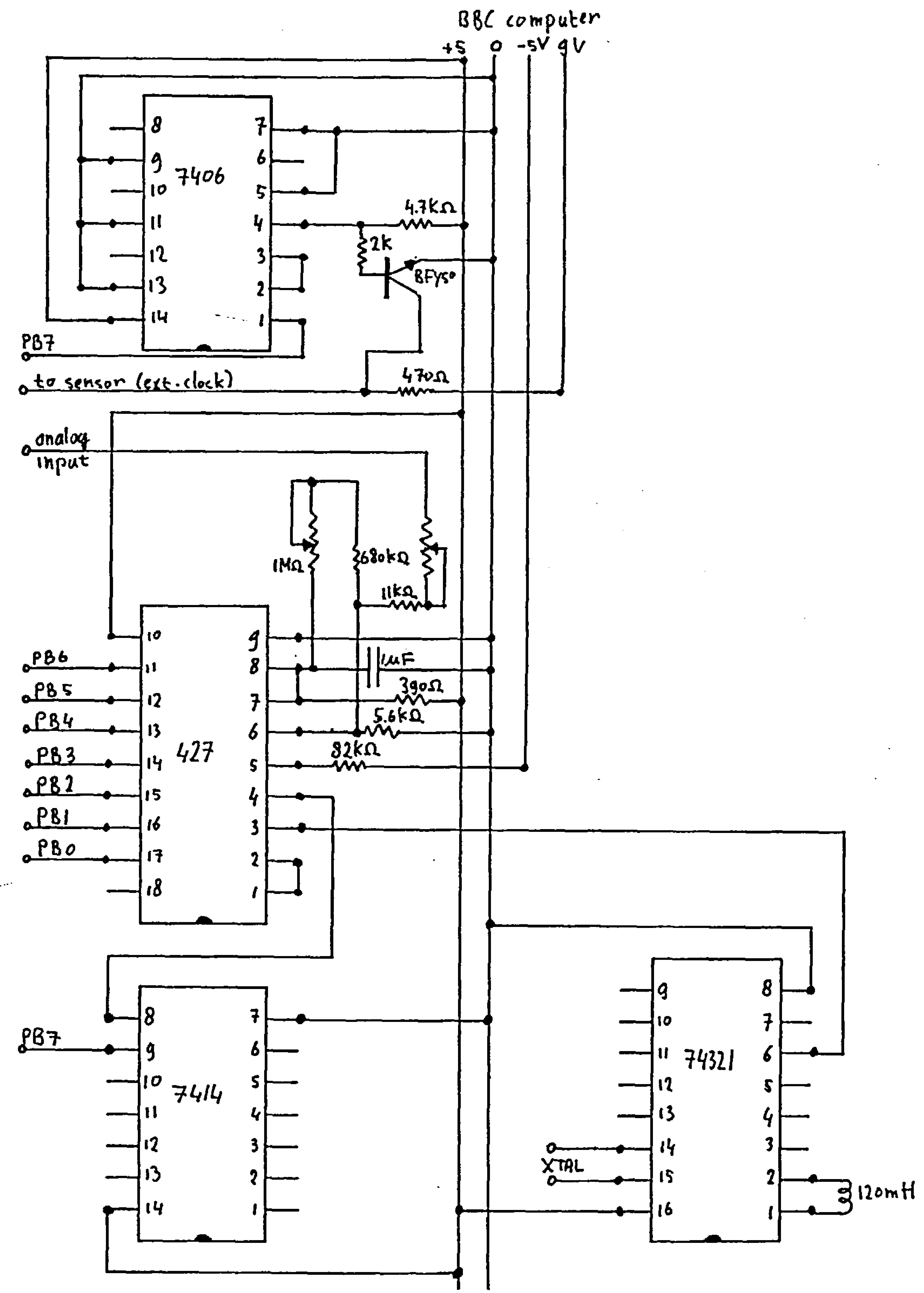


Figure 12.11. Electronic circuit diagram for interfacing the BBC computer and the optical sensor-head.

12.2.2. Software.

The computer program driving the sensor-head and analysing the optical data, is permanently stored in an EPROM. The program is written in a mixture of BBC Basic and 6502 assembly language. The assembly language performs the data logging and subsequent procesing of the raw optical data. The display of various messages, final calculations and display of the results is taken care of by Basic.

Figure 12.12 shows the flowchart of the computer program. The names preceeded by a dot, printed next to the boxes, are subroutines written in assembly language. Appendix 9 contains the full listings and a brief description of these subroutines together with flowcharts and the listing of the Basic program.

The scanning rate of the diode array is automatically adjusted to ensure that the photo-diodes receive enough light. The scanning rate can be adjusted in the range of 800Hz up to 33KHz. If not enough light is received at the slowest scanning speed, then the computer will display a message whether the diode laser is switched on or the object is missing. It displays furthermore messages to aid the angular adjustment of the object in order to display the diffraction pattern as central as possible onto the photo diode array. The algorithms to convert the relative intensity amplitudes into R_q values are those published in 1986. [Rakels 1986c]. They are:

$$R_q/\lambda=(\sqrt{1-A(0)})/4 \quad ; \quad 0.00<R_q/\lambda<0.03 \qquad (12.7A)$$

$$R_q/\lambda=0.932(1-A(0))+0.167 \quad ; \quad 0.03<R_q/\lambda<0.08 \qquad (12.7B)$$

$$R_q/ =0.609A(1)[1-2A(1)/(1-A(0))]+0.0496; \quad 0.08<R_q/ <0.125 \quad (12.7C)$$

where A(n) are the relative intensity amplitudes of the diffraction pattern. The last two algoritms differ from the ones derived in this thesis (see page 168). The above algorithms A,B and C are obtained by a curve fitting process applied to the same type of curves as shown on the pages 163 and 164. Further details on this can be found in the quoted reference. Great care has to be taken when applying algorithm C since it is rather unstable due to a division by [1-A(0)]. In its validity range this term equals nearly zero, so a small error in the estimation of A(0) causes big errors in the final result. However the roughness range of these three algorithms is 25% larger then of the algorithms derived in this thesis, which on the other hand are very stable.

The total time necessary for the evaluation of the roughness value is 1 or 5 seconds, depending on whether or not a diffraction pattern display on the monitor is requested.

The computer displays also the average feed spacing of the turning marks. This value is obtained by conversion of the average spacing of the diffraction peaks with the aid of the grating formula 7.5 on page 126. This value should not be confused with the rms surface wavelength.

The software also provides the possibility of a print-out of the results.

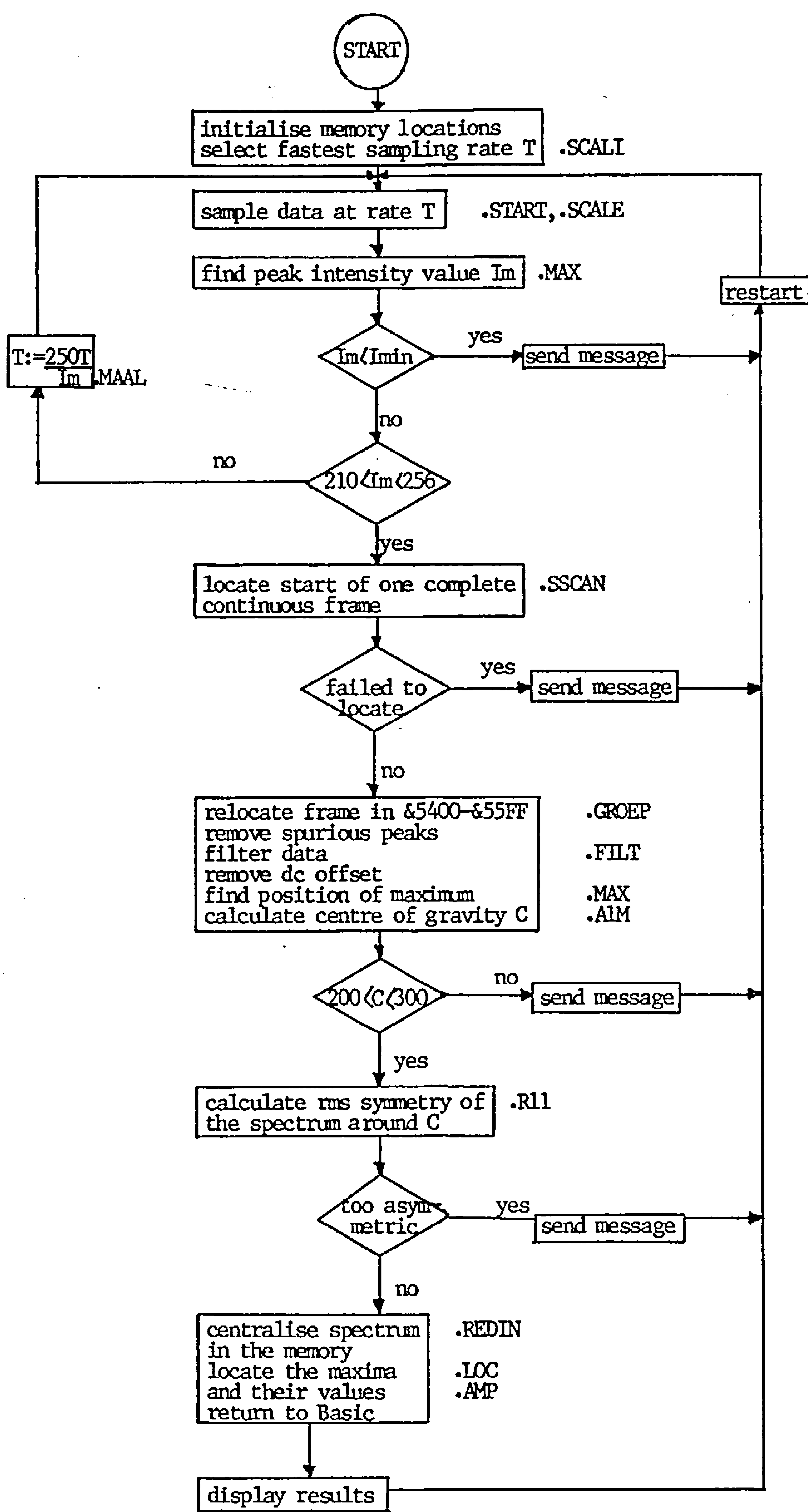


Figure 12.12. Flowchart of the computer program of the sensor.

12.3. Practical results.

Originally the algorithms 12.7A-C have been applied to estimate the R_q parameter of four different samples. Their optical diffraction patterns are depicted in figure 12.13 The spectra in figure 12.13a,b and c are published by E.G.Thwaite [1982]. The pattern in figure 12.13d is produced by a diffraction grating.

The relative amplitudes of these spectra are listed in table 12.1 alongside with the illumination wavelengths. Table 12.2 lists the results $R_{q\ opt.}$ by application of the relevant algorithms 12.7 and 9.42 on the data listed in table 12.1, it also states the R_q parameter obtained by conventional stylus instruments as reported in private communication with E.G.Thwaite. For these R_q values there is no cutoff specified, however, since there is no waviness present on his surface profile traces, these R_q values can be assumed to be accurate.

Table 12.1 The relative amplitudes of the spectra 12.13a-d.

		Amplitude					
Spectrum	(um)	A(0)	A(1)	A(-1)	A(2)	A(-2)	A(-3)
12.13a	3.39	20	5	4	--	--	--
12.13b	3.39	8	19.5	23.5	9	9.5	--
12.13c	3.39	9	25.5	25	20	16	5
12.13d	0.633	30	11.5	9.5	--	--	--

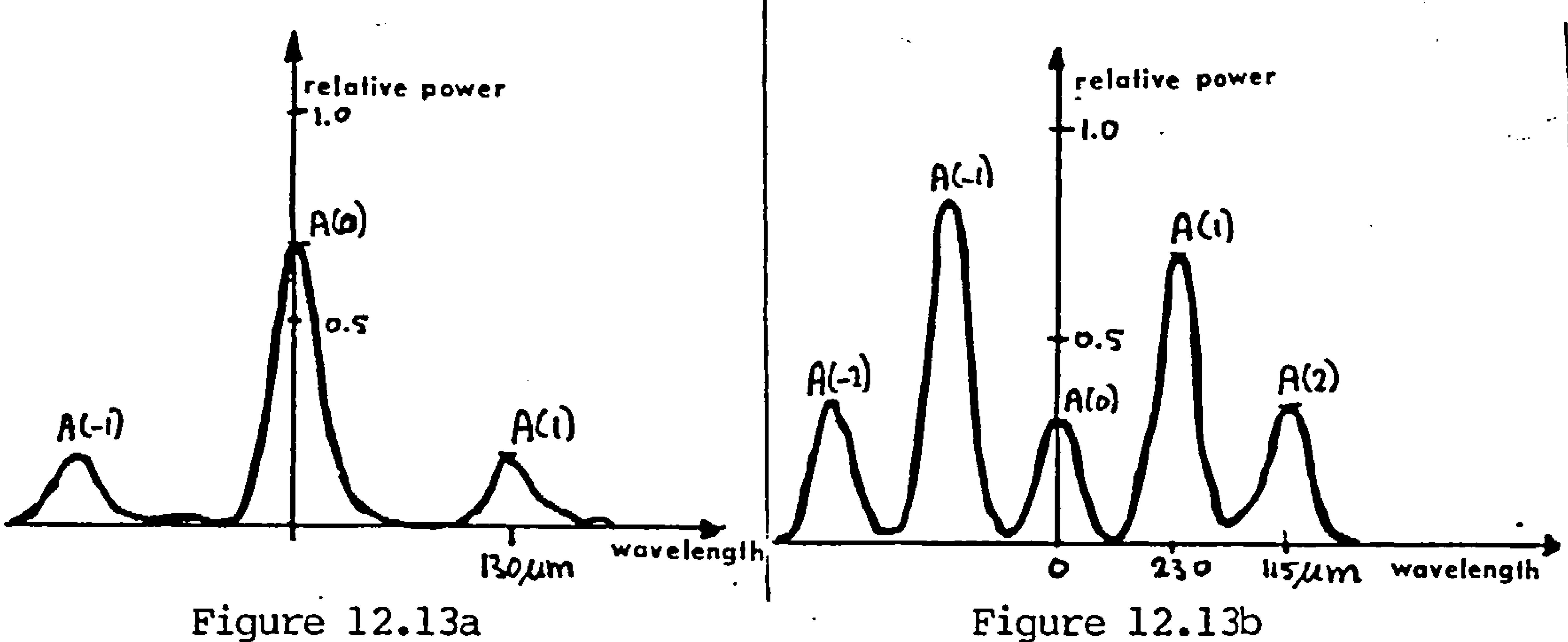


Figure 12.13a

Figure 12.13b

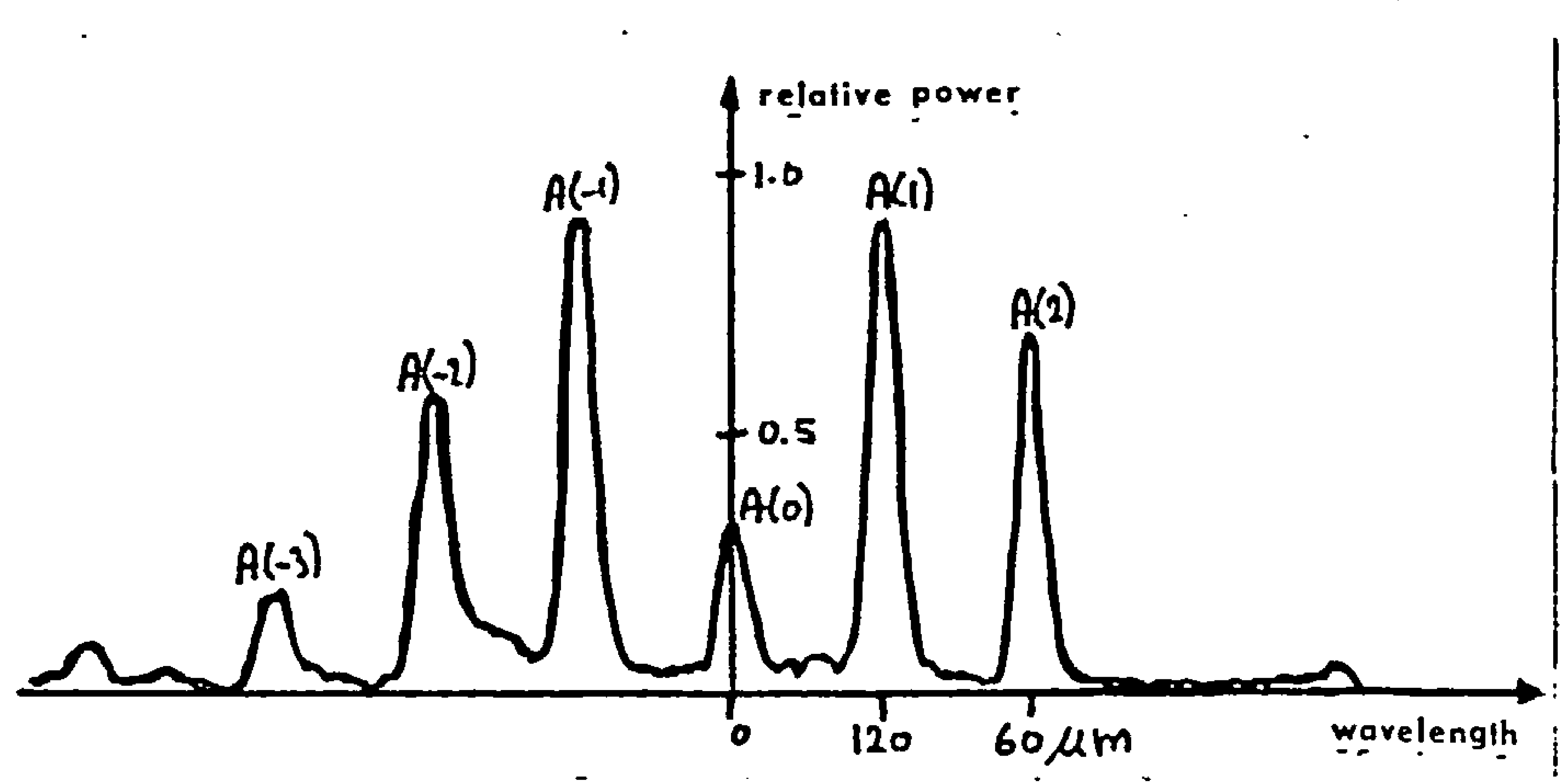


Figure 12.13c

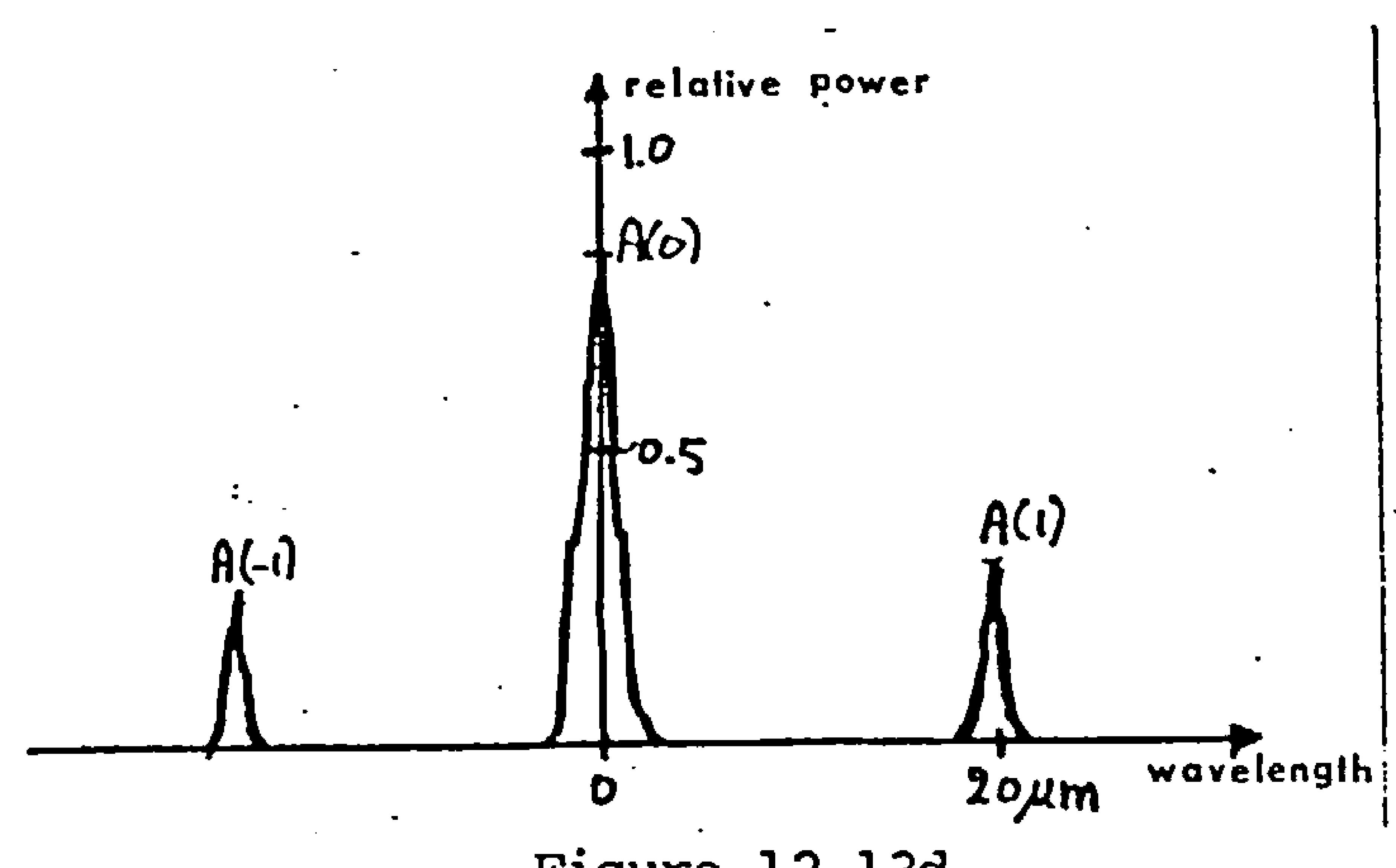


Figure 12.13d

Figure 12.13. Optical power spectra, which are used to calculate R_{qop}

(a) R_q=158nm, (b) R_q=370nm, (c) R_q=480nm, (d) R_q=31nm.

Table 12.2 Results obtained with the algorithms 9.42 and 12.7.

Spectrum	Algorithm	R_q/λ	R_{qopt}(nm)	Algorithm	R_q/λ	R_{qopt}(nm)	R_{qstyl}
12.13a	9.42b	0.047	159	12.7b	0.046	156	158
12.13b	n.a.	--	---	12.7c	0.112	380	370
12.13c	n.a.	--	---	12.7c	0.13	440	480
12.13c	9.42c	0.056	35	12.7b	0.056	35	31

The sample correlation coefficient r_{xy} of R_q and $R_{q\ opt}$ is 0.996. Although this number is based on only 4 samples, this result is important, since the critical value of r at a significance level of 1% for 4 observations and two variables is 0.990. If r_{xy} exceeds r, then the null hypothesis that there is no association between the two variables is rejected at the given level. In this case the conclusion is that correlation exists and the level of significance represents the probability that the the conclusion is wrong.

The sensor has also been evaluated independently at Rank Taylor Hobson Ltd (Leicester) [report JKT/YRC, 24.6.85]. The test surfaces used were R_a standard gratings with a 50% duty cycle and , therefore, R_q=R_a. Seven physically different gratings were used. The R_q values were firstly obtained with a stylus instrument according to (R/ISO/0.8mm). Their results are listed in table 12.3.

Table 12.3 Test results obtained by Rank Taylor Hobson Ltd.

grating	R_q (nm) stylus	R_q (nm) optical	Algorithm used
1	15	14	12.7a
2	21	21	12.7a
3	33	29	12.7b
4	35	37	12.7b
5	53	45	12.7b
6	53	45	12.7b
7	101	95-100	12.7c

From these results they came to the conclusion that the diffractometric instrument is capable of determining R_q values which fall into a special category, and not just relative roughness values.

12.4 Conclusions.

Within the surface roughness range $0<R_q/\lambda<0.10$, it is possible to estimate optically R_q values of periodic surfaces, by analysing the optical diffraction pattern. To convert the amplitudes of this pattern into a R_q value, three formulae (9.42a-b) have been derived for specific roughness ranges. On these formulae a computer algorithm can be based by adopting the following strategy:

i determine the amplitudes A(m)

ii calculate a(0) by $a(0)=A(0)/\sum_{m=-\infty}^{\infty}A(m)$.

iii apply formula 9.42a, if $0<R_q/\lambda calc<0.022$, then $R_q/\lambda=R_q/\lambda calc$, if not then go to (iv)

iv apply formula 9.42b, if $0.022<R_q/\lambda calc<0.056$, then $R_q/\lambda=R_q/\lambda calc$, if not then go to (v)

v apply formula 9.42c, if $0.056<R_q/\lambda calc<0.096$, then $R_q/\lambda=R_q/\lambda calc$, if not then the surface exceeds the range

The results of this routine are listed in table 12.2. They show a large correlation between the optically obtained values R_q and the conventionally obtained R_q values. This proves that it is feasible to obtain R_q values by optical diffraction techniques.

Final conclusions.

Using the Kirchhoff diffraction integral as a starting point, three algorithms have been derived for the estimation of the R_q surface roughness parameter for periodic surfaces. It has also been shown that the second moment of the diffraction pattern equals twice the mean square of the surface slopes.

Using a geometrical reflection approach it has been shown that the validity range of the Kirchhoff diffraction approach is valid when the maximum surface slope does not exceed 0.4549.

The number M of diffraction orders necessary to contain 95% of the reflected light equals $M=4\pi h/D$, where h is the amplitude and D the wavelength of the surface profile. The maximum diffraction angle related to this value of M is determined by the grating formula:

$$\sin\Theta_M=M\ /D$$

This angle Θ_M can be related to the back focal length, F, of a transform lens and the half length, h/2, of the diffraction plane by:

$$\tan\Theta_M=h/2F$$

These last two rules are very useful for the design of an instrument which has to capture diffraction patterns from periodic surfaces.

Depending on the illumination wavelength, the useful range of this technique is about up to R_q =100nm. This roughness range is

normally achieved in diamond turning. It might appear that this only represents a small fraction of the totality of manufactured surfaces, but this fraction is not insignificant. Examples of these surfaces are photocopier drums, hard disks for computers but also germanium lenses. It is also interesting to note that these surfaces cannot be measured by stylus techniques because of possible damage.

The surface wavelength to which this technique applies is between the limits of 10um to 1000um depending on the optics used.

References.

ABBOTT, E.J. and FIRESTONE, F.A. 1933, "Specifying surface quality." Mech. Eng. 55, pp. 569.

ABBOTT, E.J. and GOLDSCHMIDT, E. 1937, "Surface quality." Mech. Engng. 59, pp. 813.

ABBOTT,E.J., BOUSKY,S. and WILLIAMSON, D.E. 1938, "The profilometer." Mech. Eng. 60, pp. 205.

ABRAMOWITZ, M and STEGUN, I.A. 1972 "Handbook of Mathematical Functions." Dover Publications,Inc. New York.

AGULLOR, J.B. and PAGES-FITA, J. 1974, "Performance analysis of the stylus techniques of surface roughness assessment: a random field approach." Proc. of the 15th Int. Machine Tool Des. & Res. Conf., Birmingham, U.K., pp. 349.

AHLERS, R.J. 1985, "White light method: A new sensor for the optical evaluation of rough surfaces." Opt. Eng. 24, pp. 423.

AJIOKA, S. 1966 , "Dynamic response of stylus." Bull. Jap. Soc. Prec. Eng. 1, pp. 228.

ALLEBACH,J.P. 1981, "Analysis of sampling-pattern dependence in time sequential sampling of spatiotemporal signals." JOSA 71, pp 99.

ARECCHI, F.T.,BERTANI, D. and CILIBERTO, S. 1979, "A fast versatile optical profilometer." Opt. Comm. 31, pp. 263.

ASAKURA, T. 1978, "Surface roughness measurement." in Speckle Metrology. ed. R.K. Erf (Academic Press, New York).

AXLINE, R.M. and FUNG, A.K. 1978, "Numerical computation of scattering from a perfect conducting random surface." IEEE Trans. on Antennas and Propagation, AP-26, pp. 482.

BAIRD, K.M. and HANES, G.R. 1967, "Interferometers" in Applied Optics and Optical Engineering, Vol IV, Chapt.9 ed Kingslake, R. Academic Press. New York. (1967).

BECKMANN,P. and SPIZZICHINO,A. 1963, " The scattering of electromagnetic waves from rough surfaces." Book. Pergamon Press, London.

BENNETT, H.E. 1961, "Specular reflectance of aluminized ground glass and the height distribution of surface irregularities." JOSA 53, pp. 1389.

BENNETT, H.E. and BENNETT, J.M. 1967, "Physics of thin films." Book eds.: Hass, G and Thurn, R.E., Vol 4 pp. 1-96. Academic Press, New York.

BENNETT, H.E. and PORTEUS, J.O. 1961, "Relation between roughness and specular reflectance at normal incidence." JOSA, 51,pp. 123.

BENNETT, J.M. 1976, "Measurement of the rms roughness, autocovariance function and other statistical properties of optical surfaces using a FECO scanning interferometer." Appl. Opt. 15, pp. 2705.

BENNETT,H.E. 1978, "Scattering characteristics of optical materials." Opt. Eng. 17, pp 480.

BENNETT,J.M. and DANCY,J.H. 1981, "Stylus profiling instrument for measuring statistical properties of smooth optical surfaces." Appl. Opt. 20, pp. 1785.

BER, A. and BRAUN, S. 1968, "Spectral analysis of surface finish." Ann. C.I.R.P., 16, pp. 53.

BERG,P.M. v.d. and FOKKEMA,J.T. 1979, "The Rayleigh hypothesis in the theory of reflection by a grating." JOSA 69, pp. 27.

BHUSHAN, B., WYANT, J.C. and KOLIOPOULUS, C.L. 1985, "Measurement of surface topography of magnetic tapes by Mirau interferometry." Appl. Opt. 24, pp. 1489.

BIDWELL, J.B. 1953, "Movable-anode tube gages surface roughness." Electronics, 26, pp. 181.

BIKERMAN,J.J. 1970, "Physical surfaces." Book. Academic Press, New York.

BLADEL,J van. 1964, "Electromagnetic fields." Book. MacGrawhill, New York.

BRODMANN,R. 1983a, "Optisches Rauheitsmessgeraet fuer die Fertigung." Feinwerktechnik & Messtechnik 91, pp63.

BRODMANN,R. 1983b, "Beruhrungslose automatische Messung der Oberflaechenrauheit von Tonkoepfen." Feinwerktechnik & Messtechnik 91, pp369.

BRODMANN,R., GAST,Th. and THURN,G. 1984, "An optical instrument for measuring the surface roughness in production control." CIRP Annals 33,pp. 403.

BRODMANN,R.,GERSTDORFER,O. and THURN,G. 1985, "Optical roughness measuring instrument for fine machined surfaces." Opt. Eng. 24, pp. 408.

BRODMANN,R.,PAISDZIOR,H.,RAU,N. and HUEBNER,G.1986, "Optische Messung der Rauheit und Welligkeit feinbearbeiteter Oberflaechen." Proc. Int. Conf. on Modern Production and Production Metrology. TU Vienna, paper113.3

BRODMANN, R. and THURN, G. 1986, "Roughness measurement of ground, turned and shot-peened surfaces by the light scattering method." Wear 109, pp. 1.

CHANDLEY,P.J. and WELFORD,W.T. 1975, "A reformulation of some results of P. Beckmann for scattering from rough surfaces." Opt. Quant. Elec. 7, pp. 393.

CHANDLEY, P.J.. 1976a, "Surface roughness measurements for coherent light scattering." Opt. Quant. Elec. 8, pp. 323.

CHANDLEY, P.J. 1976b, "Determination of the autocorrelation function of height on a rough surface from coherent light scattering." Opt. Quant. Elec. 8, pp. 329.

CHETWYND,D.G.,McKEE,F.A. and RAKELS,J.H. 1982, "Machined surfaces: Final texture and underlying structure." Wear,83.

CHIESORIN,P. and LONARDO,P.M. 1977, " A new sensor of surface roughnes for process control system." Soc. Manuf. Enging. (USA), MS77, paper 217.

CHINICK, H.P. 1968, "LVDT puts precision in surface texture measurement." Cutting Tool Engng. 20, pp. 13.

CHURCH, E.L. and ZAVADA, J.H. 1975, "Residual surface roughness of diamond turned optics." App. Opt. 14, pp. 1788.

CHURCH,E.L., JENKINSON,H.A. and ZAVADA,J.M. 1977, "Measurement of the finish of diamond turned surfaces by differential light scattering." Opt. Eng. 16, pp. 360.

CHURCH,E.L. 1979, "The measurement of surface texture and topography by differential light scattering." Wear 57, pp. 93.

DAMIR, M.N.H. 1973, "Error in measurement due to stylus kinematics." Wear, 26, 2, pp. 219.

DAINTY, J.C. 1975, "Laser speckle and related phenomena." Springer Verlag Berlin.

DAINTY, J.C. 1976, "The statistics of speckle patterns." in Progress in Optics vol. 14 ed. E. Wolf (North-Holland, Amsterdam).

DAVIES, H. 1954, "The reflection of electromagnetic waves from a rough surface." Proc. Inst. Elec. Engrs. 101, pp. 209.

DEPEW, C.A. and WEIR, R.D. 1971, "Surface roughness determination by the measurement of reflectance." Appl. Opt. 10, pp. 969.

DETRIO,J.A. and MINER,S.M. 1985, "Standardised total integrated scatter measurements of optical surfaces." Opt. Eng. 24, pp. 419

DIETZ, R.W. and BENNETT, J.M. 1967, "Smoothness and thermal stability of Cer-Vit optical material." Appl. Opt. 6, pp. 1275.

DEUTSCH,S. 1965, "Psuedo-random dot scan television systems." IEEE Trans. Broadcasting, BC-11, pp. 11.

DUPUY,O. 1967/68, "High-precision optical profilometer for the study of micro-geometrical surface defects." Proc. Instn. Mech. Engrs.182 Pt3K, pp. 255.

DUTSCHKE,W. and EISSLER,W. 1978,"A new sensor for measuring the surface roughness in-process on a grinding machine." Proc. 3rd Int. Conf. on Automated Inspection and Product Control (Nottingham), pp. 19.

EASTMAN,J.M. 1980,"The scanning Fizeau interferometer: an automated instrument for characterising optical surfaces." Opt. Eng. 19, pp. 810.

ELSON, J.M. and BENNETT J.M. 1979, "Relation between the angular dependence of scattering and statistical properties of optical surfaces." JOSA 69, pp. 1.

ERF, R.K. 1978, "Speckle metrology." Academic Press, New York.

ERTL, F. 1978, "Aufbau und Untersuchung eines beruehrungslos optisch arbeitenden Laengenmessverfahrens fuer den Einsatz in der Fertigung." Dissertation TH Darmstadt.

ESTILL, W.B. and MOODY, J.C. 1966, "Deformation caused by stylus tracking on thin gold film." I.S.A. Trans., 5, pp. 373.

FAINMAN, Y., LENZ, E. and SHAMIR, J. 1982, "Optical profilometer: a new method for high sensitivity and wide dynamic range." Appl. Opt. 21, pp. 3200.

FENSTERMAKER,C.A. and McCRACKIN,F.L. 1969. " Errors arising from surface roughness in ellipsometric measurement of the refractive index of a surface." Surface Science 16, pp. 85.

FORTUIN, L. 1970, "Survey of literature on reflection and scattering of sound waves at the sea surface." JOSA 47, pp. 1209.

FUJII, H. and LIT, J.W.Y. 1978, "Surface roughness measurement using dichromatic speckle pattern: An experimental study." Appl. Opt. 17, pp. 2690.

GAST,Th. and THURN,G. 1984, "Automatisiertes Messsystem zur Oberflaechenpruefung nach der Streulichtmethode." Proc. VI Oberflaechen Kolloquium, TH Karl-Marx-Stadt, paper 38

GIGLIO, M., MUSAZZI, S. and PERINI, U. 1979, "Surface roughness by means of speckle wavelength decorrelation." Opt. Comm. 28, pp. 166.

GOLD, R. 1964, " An iterative unfolding method for responce matrices" Report ANL-6984. Argonne National Laboratory, Argonne, U.S.A.

GOODMAN, J.W. 1976, "Some fundamental properties of speckle." JOSA 66 pp. 1145.

GRANEEK,M and WUNSCH,H.L. 1952,"Application of pnuematic gauging to the measurement of surface finish." Machinery 81, pp. 701.

GRAY, G.G. and JOHNSON, K.L. 1972, "The dynamic response of elastic bodies in rolling contact to random roughness of their surfaces." J. Sound & Vibration 22, pp. 323.

GREENWOOD,J.A. and TRIPP,J.H. 1967, "The elastic contact of rough spheres." Trans. ASME: J. Appl. Mech. 34E, pp. 153.

GUERRERO, J.L. and BLACK, J.T. 1972, "Stylus tracer resolution and surface damage as determined by scanning electronic electron microscopy." Trans. ASME Ser. B. J. Eng. Ind., 94, pp. 1087.

HAESING,J. 1961, "Determining surface finish of workpieces by means of surface standards." Microtecnic 15, pp. 24.

HAMMOND, D.L. 1970, "Surface finish analysis through analog techniques." S.A.E. Paper No. 700142.

HARRISON, R.E.W. 1931, "A survey of surface quality standards and tolerance costs based on 1929-1930 Precision-Grinding practice." Trans. ASME 53, pp. 11.

HECHT,E. and ZAJAC, A. 1974, "Optics" Book, Addison-Wesley Publ. Co.

HENRY, J.J. and FENECH. H. 1964, "The use of analog computers for determining surface parameters required for prediction of thermal contact conductance." Trans. ASME Ser. C. J. Heat Transfer 86, pp. 543.

HENSLER,D.H. 1972, "Light scattering from fused polycristalline aluminium oxide surfaces." Appl. Opt. 11, pp. 2522.

HLLMANN,W., KRANZ,O. and ECKOLT,K. 1984, "Reliability of roughness measurements using contact stylus instruments with particular reference to results of recent research at the Physikalisch Technische Bundesanstalt." Wear 97, pp. 27.

HINGLE, H.T. and RAKELS, J.H. 1983, " The practical application of diffraction techniques to assess surface finish of diamond turned parts." CIRP Annals 32/1, pp. 499.

HINGLE,H.T. and RAKELS,J.H. 1984, "In-Prozess-Messung an diamantgedrehten Oberflaechen durch optische Beugungsmethoden." VI Oberflaechenkolloquium, Karl-Marx-Stadt, paper 45.

HIRST,W. and HOLLANDER,A.E. 1974, "Surface finish and damage in sliding." Proc. Roy. Soc. Lond., A337, pp. 79.

HOWES, V.R. 1974, "An angle profile technique for surface studies." Metallography., 7, pp. 431.

IKUNO,H and YASUURA,K 1973, "Improved point-matching method with application to scattering from a periodic surface." IEEE Trans. on Antennas and Propagation, AP21, pp. 657.

JANSSON,P.A. 1984, "Deconvolution with application to spectroscopy." Book. Academic Press, Orlando, U.S.A.

JIRACEK,G.R. 1973, "Numerical comparisons of a modified Rayleigh approach with other rough surface EM scattering solutions." IEEE Trans. on Antennas and Propagation, AP21, pp. 393.

JOST,H.P. 1944, "A case for the quantitative inspection of surface finish." Machinery 65, pp. 483.

JUNGLES, J. and WHITEHOUSE, D.J. 1970, "An investigation of the shape and dimension of some diamond styli." J. Phys.E: Sci. Instrum., 3, pp. 437.

KALHOR, H.A. and NEUREUTHER, A.R. 1971, "Numerical method for the analysis of diffraction gratings." JOSA 61, pp. 43

KALISZER,H. GRIEVE, D.J. and ROWE, G.W. 1970, "Digital computation of surface topography." Proc. of the 11th Int. Machine Tool Des. & Res. Conf., Birmingham, U.K. September, 1970, pp. 543.

KAMNEV,V.V. 1966, "Integral evaluation of surface roughness." Measurement Techn. 2, pp. 261.

KAYSER, J.F. 1943, "Optical cut method for the determination of surface roughness." Foundry Trade J., 70, pp. 137.

KOEHLER, W.F. and WHITE, W.C. 1955, " Multiple-beam fringes of equal chromatic order. Part VI. Method of measuring roughness." JOSA 45, pp. 1011.

KOEHLER, W.F. 1953, "Multiple-beam fringes of equal chromatic order. Part II. Mechanism of polishing glass." JOSA 43, pp. 743.

KRANZ, O. 1980, "Untersuchungen des Abtastvorganges bei der Rauheitsmessung." P.T.B. Bericht PTB-ME-29.

LANG, J.E. and SCOTT, G.D. 1968 "Resolution limits in multiple-beam interferometry." JOSA 58, pp. 81.

LEADER, J.C. 1971, "Biderectional scattering of electromagnetic waves from rough surfaces." J. Appl. Phys. 42, pp. 4808.

LEGER, D. and PERRIN, J.C. 1976, "Real-time measurement of surface roughness by correlation of speckle pattern." JOSA 66, pp. 1210.

LIPPMANN, B.A. 1953, "Note on the theory of gratings." JOSA 43, pp. 408

LONARDO, P.M. 1974, "Measurement of smooth surface roughness by means of a photometric method." CIRP Annals 23, pp. 189.

LONARDO, P.M. 1978, "A new optical sensor for in-process detection of surface roughness." CIRP Annals 27/1, pp. 531.

LOU,D.Y, MARTINEZ,A. and STANTO,D. 1984, "Surface profile measurement with a dual beam optical system." Appl. Opt. 23, pp. 746.

LYNCH,P.J. 1970,"Curvature corrections to rough surface scattering at high frequencies." J. Acoust. Soc. Am. 47, pp 804.

MacBEAN, M.D.A. 1984, "Oblique interferometry of rough surfaces using a novel dove prism spectrometer." Appl. Opt. 23, pp. 4024.

MAKOSCH, G. and DROLLINGER, B. 1984, "Surface profile measurement with a scanning differential ac interferometer." Appl. Opt. 23, pp. 4544.

MATA MENDEZ,O., ROGER,A. and MAYSTRE,D. 1983, "Numerical solution for an inverse scattering problem of non-periodic rough surfaces." Appl. Phys. B32, pp. 199.

MAYSTRE,D. 1983, "Electromagnetic scattering from perfectly conducting rough surfaces in the resonance region." IEEE Trans. on Antennas and Propagation AP-31, pp. 885.

MAYSTRE,D. 1984, "Rigorous theory of light scattering from rough surfaces." J. Optics (Paris) 15, pp. 43.

MEECHAM,W.C. 1956, "Variational method for the calculation of the distribution of energy reflected from a periodic surface." J. Appl. Phys. 27, pp. 361.

MENDEZ, J.A. and ROBLIN, M.L. 1975, "Utilazition des franges d'interference en lumiere diffuse pour l'etude de l'etat de surface d'un diffuseur." Opt. Comm. 11, pp. 172.

MERCHANT, M.E. 1971, "Delphi type forecast of the future of production engineering." CIRP Annals 20, pp.213.

MILLAR,R.F. 1968, "On the Rayleigh assumption in scattering by a periodic surface." Proc. Camb. Phil. Soc. 65, pp. 773.

MILLAR,R.F. 1970a, "Singularities of two-dimensional exterior solutions of the Helmholtz equation." Proc. Camb. Phil. Soc. 69, pp. 175.

MILLAR,R.F. 1970b, "On the Rayleigh assumption in scattering by a periodic surface.II" Proc. Camb. Phil. Soc 69, pp. 217.

MILLAR,R.F. 1973, "The Rayleigh hypothesis and a related least square solution to scattering problems for periodic surfaces and other scatterers." Radio Science 8, pp. 786.

MIGNOT, J. and GORECKI, C. 1983, "Measurement of surface roughness: comparison between a defect-of-focus optical technique and the classical stylus technique." Wear, 87, pp. 38.

MITSUI,K. and SATO,H. 1976, "Development of an in process sensor for surface roughness by laser beam." Proc. 16th. Int. Conf. Mach. Tool and Res., pp. 171.

MOODY, J.C. 1968, "Measurement of ultrafine surface finishes." I.S.A. Trans., 7, pp. 67.

MOTYCKA,J. 1969, "Proposed interferometric method of measurement of roughness and autocorrelation function in smooth-finished surfaces." Appl. Opt. 8, pp. 1435.

MYERS,N.O. 1962 ,"Characterization of surface roughness." Wear, 5, pp. 182.

NAKAMURA, T. 1966, "On deformation of surface roughness curves caused by finite radius of stylus tip and tilting of stylus holder arm." Bull. Jap. Soc. Prec. Eng. 1, pp. 240.

NARA, J. 1966, "On CLA value obtained with direct reading surface roughness testers- effects of skid and high pass filter." Bull. Jap. Soc. Prec. Eng. 1, pp. 263.

NARA, J. 1969, "Two-dimensional representation of surface roughness." CIRP Annals 17, pp. 485.

NARA, J. 1971, "About the standardization and spectral measurement of surface waviness." CIRP Annals 19, pp. 687.

NAYAK, P.R. 1971, "Random process model of rough surfaces." Trans. ASME Ser. F. J. Lubr. Tech., 93, 3, pp. 398.

NELSON, H.R. 1969, "Taper sectioning as a means of describing the surface contour of metals." Proc. Conf. on Friction & Surface Finish, pp. 217.

NEVIERE,M. and CADILHAC,M. 1970, "Sur la validite du development de Rayleigh." Opt. Comm. 2, pp. 235.

NICOLLS, M.O. 1975, "The measurement of surface finish. Part 2." Metrology and Inspection, pp. 27.

OHLIDAL,I.,LUKES,F. and NAVRATAL,K. 1974, " Rough surface studies by optical methods." Surface Science 45, pp. 91.

OONISHI, Y. 1966, "Measurement of bearing curves of surface profiles." Bull. Jap. Soc. Prec. Eng. 1, pp. 249.

OSMAN, M.O.M. and SANKAR, T.S. 1975, "Profile characteristics of Manufactured surface using random function excursion technique. Part 2: application." Trans. ASME Ser. B. Engng. Ind., 97, pp. 196.

PARRY, G. 1974, "Some effects of surface roughness on the appearance of speckle in polychromatic light." Opt. Comm. 12, pp. 75.

PEKLINEK,J. 1963,"Contribution to the theory of surface characterization. " CIRP Annals 12, pp.173.

PEKLINEK,J. 1967 ,"Investigation of the surface typology." CIRP Annals 15/4.

PERES, N.J.C. 1953, "Geometrical considerations arising from the use of square wave calibrations standards of surface finish." Aust. J. Appl. Phys., 4, pp. 380.

PESANTE, M. 1955, "A new instrument for measuring surface roughness." Microtecnic, 9, pp. 27.

PESANTE, M. 1964, "Determination of surface roughness typology by means of amplitude density curves." CIRP Annals 12, pp. 61.

PETERS,J. 1965, "Messung des Mittenrauheitswertes zylindrischen Teile waehrend des Schleifens." VDI Berichte, 90, pp. 27.

PETERS,J. van HERCK,P. and SASTRODINOTO,M. 1979, "Assessment of surface topology analysis techniques." CIRP Annals 28, pp. 539.

PETIT,R. 1966, "Diffraction d'une onde plane par un reseau metallique." Revue d'Optique 45, pp. 249.

PETIT,R and CADILHAC,M 1966, "Sur la diffraction d'une onde plane par un reseau infiniment conducteur." C.R.Acad.Sc.Paris 262 serie B, pp. 468.

PETIT,R. 1980, "Electromagnetic theory of gratings." Book. Springer Verlag, Berlin heidelberg and New York.

PETTIGREW, R.M. and HANCOCK, F.J. 1979, "An optical profilometer." Prec. Eng. 1, pp. 133.

PORTEUS, J.O. 1963, "Relation between the height distribution of a rough surface and reflectance at normal incidence." JOSA 53, pp. 1394.

POWELL,R.W. 1957, "Experiments using simple thermal comparator for measurements of thermal conductivity, surface roughness and thickness of foils or of surface deposits." J. Sci. Instr., 34, pp. 485.

QUINEY, R.G. AUSTIN, F.R. and SARGENT, L.B. 1967, "The measurement of surface roughness and profiles on metals." ASME Trans. 19, pp. 193.

RABINOWICS, E. 1950, "Taper sectioning for the examination of metal surfaces." Metal Industry, 76, pp. 83.

RADHAKRISNAN,V. and SAGAR,V. 1970 ,"Surface roughness assessment by means of pnuematic measurement." Proc. 4th. All-India Machine Tool Des. & Res. Conf. (Indian Inst. Technology, Madras).

RADHAKRISNAN,V. 1977,"Application of inductive heads for non-contact measurement of surface finish." Proc. Int. Conf. Prodn. Engng.2, pp. 80. (Inst Engrs Calcutta,India).

RADHAKRISHNAN, V. 1970, "Effect of stylus radius on the roughness values measured with tracing stylus instruments." Wear 16, pp. 325.

RAJA,J. and WHITEHOUSE,D.J. 1984 ,"An investigation into the possibility of using surface profiles for machine tool surveillance." Int. J. Prod. Res. ,Vol 22, pp. 453.

RAKELS,J.H. 1975, "Een eerste aanzet tot de toetsing van..." Ingenieurs thesis, EM75-11, Technical University of Eindhoven.

RAKELS,J.H. and HINGLE,H.T. 1986a, "The use of optical diffraction techniques to obtain information about surface finish, tool shape and machine tool condition." WEAR 109, pp. 259.

RAKELS,J.H. 1986b, "Diffraction, an old optical phenomenon, used as an advanced metrology tool." International Conference on "Modern production and production metrology" T.U. Vienna, pp. 79.3.

RAKELS,J.H. 1986c, "The use of Bessel functions to extend the range of optical diffraction techniques for in-process surface finish measurements of high precision turned parts." J. Phys. E: Sci Instrum. 19, pp. 76.

RAKELS,J.H. 1987, "Computer simulation of an in-process surface finish sensor." SPIE proceedings Vol. 803.

RAU,N.,HUEBNER,G. and STAIGER,W. 1985a, "Optische Oberflaechenanalyse zum Bestimmen von Gestaltabweichungen." Werkstatt und Betrieb, 118, pp. 725.

RAU,N. and HUEBNER,G 1985b, "The assessment of surface structure with an optical sensor." Book "Towards the Factory of the Future." ed. H.J. Bollinger and H.J. Warnecke, Springer Verlag, pp. 550.

RAYLEIGH,LORD 1907, "On the dynamical theory of gratings." Proc. Roy. Soc., Ser A,79, pp. 399.

RAYLEIGH,LORD 1945, "The theory of sound." vol 2, Dover, New York.

REASON, R.E. 1944a, "Surface finish and its measurement." J.Inst. Prod. Engrs., 23, pp. 347.

REASON, R.E. HOPKINS, M.R. and GARROD, R.I. 1944b, "Report on the measurement of surface finish by stylus methods." Rank Taylor Hobson, Leicester, U.K. Report.

REASON, R.E. 1951, "Surface finish." Australasian Engr. 44, pp. 48.

REASON, R.E. 1954, "The trend of surface measurement." J. Inst. Prod. Engrs., 33, pp. 263.

REASON, R.E. 1956 "Significance and measurement of surface finish part 3: -conclusion. The interference method; roundness measurement; the best use of reference standards." Grinding & Finishing 2, pp. 33.

REASON, R.E. 1956, "Significance and measurement of surface part 2: how transducers affect instrument performance; how to select proper cut-off values." Grinding & Finishing 2, pp. 32.

REASON, R.E. 1960, "The measurement of surface texture."Modern workshop technology," 2 - processes Ed. Baker H.W. Book: publ. by Macmillan Ltd., London, U.K. Ch. 23, pp. 585.

RIBBENS,W.B. 1972, "Surface roughness measurement by holographic interferometry." Appl. Opt. 11, pp. 807.

RIBBENS,W.B. 1974, "Surface roughness measurement by two wavelengths holographic interferometry." Appl. Opt. 13, pp. 1085.

RUFFING, B. 1986, "Beruehrungslos-optische Rauheitsmessung technischer Oberflaechen mit dem Speckle Korrelationsverfahren." Proc. Int. Conf. on Modern Production and Production Metrology, April 1986, Technical University, Vienna.

SAWATARI, T. and ZIPIN, R.B. 1979, "Optical profile transducer." Opt. Eng. 18,, pp. 222.

SAYLES,R.S. and THOMAS,T.R. 1979, "Measurements of the statistical properties microtopography of engineering surfaces." Trans. ASME J.Lub. Techn. 101F, pp. 409.

SCHLESINGER, G. 1942, "Surface finish." Inst. of Prod. Engrs., Report.

SCHMALTZ, G. 1936, "Technische Oberflaechenkunde" Book. Springer Verlag, Berlin.

SCHWARTZ, N. and BROWN R. 1966, "A stylus method for evaluating the thickness of thin films and substrate surface roughness." Trans. of the 8th National Vacuum Symp., pp. 836.

SELVAM, M.S. 1975, "Tool vibration and its influence on surface roughness in turning." Wear 35, pp. 149.

SHAW, M.C. and PEKLINEK, J. 1963, "A light projection technique for studying surface topology." CIRP Annals 12/2, pp. 93.

SHERWOOD,K.F. and CROOKALL,J.R. 1967/68,"Surface finish assessment by an electrical capacitance technique." Proc. Instn. Mech. Engrs.182, part 3K, pp. 344.

SHIRAISI, M. 1983, In-process measurement of surface roughness in turning by laser beams."Trans. ASME J. Eng. for Ind. 103, pp. 203.

SILIN, R.I. and FREDERICK, J.R. 1967, "Instruments for the analysis of the profile of machined surfaces." CIRP Annals 15, pp. 295.

SILVER,S. 1949, "Microwave antenna theory and design." McGraw-Hill book company,inc, New York-Toronto-London.

SMITH, T. 1976, "Effect of surface roughness on ellipsometry of aluminium." Surface Science, 56, pp. 252.

SOMMARGREN, G.E. 1981, "An optical measurement of surface profile." Prec. Eng., 3, pp. 131.

SOMMERFELD, A. 1967, "Optics" Book. Academic Press, New-York.

SPRAGG, R.C. and WHITEHOUSE, D.J. 1970/71, "A new unified approach to surface metrology." Proc. I. Mech. E., 185, pp. 697.

SPRAGG, R.C. and WHITEHOUSE, D.J. 1974, "An average wavelength parameter for surface metrology." Rev. M. Mec. 20, pp. 293.

SPRAGUE, R.A. 1972, "Surface roughness measurement using white light speckle." Appl. Opt. 11, pp. 2811.

STANSBERG,C.T. 1979, "Surface roughness measurements by means of polychromatic speckle patterns." Appl. Opt. 18, pp. 4051.

STAUFERT,G. 1979,"Description of roughness profiles by separating the random and periodic components." Wear 57, pp. 185.

STOVER,J.C. 1975, "Roughness characterization of smooth machined surfaces by light scattering." Appl. Opt. 14, pp.1796.

STRATTON,J.A. 1941, "Electromagnetic theory." Book. McGraw-Hill Book Company,Inc New York and London.

STROKE,G.W. 1960, "Diffraction par les reseaux optiques." Revue d'Optique 39, pp. 356.

TALLIAN, T.E., CHIU, Y.P. HUTTENLOCHER, D.F., KAMENSHINE, J.A.,SIBLEY, L.B. and SINDLINGER, N.E. 1964, "Lubricant films in rolling contact of rough surfaces." A.S.L.E. Trans., 7, pp.109.

TALLIAN, T.E. and McCOOL, J.I. 1968, "The observation of individual asperity interactions in lubricated point contact." A.S.L.E. Trans., 11, 2, pp. 176.

TANIGUCHI,N. 1983, "Current status in, and future trends of, ultraprecision machining and ultrafine materials processing." CIRP Annal 32/2, pp 1.

TANNER, L.H. and FAHOUM, P. 1976a, "A study of the surface parameters of ground and lapped metal surfaces using specular and diffuse reflection of laser light." Wear 36, pp. 299.

TANNER, L.H. 1976b, "The use of laser light in the study of metal surfaces." Opt. and Laser Techn. 8, pp. 113.

TANNER,L.H. 1981, "A self-balencing pneumatic potentiometer and Wheatstone bridge with electrical readout." Prec. Eng. 4, pp. 202.

TARASOV, L.P. 1945, "Relation of surface roughness readings to actual surface profile." Trans. ASME 67, pp. 189.

TEAGUE,E.C., SCIRE,F.E. and VORBURGER,T.V. 1982, "Sinusoidal profile precision roughness specimens." Wear 93, pp. 61.

THOMAS,T.R and PROBERT,S.D. 1970,"Establishment of contact parameters from surface profiles." J. Phys. D: Appl. Phys., 3,pp. 277.

THOMAS,T.R. and PROBERT,S.D. 1972,"Correlations for thermal contact conductance in vacuo." Trans. ASME. J. Heat Transfer, 94C, pp. 176.

THOMAS,T.R.1979,"Calculation of elastic contact sytresses for rough curved surfaces." ASLE Trans. 22, pp. 184.

THOMAS,T.R.1981, "Characterization of surface roughness." Prec. Eng. 3, pp. 97.

THOMAS,T.R. and CHARLTON,G., 1981,"Variations of roughness parameters on some typical manufactured surfaces." Prec. Eng. 3, pp. 91.

THOMAS,T.R. 1982, "Rough surfaces." book.Ed. Thomas, Longmans, London and New York.

THWAITE, E.G. 1980, "Power spectra of rough surfaces obtained by optical Fourier transformation." CIRP Annals 29, pp. 419.

THWAITE, E.G. 1982a, "The extension of optical annular scattering techniques to the measurement of intermediate scale roughness." CIRP Annals 31, pp. 463.

THWAITE, E.G. 1982b, "A quantitative comparison of the wavelength spectrum of a surface obtained by optical Fourier transformation with calculations from profile measurements." Wear 83, pp. 181.

THWAITE,E.G. 1984, "Measurement and control of surface finish." Prec. Eng. 6, pp. 207.

TOLANSKY, S. 1948, "Multiple-beam interferometry of surfaces and films." Book. Clarendon Press, Oxford.

TOLANSKY, S. 1960, "Surface microtopography." Book Interscience, New York.

TORNEBOHM,H. 1936, "Modern tolerance requirements and their scientific determination." Mech. Eng. 58, pp. 411.

TRIBILLON,G. 1974, "Correlation entre deux speckles obtenus avec deux longueurs de l'onde- application a la measure de la rogusite moyenne." Opt. Comm. 11, pp. 172.

TSUKADO, T. and ANNO, Y. 1974, "An evaluation of machined surfaces topography. (1st report, on slope distribution of projections)." Bull. Jap. Soc. Prec. Eng. 8, pp. 141.

TSUKADA, T. and ANNO, Y. 1975, "An evaluation of machined surface topography. (2nd report, on statistics of surface asperity heights). Bull. Jap. Soc. Prec. Eng. 9, pp. 1.

TUCKER, R.C. and MEYERHOFF, R.W.1969, "A SEM study of surface roughness measurement." Proc. 2nd Annual Scanning Electron Microscopy Synp., Illinois Inst. of Technol. Res. Inst., Chicago, U.S.A., pp. 389.

UCHIDA, S-N, SATO, H. and O-HORI, M. 1979, "Two dimensional measurement of surface roughness by the light sectioning method. " CIRP Annals 28/1, pp. 419.

UNDERWOOD, A.F. and BIDDWELL, J.B. 1953, "New instrument for roughness measurement." Mach. & Tool Blue Book, 49, pp. 202.

VIENOT, J.C., GOEDGEBEUR J.P. and LACOURT, A. 1977, "Space and time variables in optics and holography: recent experimental aspects. " Appl. Opt. 15, pp. 454.

VORBURGER, T.V. and LUDEMA, K.C. 1980, "Ellipsometry of rough surfaces." Appl. Opt. 19, pp. 561.

VORBURGER,T.V. and TEAGUE,E.C. 1981 "Optical techniques for on-line measurement of surface topography." Prec. Eng. 2, pp 62.

VORBURGER,T.V., GILSINN,D.E., SCIRE,F.E., GIAUQUE,C.H.W. and TEAGUE, E.C. 1986, "Optical measurement of the roughness of sinusoidal surfaces." Wear 109, pp. 15.

WAGER, J.G. 1967, "Surface effects in pneumatic gauging." Int. J. Mach. Tool Des. Res., 7, pp. 1.

WAY, S. 1969, "Description and observation of metal surfaces." Proc. Conf on Friction & Surface Finish, pp. 44.

WEBSTER,J.A. and KALISZER,H. 1980, "In-process measurement of surface roughness and waviness." NELEX 80, Paper 5.6, National Engineering Laboratory East Kilbride, Glasgow.

WELFORD,W.T. 1977, "Review- Optical estimation of statistics of surface roughness from light scattering measurements." Opt, Quant. Elec. 9, pp. 269.

WELFORD,W.T. 1980, "Non contacting measurement of surface finish." SPIE 235, pp. 118.

WESTBERG, J. 1967/68, "Development of objective methods for judging the quality of ground and polished surfaces in production." Proc. Instn. of Mech. Engrs. 182 part 3K, pp. 260.

WHITEFIELD, R.J. 1975, "Noncontact optical profilometer." Appl. Opt. 14, pp. 2480.

WHITEHOUSE,D.J. 1967/68, "Improved type of wavefilter for use in surface-finish measurement." Proc. Instn. Mech. Engrs. Vol 182, pt.3K. pp.306.

WHITEHOUSE,D.J. and ARCHARD,J.F. 1970, "The properties of random surfaces and the significance in their contact." Proc. Roy. Soc. A316, pp. 97.

WHITEHOUSE,D.J. 1971 ,"Typology of manufactured surfaces." CIRP Annals 19/3. pp.417.

WHITEHOUSE, D.J. 1974a, "Theoretical analysis of stylus integration." CIRP Annals, 23, pp. 81.

WHITEHOUSE, D.J. 1974b, "The measaurement and analysis of surfaces." Tribology Int., 7, pp. 249.

WHITEHOUSE,D.J.,van HERCK,P.,de BRUIN,W. and van LUTTERVELT,C.A. 1974c, "Assessement of surface typology analysis techniques in turning" CIRP Annals 23/2.

WHITEHOUSE,D.J. 1974d, "Stylus techniques." in Characterization of Solid Surfaces, eds Kane,P.E. and Larrabee, G.P..Plenum New York.

WHITEHOUSE,D.J. 1978 ,"Surfaces- a link between manufacture and function." Proc. Instn Mech Engrs Vol 192: I. Mech. E. pp. 179.

WHITEHOUSE,D.J. and PHILIPS,M.J. 1978 ,"Discrete properties of random surfaces." Phil. Trans. Roy. Soc. A290, pp. 267.

WHITEHOUSE,D.J. 1982 ,"The parameter rash - is there a cure?" Wear 83 pp. 75.

WHITEHOUSE, D.J. and PHILIPS, M.J. 1985, "Sampling in a two-dimensional plane." J. Phys. A: Math. Gen. 18. pp. 2465.

WHITMAN, G. and SCHWERING, F. 1977, "Scattering by periodic metal surfaces with sinusoidal height profiles- A theoretical approach." IEEE Trans. on Antennas and Propagation, AP-25, pp. 869.

WILLIAMSON, D.E. 1947, "Tracer-point sharpness as affecting roughness measurements." Trans. ASLE 69, pp. 319.

WILLIAMSON, J.B.P. and HUNT, R.T. 1967/68, "Microtopography of surfaces." Proc. Inst. MEch. Engrs. Vol 182,part 3K, pp. 21.

WILLIAMSON, J.B.P. 1968a, "Physical aspects of a surface." A.S.T.M.E. Paper No. EM68-513,

WILLIAMSON, J.B.P. 1968b, "Topography of solid surfaces." In; Interdisciplinary approach to friction and wear. Ku, P.M. (Ed). NASA SP-181, pp. 85.

WILLN, J.E. 1972, "Characterisation of cylinder bore surface finish- A review of profile analysis." Wear 19, 2, pp. 143.

WIRGIN,A. 1980a, "On Rayleigh's theory of sinusoidal diffraction gratings." Opt. Acta 27, pp. 1671.

WIRGIN,A. 1980b, "Reflection from a corrugated surface." J. Acous. Soc. Am. 68, pp. 692.

WIRGIN,A. 1982, "Reflection by a grating: Rayleigh methods: components." JOSA 72, pp. 963.

WIRGIN, A. 1983, "Scattering from sinusoidal gratings: an evaluation of the Kirchoff approximation." JOSA 73, pp. 1028.

WOLFF, H. 1967,68, "Surface testing using the Hommel tester type T." Conf. on Properties & Metrology of Surfaces, Oxford. Proc. I. Mech. E., 182, Part 3K, pp. 279.

WYKES, C. 1977, "Decorrelation effects in speckle pattern interferometry. " Opt. Acta 24, pp. 517.

YAKOVLEV,E.A. 1965, "Calculation of the distribution of intensities by a diffraction grating in polarised light." Opt. Spektr. 19, pp. 417.

YANKEE,W.H., 1979,"Manufacturing processes." book. Prentice-Hall, Inc, Englewood, N.J.

YOUNG,R.d.,VORBURGER,T.V.and TEAGUE,E.C.,1980," In-process and on-line measurement of surface finish." CIRP Annals 29/1, pp. 435.

ZAKI,K.A. and NEUREUTHER,A.R. 1971, "Scattering from a perfectly conducting surface with a sinusoidal height profile: TE polarization." IEEE Trans. on Antennas and Propagation, AP19, pp. 208.

Appendix A0 Machining processes.

A0.1 Milling.

Together with the lathe, the milling machine is regarded as one of the most versatile machine tools. Practically all shapes and sizes of both flat and curved surfaces, on both the inside and the outside of workpieces can usually be machined by one or another milling method. It is an uniquely adaptable and economical method of producing a wide variety of machining operations in the manufacture of only a few parts to an almost unlimited production quantity. Milling offers the additional advantage of providing methods of repetitively creating accurate parts for interchangeable manufacture.

Process description.

Essentially, machined surfaces are formed by the action of a rotating multiple-tooth cutter, sometimes in a single pass of the work. The work may be held in a vise, a three-jaw chuck, an index head, a rotary table, between centres, in a special fixture, or bolted to the machine table. In most cases, the work is fed against the cutter. The speed of the cutting tool and the rate at which the workpiece travels is dependent upon the kind of material being machined. Optimum cutting effiency is also related to machine capacity, cutter design, workpiece requirements, workholding fixture design, and other governing factors. The metal removal rate of

milling machines, compared to lathes, shapers and planers, is usually considerably greater.

Types of milling operations.

In most conventional milling operations, the surfaces generated by a rotating toothed cutter may be classified into two general categories, peripheral and face milling.

In peripheral milling, the cutter rotates about an axis that is parallel to the surface being cut. Both flat and formed surfaces are produced in this way.

There are two different methods of generating surfaces by peripheral milling operations: up (or conventional) milling and down (or climb) milling.

Up milling is the condition when the work is fed against the direction of the rotating milling cutter. The chip is very thin where the tooth first contacts the work and increases in thickness to a maximum where the tooth breaks out of the work. The initial tooth contact usually occurs in clean metal and ends by lifting or peeling off the rough surface scale. This is the preferred method for machining sand castings, forgings, or metals that have a rough or hard abrasive surface scale.

Down milling differs from up milling in that the work is fed in the same direction as the rotation of the cutter. The cutter tooth starts into the work with a maximum cut thickness and ends with a thin chip, resulting in less cutter wear. Under the condition that there is no backlash in the feed system, down milling usually

produces a better surface finish on harder steels than does up milling. It is the method that is usually selected for milling operations on slender and intricate parts.

In face milling, the cutter rotates about an axis that is generally at right angles to the surface being cut. High-speed face milling with carbide-tipped cutters produces the best surface finish.

A0.2. Shaping.

The shaper is a machine that produces flat surfaces in horizontal, vertical, and angular planes. Two advantages of shapers are: (1) they use relatively inexpensive tools, and (2) for most types of work, they require only a short time for setup.

The leading disadvantage of a shaper is that its rate of metal removal is slow in comparison with a milling machine, because of the lost cutting time of the return stroke. Also, milling machines remove metal more rapidly because of the rapid cutting action of the multiple-tooth cutters.

Process description.

In shaping, a reciprocating single-point tool bit is used which is accurately guided as it moves back and forth in a straight line across a rigidly held workpiece. A ram, which moves back and forth, carries the tool holder and the cutting tool. The machined surface on the workpiece is generated by the cutting action of the tool bit as it peels off a chip on its forward stroke. As the ram returns, the

table that holds the workpiece feeds crosswise with a preset incremental distance equal to the feed desired. Since the return stroke of the ram represents nonproductive or lost time, the ram travels up to two times faster on the return stroke than on the forward or cutting stroke.

Types of shapers.

Shapers are classified into three distinctly different types: horizontal, vertical, and special-purpose. Most of these types are designed and designated specially for light-, medium- or standard-, or heavy-duty work.

Horizontal shapers are by far the most common of the three types. They are available with either a push-cut or a pull-cut stroke. Most have a universal table fitted which may be swivelled at any desired angle for making cuts.

Generally flat surfaces are produced by horizontal shapers. Curved or irregular surfaces and other miscellaneous shapes can be machined by carefully handfeeding the tool bit along a scribed line. Elaborate shapes can also be automatically produced by using machines equipped with hydraulically operated attachments for automatically contouring and duplicating

Vertical shapers are also known as slotters. They have a vertical ram and a hand- or power-operated rotary table. On some machines, the ram may be inclined when cutting inclined surfaces. Also included in this machine classification is a keyseater, which is a shaper principally designed for cutting keyways on the inside of

gears, pulleys, cams, and wheel hubs.

Special purpose shapers are principally confined to gear-cutting operations, like spur and helical gears. The tool bit, or cutter, is shaped like the gears it produces, except that each tooth is ground or relieved in such a way as to produce a cutting edge. The cutter is clamped on a ram and is gradually stroked into contact with one or more slowly rotating gear blanks. The relative motion of the gear cutter generates or shapes the desired gear teeth.

Shaper tools.

Cutting tools used on shapers are essentially the same as those used on the lathe except for slight differences in the tool angles. Tool bits are made of cast alloy and cemented carbides, but cutting tools made of high-speed steel are the most common. The main advantage of the single point cutting tool bit used on the shaper as compared to multiple point cutting tools used on broaching or milling machines, for example, is that they may be made and sharpened more easily.

The major disadvantage of a shaper is that its rate of metal removal is slow in comparison with a milling machine.

A0.3 Planing.

The planer is a machine which, like the shaper, produces flat surfaces in horizontal, vertical, or angular planes. There are two major differences between the shaper and the planer. The first

difference is that most planers operate with an action opposite to that of the shaper; that is the workpiece reciprocates past one or more stationary single-point cutting tools. A second difference is that planers are designed to accommodate workpieces which are far greater in size than those machined on the shaper.

Process description.

The work is clamped onto a large table which is constructed with T-slots and special holes that provide a means for positioning and securely holding the work. A wide assortment of clamps, jacks, stop pins, wedges, and special work-holding attachments are available to attach parts or fixtures to the table. The table rides in V-grooves on the bed of the machine and, accordingly, is accurately guided as it travels back and forth. Cutting tools are held in tool heads that can travel from side to side. The toolheads are mounted on a horizontal cross rail that can be moved up and down. Productivity is increased by using multiple toolheads, which permits several simultaneous cuts to be made with each relatively slow stroke of the table. Unlike the shaper, cutting may occur during both directions of the table travel. Toolheads may be installed with two or more cutter bits which are arranged in such a way that one or two cut in one direction of table travel and the remaining cut while the table travels in the opposite direction.

Whilst in most types of planers the workpiece reciprocates, with a pit-type planer the table remains stationary and the tool

reciprocates. This type of planer is used for extremely heavy work when the weight of the workpiece makes reciprocation difficult. Massive rails are mounted on both sides of the table. Toolheads, usually restricted to two in number, are mounted on the cross rail for two-way planing.

As in shaping, two or three cuts are considered normal practice for producing an acceptably finished surface.

A0.4 Broaching.

Broaching is a modern method of stock removal that has almost unlimited applications in metal working. It can also be used as a method to work wood, hard vulcanised fibre and other composites, hard rubber, graphite, and some plastics. Metals and alloys are the most commonly broached materials, however.

Process description.

The machine used may be hand-operated or electromechanically or hydraulically operated. Both the cutting tool (broach) and the workpiece are held in rigid fixtures. The principal function of the machine is to provide the speed necessary for cutting to take place. Cutting speed or travel of a broach across the work is usually 15 to 25 times faster than in milling. The feed (or removal rate) is regulated by the broach, which consists of a hardened steel bar with a series of cutting teeth. Each tooth is made progressively higher, to remove successively larger amounts of materials. Thus, the cut

grows deeper as the operation progresses. Cutting fluid is generously applied along the cut.

The process of surface finishing by broaching may be compared roughly to planing, except that in broaching a multitoothed tool is used. In both cases, the cutting tool may be either passed across a fixed workpiece or the tool may be held stationary and the work moved in a continuous stroke. In some machines, the tool is pushed along the surface to be broached while other machines pull the broaches through or over the work.

Broaching tools.

Except for variations in tooth sizes, teeth on a broaching tool resemble those on a wood rasp. The design of a broaching tool is based upon a concept unique to the process, in which rough, semifinish, and finish cutting teeth are combined in one tool. The difference in height between each tooth, or tooth rise, usually is greater along the roughing section and less along the finishing section. All finishing teeth are the same size.

Broaching tools may be of one-piece construction or they may consist of a series of segmented sections, which are assembled end to end on a moving slide. The entire length of such broaches can be supported. Sectional construction results in a broach which is easier and cheaper to construct and sharpen. Another advantage is that broken sections can be readily replaced without discarding other sections.

Burnishing broaches are used on parts when the surface finish

and accuracy requirements are critical. Burnishers are generally designed as push broaches. They are designed to polish rather than to cut a hole. Burnishing is generally restricted to soft, ductile materials.

Broaching machines.

In general, there are two types of broaching machines, horizontal and vertical, based upon the direction of broach travel. Most broaching machines are hydraulically powered, but some are driven electromechanically.

A0.5 Grinding.

Precision grinding and rough grinding each consist of forming surfaces by the use of a rotating abrasive wheel composed of many small hard bonded abrasive grains. Each individual and irregularly shaped grain acts as a cutting tool. Rough grinding is a commonly used method for removing excess material from castings, forgings, and weldments, or as a method of removing or snagging thin fins, sharp corners, or other unwanted projections from various shapes of workpieces. Small parts are often handheld and moved into contact with a rotating abrasive wheel. Surfaces on larger work are ground by manually moving a portable abrasive tool over the workpiece surfaces.

Description of the process.

Precision grinding is the principal production method of cutting materials that are too hard to cut by other conventional tools or for producing surfaces on parts to tolerance or finish requirements more exacting than can be achieved by other manufacturing methods. Surfaces on workpieces are simultaneously produced by grinding to an accurate size and with superior surface finish. The process of precision grinding is principally associated with the removal of small amounts of material to close tolerances and should not be confused with the process of abrasive machining. Abrasive machining relates to the rapid removal of relatively larger amounts of stock.

Abrasives.

Abrasives are substances that are both hard and tough. Efficient cutting action can only take place when the abrasive is sufficiently hard to penetrate and scratch the workpiece material. Toughness is a property that prevents premature fracturing and breaking away of the abrasive grains at the beginning of the cut. Maximum cutting efficiency of a grinding wheel is obtained when the dulled abrasive grains are continually fractured, thus exposing a succession of fresh new cutting edges.

The cutting action of a grinding wheel is similar to that of other cutting tools having multiple teeth. The pores on grinding wheels and the tooth spaces on milling cutters or broaching tools,

for example provide convenient spaces for the chips to escape and assist in the passage of cutting fluid to the cutting area.

Grinding fluids.

Grinding fluids are used when grinding metals, as well as for most nonmetallic materials. The function of grinding fluids is twofold: to cool the workpiece and to maintain the wheel face. The amount of heat generated is dependent upon the sharpness of the abrasive grains. Grinding fluids carry the chips away and improve the surface finish of the workpiece. In conventional dry grinding, the air acts as a grinding fluid. The oxygen in the air oxidises the workpiece surface. The action is sufficient to prevent the tendency of the chips to weld back onto the ground surface. Water-based fluids consisting of soluble-oil emulsions or synthetic compounds are also used as well as grinding oils, which are mineral oils containing fatty materials. Grinding fluids are applied in ample volume directly to the grinding area.

Grinding machines.

Broadly considered, there are three main categories of precision grinding machines for production applications: surface grinders, external cylindrical grinders, and internal cylindrical grinders.

The term surface grinding is understood to relate specifically

to the production of flat or plane surfaces. The two main types of surface grinders, horizontal and vertical, are classified according to the position of their wheel spindle. Each type of machine is available with a rotary or reciprocating work table.

External cylindrical grinders consist of two basic types: centre type and centreless. The chucking-type external cylindrical grinder is used for grinding relatively small parts with shapes that are adaptable for holding in a chuck or a collet. Chucking grinders are particularly useful for parts with shapes that cannot be readily held between centres.

Internal cylindrical grinders are designed to produce internally ground surfaces or holes on parts. There are three types of internal grinding machines: one that holds the workpiece in a rotating chuck, one in which the workpiece is rotated by the by the outside diameter between rolls and the planetary type, for heavy work, in which the work is held stationary.

In addition to the three main categories of production grinding machines previously discussed, there are a number of other standard and special types, each intended to accomplish a variety of grinding functions. Examples are tool grinders and multiple wheel camshaft grinders. A special form of flat surface grinder, the abrasive belt machine, is also commonly found in production shops, as well as disc grinders, cutter grinders, and gear grinders.

A0.6 Chemical milling.

Chemical milling is also known as photochemical machining, chemical machining, photofabrication, or photoetching. It is a well established process used to shape metals by removing unwanted metal by a controlled chemical attack. In principle, chemical milling originally evolved from the well established process of photoengraving, but the techniques, materials, controls, and purposes presently differ markedly from those of its forerunners. An early application of chemical milling was in producing printed circuit boards.

Description of the process.

There are two categories of chemical milling: blanking and contour machining. The basic difference in the two types of chemical milling is that blanking is a process used to etch entirely through a metal part, while contour machining is a method of selectively etching an area to some desired depth.

A0.7 Electrical discharge machining.

Electrical discharge machining, the process normally referred to as EDM, came into industrial use shortly after World War II. Its initial applications were in "tap busting". It was soon discovered,

however, that the process of electrical erosion could be controlled to machine cavities and holes. It was observed that the shape formed by this cutting is virtually a direct conforming image of the tool. EDM is now unquestionably recognised as an important precision machine-tool forming process for producing internal shapes on workpieces, traditionally the most difficult type of operation.

Description of the process.

Electrical discharge machining removes metal by the eroding action of small electrical sparks. The energy source for these sparks can be either an intermittently discharged capacitor or an electronically switched dc power source.

Normally, the power supply is connected so as to give the tool (electrode) negative polarity, and the workpiece positive polarity. In operation, both the electrode and workpiece are immersed in a dielectric fluid. As the negatively charged electrode approaches the positively charged workpiece, an electrical strain builds up across the dielectric until it ionises. Electrons then flow from the electrode, through the ionised path in the dielectric, to the workpiece. These electrons further ionise the dielectric, causing the electron flow to greatly increase, thus initiating an electronic avalanche, a spark. The spark, striking the workpiece at extremely high speed, melts and vapourises a small amount of material, leaving a crater in the workpiece surface. The vapourised metal, floating in the dielectric as a metallic cloud, condenses into small particles

that must be quickly flushed away.

In theory, any conductive material can be cut by electrical discharge machining. Workpiece hardness is of no concern in this process. The process results in completely burr-free parts.

Dielectric fluid.

A dielectric is a material that does not conduct electricity. It is pumped through the tool. The dielectric fluid used for this cutting process performs several functions:

i it aids in keeping the electrode and the workpiece cool.

ii it washes the debris out of the cut.

iii when ionised, it aids in stabilising the cut by maintaining a path of finely divided electrically conductive material in the gap.

Electrode material.

The prime requirements of any electrode material is that it must be electrically conductive and maintain a good tool-to-workpiece wear ratio. Electrode tool materials perform with varying degrees of success on different workpiece materials. Graphite and copper have been found to be the best materials for general use.

A0.8 Electron beam machining

Applications of electron beam technology other than welding are still in a state of development. It appears that much progress is still needed in research and technology in machining applications before it can be considered a well established, mature manufacturing process. Although not widespread, high volume operations such as drilling and perforating are currently being performed on commercial products. Of particular interest is the adaptability of the process to superalloys and refractory meterials, as well as to nonconducting materials such as ceramics or quartz.

Description of the process.

An electron beam may be defined as a stream of electrons all moving with about the same velocity and in the same direction so as to form a beam. Electron beam machining (EBM) is a thermoelectric process in which electrons are accelerated to ultrahigh velocities of nearly three-fourths the speed of light. Material is instantly removed by a melting and vapourising action caused by concentrating a high velocity narrow stream of electrons on a precisely limited area upon the workpiece. Upon impact with the workpiece, the kinetic energy of the electrons in the beam is converted to powerful heat energy. The beam power is continuously variable. It can pulsed on and off at rates up to at least 10 kHz, with typical pulse width from 10 to 80 us. EBM work is performed in a high vacuum chamber to eliminate the scattering of the beam of electrons as they contact gas molecules

on the workpiece.

A major factor in the effiency of the process is in conducting the heat away from the workpiece at the beam focus point. The cutting action is accomplished by a pulsing technique. This consists of repeatedly impinging the electron beam upon the work piece for a few milliseconds and then turning it off for a certain period of time. Careful control must be exercised over the heating and cooling cycle locally at the cut to prevent welding the surrounding metal. Short pulses are necessary to prevent ceramic material from cracking during machining.

Because the beam impact point is focussed on such a small spot, the precise position of the sharp point of the beam on the workpiece is of critical importance.

There is no mechanical contact between the tool and the workpiece. The distance from the electron gun to the workpiece is usually 100 mm.

A0.9 Ion beam machining.

Ion beam machining is an atomic-bit machining process and gives very high resolution. This method usually employs ions of argon or other inert gases with high kinetic energy in the order of 10keV. Ions which are the "tools" are caused to bombard and eject atoms from the workpiece surface by elastic collision. This phenomenon is called "ion sputter". The penetration depth of an ion at 1 keV is estimated from electron diffraction patterns to be about 5um.

Generally the method does not generate heat nor cause mechanical strain damage in the machined surface layer, but some of the ions are retained, substituting displaced workpiece atoms. This process has a machining resolution of 10 nm which contrasts with that of electron beam (thermal) machining.

Appendix A1 Summary of vector identities and vector operations.

In Cartesian coordinates a vector quantity $\vec{A}$ is written in the form:

$$\vec{A} = A_x \vec{e}_x + A_y \vec{e}_y + A_z \vec{e}_z \quad \text{(A1.1)}$$

where $\vec{e}_x$, $\vec{e}_y$ and $\vec{e}_z$ are unit vectors along the x,y and z axes respectively.

The magnitude of the vector $\vec{A}$ is the scalar

$$A = (A_x^2 + A_y^2 + A_z^2)^{\frac{1}{2}} \quad \text{(A1.2)}$$

Vectors can be added and subtracted:

$$\vec{A} \pm \vec{B} = (A_x \pm B_x)\vec{e}_x + (A_y \pm B_y)\vec{e}_y + (A_z \pm B_z)\vec{e}_z \quad \text{(A1.3)}$$

The scalar product

$$\vec{A} \cdot \vec{B} = A_x B_x + A_y B_y + A_z B_z \quad \text{(A1.4)}$$

obeys the commutative and distributive rules:

$$\vec{A} \cdot \vec{B} = \vec{B} \cdot \vec{A} \quad \text{(A1.5)}$$

$$\vec{A} \cdot (\vec{B} + \vec{C}) = \vec{A} \cdot \vec{B} + \vec{A} \vec{C} \quad \text{(A1.6)}$$

The vector product

$$\vec{A}\times\vec{B}=(A_yB_z-A_zB_y)\vec{e}_x+(A_zB_x-A_xB_z)\vec{e}_y+(A_xB_y-A_yB_x)\vec{e}_z \quad \text{(A1.7)}$$

follows the distributive rule

$$\vec{A}\times(\vec{B}+\vec{C})=\vec{A}\times\vec{B}+\vec{A}\times\vec{C} \quad \text{(A1.8)}$$

but not the commutative

$$\vec{A}\times\vec{B}=-\vec{B}\times\vec{A} \quad \text{(A1.9)}$$

The time derivative of $\vec{A}$ is:

$$\frac{d\vec{A}}{dt}=\frac{dA_x}{dt}\vec{e}_x+\frac{dA_y}{dt}\vec{e}_y+\frac{dA_z}{dt}\vec{e}_z \quad \text{(A1.10)}$$

The del operator is defined as:

$$\nabla=\frac{\partial}{\partial x}\vec{e}_x+\frac{\partial}{\partial y}\vec{e}_y+\frac{\partial}{\partial z}\vec{e}_z \quad \text{(A1.11)}$$

and the gradient of a scalar function f is written as:

$$\nabla f=\frac{\partial f}{\partial x}\vec{e}_x+\frac{\partial f}{\partial y}\vec{e}_y+\frac{\partial f}{\partial z}\vec{e}_z \quad \text{(A1.12)}$$

The div of $\vec{A}$ is

$$\nabla\cdot\vec{A}=\frac{\partial A_x}{\partial x}+\frac{\partial A_y}{\partial y}+\frac{\partial A_z}{\partial z} \quad \text{(A1.13)}$$

and the curl of $\vec{A}$ is defined by:

$$\nabla \times \vec{A} = \begin{vmatrix} \vec{e}_x & \vec{e}_y & \vec{e}_z \\ \frac{\partial}{\partial x} & \frac{\partial}{\partial y} & \frac{\partial}{\partial z} \\ A_x & A_y & A_z \end{vmatrix} \tag{A1.14}$$

The Laplacian is the div of the gradient

$$\nabla \cdot \nabla f = \nabla^2 f = \frac{\partial^2 f}{\partial x^2} + \frac{\partial^2 f}{\partial y^2} + \frac{\partial^2 f}{\partial z^2} \tag{A1.15}$$

The Laplacian of a vector $\vec{A}$ is

$$\nabla^2 \vec{A} = \nabla(\nabla \cdot \vec{A}) - \nabla \times \nabla \times \vec{A} \tag{A1.16}$$

A more extensive collection of vector identities and theorems can be found in [Van Bladel 1964]

Appendix A2 Mirror effect of a perfect, plane conductor.

In the thesis [Rakels 1975], formulae have been derived to predict the eddy current distributions in a plane, conducting metal sheet in the presence of a moving conductor (a version of a time varying field). Some of the results obtained therein, are equally applicable to prove that a perfect conductor acts as a mirror, i.e. that if an electromagnetic source Q is present at the position (x_o,y_o,z_o), then the field near the conductor, with the surface in the z=0 plane, can be represented by the field caused by a source Q' positioned at (x,y,-z).

To use the results obtained in that thesis, one has to remember that a plane metal reflector can be represented by a thin conducting sheet, since at optical frequencies the skin depth of the reflector will be extremely small, so that effectively the same conditions exist as mentioned in the initial lines.

The main argument followed is, that in this particular case a relationship can be established between the magnetic vector potentials $\vec{A}^o$ and $\vec{A}^i$ of respectively the inducing member and the induced currents. This relationship is:

$$\frac{\partial}{\partial t}\left(\vec{A}^o + \vec{A}^i\right) = w\,\frac{\partial}{\partial z}\vec{A}^i \qquad (z = \delta/2) \tag{A2.1}$$

where $w=2/u_o\sigma\delta$. The geometry used in this equation and the other quoted results can be found in the photocopied pages of the aforementioned thesis (renumbered A2.3 to A2.13).

If $\vec{A}^o$ is expressed by $f(x_o,y_o,z_o,t)$, then the solution of (A2.1)

is according to page A2.12

$$\vec{A}^{i}(z=\delta/2) = -\int_{0}^{\infty} \frac{\partial}{\partial t} f(x_o, y_o, -z_o - w\tau, t-\tau)\, d\tau \qquad \text{(A2.2)}$$

In the optical situation the expression for $\vec{A}^{o}$ can be written as a product of a spatial and time dependent functions, namely:

$$\vec{A}^{o} = f(x_o, y_o, z_o)\, e^{i\omega t} \qquad \text{(A2.3)}$$

insertion in equation (A2.2) and with w=0 ($\sigma = \infty$), yields:

$$\vec{A}^{i}(z=\delta/2) = -f(x_o, y_o, -z_o) \int_{0}^{-\infty} \frac{d}{dt} e^{i\omega(t-\tau)}\, d\tau \qquad \text{(A2.4a)}$$

or

$$\vec{A}^{i}(z=\delta/2) = C \cdot f(x_o, y_o, -z_o)\, e^{i\omega t} \qquad \text{(A2.4b)}$$

Comparing this result with (A2.3) shows that the magnetic vector potential $\vec{A}^{i}$ at the surface of the conductor (z=d/2) is proportional to a source equal to $\vec{A}^{o}$ but located at $(x_o, y_o, -z_o)$, so its position has been mirrored relative to the surface. The total magnetic potential at the surface is the summation of the original $\vec{A}^{o}$ and the induced $\vec{A}^{i}$. From these magnetic potentials the other electromagnetic quantities can be calculated by using Maxwell's equations.

From the foregoing it can be concluded that if an optical field U^{o} is produced by a source $Q(x_o, y_o, z_o)$, then the total field at the surface of the conductor can be seen as the summation of the original field U^{o} caused by the source at $Q(x_o, y_o, z_o)$ and a field proportional to U^{o} with its source located at $Q'(x_o, y_o, -z_o)$.

2.0 BEREKENING VAN DE WERVELSTROMEN IN EEN ONEINDIG UITGEBREIDE VLAKKE PLAAT, TEN GEVOLGE VAN BEWEGENDE STROOMVOERENDE GELEIDERS, DIE ZICH BOVEN EN/OF ONDER DE PLAAT BEVINDEN.

2.1 Het ontstaan van wervelstromen.

In een geleider geldt de wet van Ohm:

$$\overline{J} = \sigma \overline{E}$$

Deze wet legt een verband tussen de elektrische veldsterkte $\overline{E}$ en de stroomdichtheid $\overline{J}$ in een geleider. De evenredigheidsconstante tussen beide grootheden is het geleidingsvermogen σ.
De elektrische veldsterkte moet ook voldoen aan:

$$\nabla \times \overline{E} = -\frac{\partial \overline{B}}{\partial t}$$

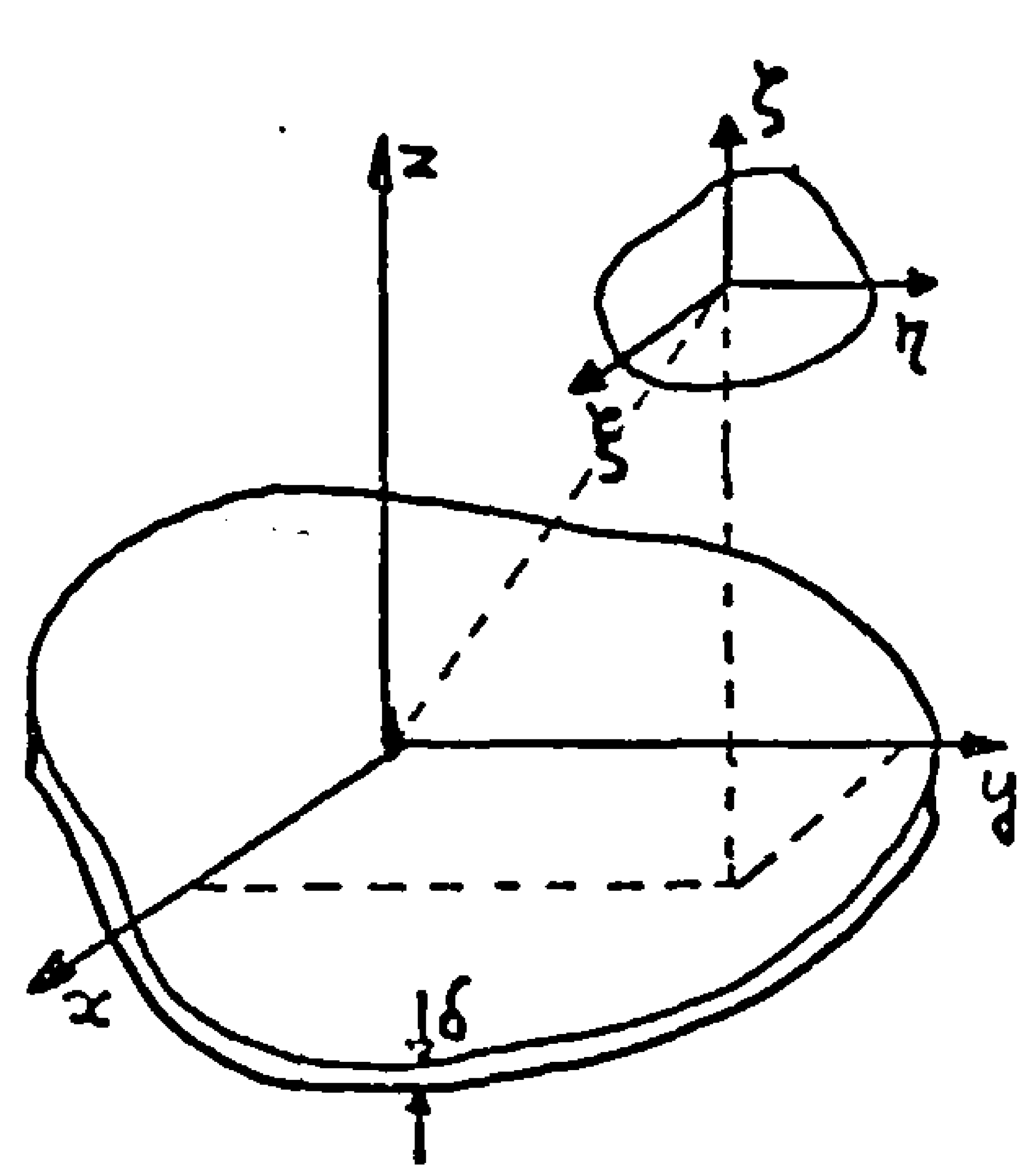
waarin $\overline{B}$ de magnetische inductie is.
Uit deze formules blijkt dat er wervelstromen in een geleider ($\sigma \neq 0$) ontstaan, als deze zich in een tijdsafhankelijk magneetveld bevindt.

2.2 Problaemstelling.

Er wordt van de volgende configuratie uitgegaan;

1° de plaat is oneindig uitgebreid, heeft een dikte $\delta \to 0$ en heeft het vlak z = 0 als symmetrievlak.
2° het geleidingsvermogen σ van de plaat heeft overal dezelfde constante waarde
3° $\lim\limits_{\substack{\sigma \to \infty \\ \delta \to 0}} (\sigma . \delta)$ = eindig
4° de magnetische permeabiliteit en de permittiviteit is over de gehele ruimte gelijk aan μ_0 resp. ε_0.
5° de stroomvoerende geleiders bewegen zich evenwijdig aan de plaat
6° de inducerende geleiders liggen ieder voor zich in vlakken waarvoor geldt: z = constant .
7° in de plaat vloeien alleen stromen in de x- en y-richting.
8° de waarnemer staat stil ten opzichte van de plaat.
9° aan de plaat is het coördinatenstelsel x,y en z vastverbonden
10° aan de geleiders is het coördinatenstelsel ξ, η en ζ vastverbonden.
11° de stromen in de plaat zyn onafhankelijk van z.

Hieronder staat een schets van een mogelijke configuratie.



2.3 Berekening van wervelstromen.

Bij de berekeningen wordt gebruik gemaakt van de vergelijkingen van Maxwell, de daaruit volgende continuiteitsvergelijking en de wet van Ohm.

$$\nabla \times \bar{E} = -\frac{\partial \bar{B}}{\partial t} \quad ; \quad \nabla \times \bar{H} = \bar{J} + \frac{\partial \bar{D}}{\partial t}$$

$$\nabla \cdot \bar{B} = 0 \quad ; \quad \nabla \cdot \bar{D} = \rho$$

$$\nabla \cdot \bar{J} + \frac{\partial \rho}{\partial t} = 0$$

$$\bar{J} = \sigma \bar{E}$$

verder geldt nog:

$$\bar{D} = \varepsilon \bar{E} \text{ en } \bar{B} = \mu \bar{H} .$$

Deze vergelijkingen zijn in goede geleiders te vereenvoudigen. Men kan dan de ladingsdichtheid binnenin isotrope homogene geleiders als volgt berekenen:

$$\frac{\partial \rho}{\partial t} + \nabla \cdot \bar{J} = \sigma \nabla \cdot \bar{E} + \frac{\partial \rho}{\partial t} = \frac{\sigma}{\varepsilon} \nabla \cdot \bar{D} + \frac{\partial \rho}{\partial t} = \frac{\sigma}{\varepsilon} \rho + \frac{\partial \rho}{\partial t} = 0 \quad \text{, zodat}$$

$$\rho = \rho_0 e^{-\frac{\sigma}{\varepsilon} t} = \rho_0 e^{-\frac{t}{\tau}}$$

Hieruit blijkt dat de ladingsdichtheid in de geleider afneemt. De lading zal zich eventueel aan de randen van de geleider ophopen. Bij geleiders zoals koper en aluminium is de relaxatietijd τ ongeveer 10^{-18} sec. zodat de ladingsconcentraties binnenin de geleider

bizonder snel verdwijnen. In het geval van een oneindig uitgestrekte vlakke plaat zal de lading zich eventueel aan de boven- en onderkant van de plaat bevinden. In de z-richting treedt dan een evenwicht op zodat J_z gelijk aan nul wordt. Verder mag $\frac{\partial D}{\partial t} = \frac{\varepsilon_0}{\sigma}\frac{\partial \bar{J}}{\partial t}$ verwaarloosd worden vanwege de lage frequenties. De vergelijkingen kunnen nu als volgt geschreven worden:

$$\nabla \times \bar{E} = -\frac{\partial \bar{B}}{\partial t} \quad ; \quad \nabla \times \bar{H} = \bar{J}$$

$$\nabla \cdot \bar{B} = 0 \quad ; \quad \nabla \cdot \bar{D} = 0$$

$$\nabla \cdot \bar{J} = 0$$

$$\bar{J} = \sigma \bar{E}$$

$$\bar{D} = \varepsilon \bar{E} \quad ; \quad \bar{B} = \mu_0 \bar{H}$$

2.4 Afleiding van de formule $\frac{\partial}{\partial t}(\bar{A}^o + \bar{A}^w) = w \frac{\partial}{\partial z}\bar{A}^w$

Bij de berekening van de wervelstromen wordt in dit geval gebruik gemaakt van de magnetische vectorpotentialen $\bar{A}^o$ van de inducerende en $\bar{A}^w$ van de geinduceerde stromen. Omdat deze stromen geen z-component bezitten, hebben de magnetische vectorpotentialen ook geen z-component.

De wervelstromen in de plaat zijn het gevolg van de magnetische inductie $\bar{B}$. Deze $\bar{B}$ wordt veroorzaakt door de inducerende en de geinduceerde stromen.

$$\bar{B} = \bar{B}^o + \bar{B}^w \tag{1}$$

waarbij $\bar{B}^o$ het veld van de inducerende en $\bar{B}^w$ het veld van de geïnduceerde stromen is.

Voor $\bar{B}^o$ en $\bar{B}^w$ kan men het volgende schrijven:

$$\bar{B}^o = \nabla \times \bar{A}^o \quad ; \quad \bar{B}^w = \nabla \times \bar{A}^w \tag{2}$$

Uit (1) en (2) volgt:

$$\bar{B} = \bar{B}^o + \bar{B}^w = \nabla \times \bar{A}^o + \nabla \times \bar{A}^w = \nabla \times (\bar{A}^o + \bar{A}^w) = \nabla \times \bar{A} \tag{3}$$

In de plaat geldt de wet van Faraday:

$$\nabla \times \bar{E} = -\frac{\partial \bar{B}}{\partial t} \tag{4}$$

Als men nu $\bar{B}$ vervangt door $\nabla \times \bar{A}$ en dit naar de tijd differentieert, dan kan men schrijven:

$$\nabla \times \bar{E} = -\frac{\partial}{\partial t} \times \bar{A} \qquad \text{of} \tag{5a}$$

$$\nabla \times (\bar{E} + \frac{\partial}{\partial t}\bar{A}) = \bar{0} \tag{5b}$$

Hieruit blijkt dat $\bar{E} + \frac{\partial \bar{A}}{\partial t}$ een vector is waarvan de rotatie gelijk is aan $\bar{\theta}$. Daarom is deze vector de gradient van iets. In de elektrostatica komt dit ook voor want daar is $\nabla \times \bar{E} = \bar{0}$. Daar geeft men dan als oplossing $\bar{E} = -\nabla\phi$, waarin ϕ de elektrostatische potentiaal is.

Deze oplossing kan men nu ook nemen

$$\bar{E} + \frac{\partial \bar{A}}{\partial t} = -\nabla\phi \tag{6}$$

Verder geldt in de plaat de wet van Ohm.

$$\bar{J}^{W} = \sigma \bar{E} \tag{7}$$

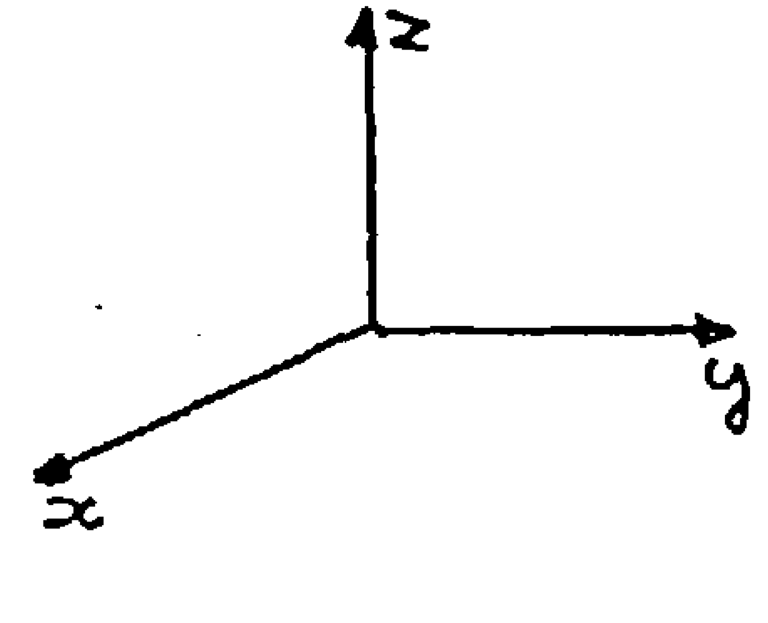
$\bar{J}^{W}$ is de geïnduceerde wervelstroom

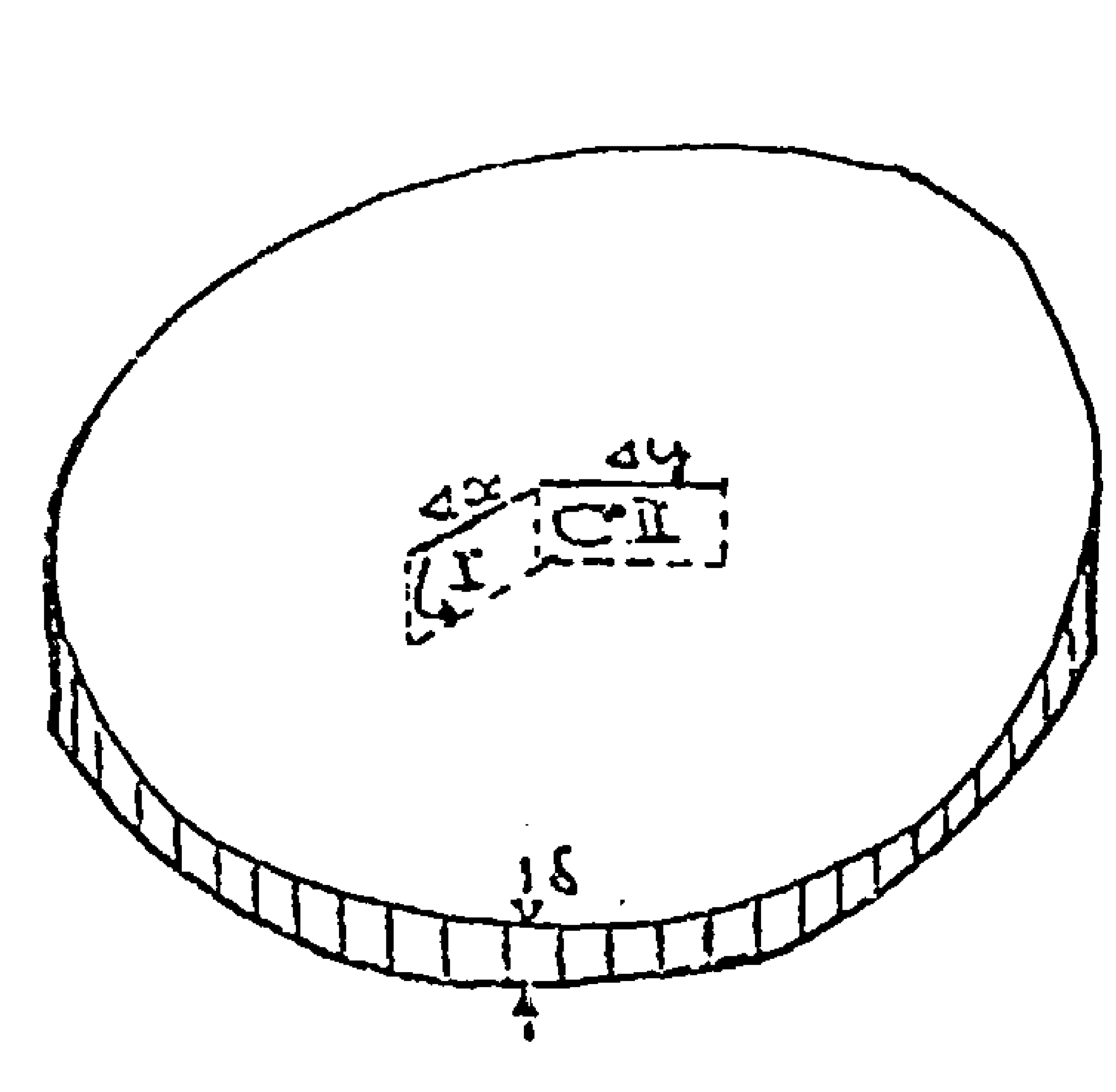
Uit de formules (6) en (7) volgt:

$$\bar{E} = -\frac{\partial \bar{A}}{\partial t} - \nabla\phi = \frac{1}{\sigma}\bar{J}^{W} \qquad \text{of} \tag{8a}$$

$$\bar{E} = -\frac{\partial}{\partial t}(\bar{A}^{o} + \bar{A}^{W}) - \nabla\phi = \frac{1}{\sigma}\bar{J}^{W} \tag{8b}$$

Met behulp van de kringintegraal voor de magnetische veldsterkte $\bar{H}^{W}$, ten gevolge van $\bar{J}^{W}$, is er een verband te leggen tussen $\bar{J}^{W}$ en $\bar{A}^{W}$ (zie onderstaande tekening)





De integratiewegen Δx en Δy worden zo klein verondersteld, dat men de magnetische veldsterkte $\overline{H}^w$ constant mag beschouwen. De integratie in de z-richting, δ, nadert tot nul zodat men het produkt van δ en $\overline{H}^w_z$ kan verwaarlozen. Vanwege de symmetrie geldt verder nog dat

$$\overline{H}^w(+\tfrac{\delta}{2}) = -\overline{H}^w(-\tfrac{\delta}{2}) \qquad (9)$$

Voor de integratieweg I geldt:

$$\oint_I \overline{H}.d\overline{l} = H^w_x(+\tfrac{\delta}{2})\Delta x - H^w_x(-\tfrac{\delta}{2})\Delta x = 2H^w_x(+\tfrac{\delta}{2})\Delta x = \delta J^w_y \Delta x \qquad (10a)$$

Voor de integratieweg II geldt:

$$\oint_{II} \overline{H}.d\overline{l} = H^w_y(+\tfrac{\delta}{2})\Delta y - H^w_y(-\tfrac{\delta}{2})\Delta y = 2H^w_y(+\tfrac{\delta}{2})\Delta y = -\delta J^w_x \Delta y \qquad (10b)$$

Verder geldt:

$$\overline{B} = \mu_0 \overline{H} \qquad (11)$$

Uit de formules (10) en (11) volgt:

$$J^w_x = \frac{-2}{\mu_0\delta} B^w_y(+\tfrac{\delta}{2}) = \frac{2}{\mu_0\delta} B^w_y(-\tfrac{\delta}{2}) \qquad (12a)$$

$$J^w_y = \frac{2}{\mu_0\delta} B^w_x(+\tfrac{\delta}{2}) = \frac{-2}{\mu_0\delta} B^w_x(-\tfrac{\delta}{2}) \qquad (12b)$$

Voor B^w_x en B^w_y kan geschreven worden omdat $A^w_z=0$

$$B^w_x = \frac{\partial A^w_z}{\partial y} - \frac{\partial A^w_y}{\partial z} = -\frac{\partial A^w_y}{\partial z} \qquad \rightarrow \qquad J^w_y = \frac{-2}{\mu_0\delta}\frac{\partial}{\partial z}A^w_y \qquad (13a)$$

$$B^w_y = \frac{\partial A^w_x}{\partial z} - \frac{\partial A^w_z}{\partial x} = \frac{\partial A^w_x}{\partial z} \qquad\qquad J^w_x = \frac{-2}{\mu_0\delta}\frac{\partial}{\partial z}A^w_x \qquad (13b)$$

Uit de formules (12) en (13) volgt:

$$\overline{J}^w = \frac{-2}{\mu_0\delta}\frac{\partial}{\partial z}A^w(+\tfrac{\delta}{2}) = \frac{2}{\mu_0\delta}\frac{\partial}{\partial z}A^w(-\tfrac{\delta}{2}) \qquad (14)$$

Uit de formules (8) en (14) volgt:

$$\frac{\partial}{\partial t}(\overline{A}^o + \overline{A}^w) + \nabla\phi = \frac{2}{\mu_0\sigma\delta}\frac{\partial}{\partial z}A^w(+\tfrac{\delta}{2}) = w\frac{\partial}{\partial z}A^w(+\tfrac{\delta}{2}) = -w\frac{\partial}{\partial z}A^w(-\tfrac{\delta}{2}) \qquad (15)$$

met $w = \frac{2}{\mu_0\sigma\delta}$ (16)

Als nu van formule (15) de divergentie wordt genomen, dan verkrijgt men het volgende resultaat omdat $\nabla.\overline{A}=0$ mag worden gesteld.

$$\nabla \cdot \nabla \phi = \nabla^2 \phi = \frac{\partial^2 \phi}{\partial x^2} + \frac{\partial^2 \phi}{\partial y^2} + \frac{\partial^2 \phi}{\partial z^2} = 0 \qquad (17)$$

Daar A_z^o en A_z^w overal gelijk zijn aan nul is ook $\frac{\partial \phi}{\partial z}$ overal gelijk aan nul zodat geldt:

$$\frac{\partial^2 \phi}{\partial z^2} 0 \qquad (18)$$

Uit (17) en (18) volgt:

$$\frac{\partial \phi}{\partial x^2} + \frac{\partial \phi}{\partial y^2} = 0 \qquad (19)$$

Formule (19) kan men ook in poolcoördinaten schrijven

$$\frac{\partial^2 \phi}{\partial r^2} + \frac{1}{r}\frac{\partial \phi}{\partial r} + \frac{1}{r}\frac{\partial^2 \phi}{\partial \varphi^2} = 0 \qquad (20)$$

met $x * r\cos\varphi$ en
$y = r\sin\varphi$

Differentiaalvergelijking (20) kan met de methode van separatie van variabelen worden opgelost door te stellen

$$\phi(r,\varphi) = R(r).P(\varphi) \qquad (21)$$

dit geeft

$$\frac{d^2R}{dr^2}.P + \frac{1}{r}\frac{dR}{dr}.P + \frac{1}{r^2}R\frac{d^2P}{d\varphi^2} 0 \qquad (22a)$$

$$\frac{1}{R}\left(r^2\frac{d^2R}{dr^2} + r\frac{dR}{dr} \right) = -\frac{1}{P}\frac{d^2P}{d\varphi^2} \qquad (22b)$$

Het linkerlid is onafhankelijk van φ, het rechterlid is onafhankelijk van r. Hieruit volgt, dat beide leden onafhankelijk van r en φ, dus constant zijn. Noem deze constante c, dan

$$r^2\frac{d^2R}{dr^2} + r\frac{dR}{dr} - cR = 0 \qquad (23)$$

$$\frac{d^2P}{d\varphi^2} + cP = 0 \qquad (24)$$

Omdat $P(\varphi)$ periodiek is met periode 2π, zijn de oplossingen van vergelijking (24) slechts bruikbaar indien:

$c = 0$ met P constant, of $c = n^2$, n natuurlijk, met $P = a\cos n\varphi$ en/of $P = b\sin n\varphi$

Voor deze c heeft (23) de oplossingen:

$R = d+e.\ln(r)$ als $c = 0$; $R = r^n$ en $R = r^{-n}$ als $c = n^2$

Op fysische gronden moeten de oplossingen waarvoor $\phi(r,\varphi)$ onbegrensd is worden verworpen. Dan blijven de volgende oplossingen over.

ϕ = constant, $\phi = r^{-n}\cos n\varphi$ en $\phi = r^{-n}\sin n\varphi$ voor alle natuurlijke n.

$$\phi(r,\varphi) = a_o + \sum_{n=1}^{\infty}(a_n r^{-n}\cos n\varphi + b_n r^{-n}\sin n\varphi) \qquad (25)$$

In deze formule zijn a_n en b_n constanten.
Van deze functie $\phi(r,\varphi)$ kan men het volgende zeggen:

1° ϕ is begrensd
2° ϕ voldoet in rechthoekige coördinaten aan de Laplace vergelijking $\nabla^2\phi = 0$
3° ϕ heeft overal continue eerste en tweede orde afgeleiden

Nu is er een definitie die zegt:

Een functie die continue eerste en tweede orde afgeleiden heeft in een omgeving van O en aan de Laplace vergelijking voldoet in O, is harmonisch in O.

ϕ voldoet overal aan deze definitie dus ϕ is harmonisch.

Nu zegt het theorema van Liouville:

Een functie, die harmonisch en begrensd is in de gehele n-dimensionale ruimte, is constant.

Hieraan wordt door ϕ voldaan. Dus geldt: ϕ= constant en

$$\nabla\phi = 0 \qquad (26)$$

Uit (15) en (26) volgt nu ;

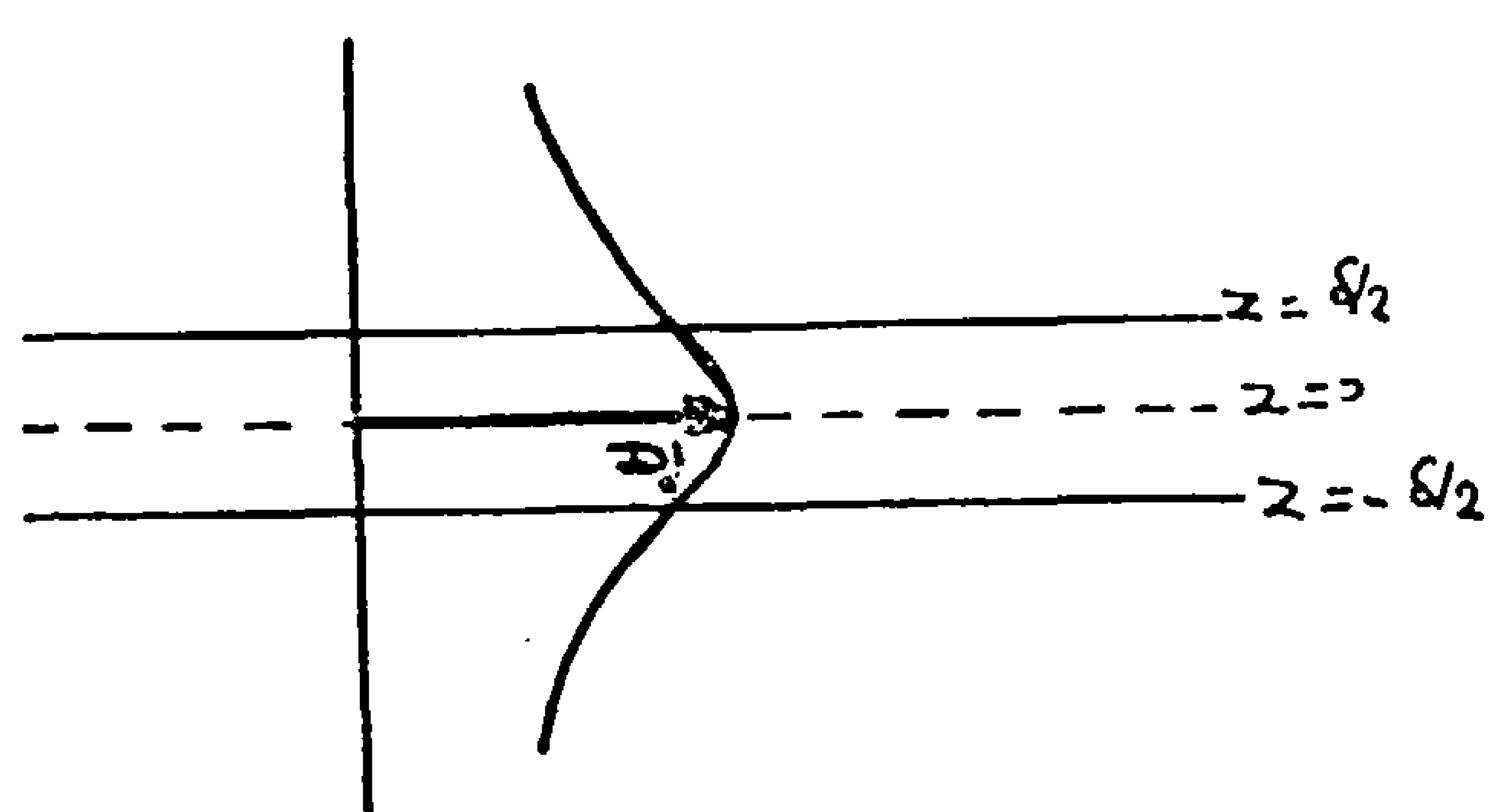
$$\frac{\partial}{\partial t}(\bar{A}^o + \bar{A}^w) = w\frac{\partial}{\partial z}\bar{A}^w(z=+\delta/2) = -w\frac{\partial}{\partial z}\bar{A}^w(z=-\delta/2) \qquad (27)$$

2.5 Afleiding van de formule $\bar{A}^w = -\int_0^\infty \frac{\partial f(x_o, y_o, \mp|z_o| \pm w\tau, t-\tau)}{\partial t} d\tau \quad (z=\mp\delta/2)$

Er wordt uitgegaan van de formule

$$\frac{\partial}{\partial t}(\bar{A}^o + \bar{A}^w) = -\frac{1}{\sigma}\bar{J}^w = \pm w \frac{\partial}{\partial z}\bar{A}^w \quad (z=\pm\delta/2) \tag{27}$$

De magnetische vectorpotentiaal, ten gevolge van de wervelstromen $\bar{J}^w$, heeft het vlak z=0, waarin de plaat zich bevindt, tot symmetrievlak. Het verloop van $\bar{A}^w$ in de buurt van de plaat is hieronder geschetst.



Voor $\bar{A}^w$ geldt

$$\frac{\partial}{\partial z}\bar{A}^w_{z=\delta/2} = -\frac{\partial}{\partial z}\bar{A}^w_{z=-\delta/2} \tag{28}$$

$$\bar{A}^w_{z=\delta/2} = \bar{A}^w_{z=-\delta/2} \approx \bar{A}^w_{z=0} \tag{29}$$

Veronderstel dat $\bar{A}^o$ constant is en er wel stromen in de plaat vloeien, dan kan voor (27) geschreven worden:

$$\frac{\partial}{\partial t}\bar{A}^w = w\frac{\partial}{\partial z}\bar{A}^w_{z=\delta/2} = -w\frac{\partial}{\partial z}\bar{A}^w_{z=-\delta/2} \tag{30}$$

De oplossingen hiervan zijn

$$\bar{A}^w = f(x,y,z+wt) \quad \text{als } z=+\delta/2 \quad \text{of} \tag{31a}$$

$$\bar{A}^w = f(x,y,z-wt) \quad \text{als } z=-\delta/2 \tag{31b}$$

Hieruit blijkt dat de magnetische vectorpotentiaal $\bar{A}^w$ voor $z=\delta/2$ gezien kan worden als het gevolg van een vaste wervelstroomverdeling die zich met een snelheid w in de negatieve z-richting van de plaat afbeweegt. Voor $z=-\delta/2$ kan $\bar{A}^w$ gezien worden als het gevolg van vaste wervelstroomverdeling die zich met een snelheid w in de positieve z-richting van de plaat afbeweegt.

Veronderstel nu dat $\vec{A}^o$ plotseling verandert. Als deze verandering zo snel gebeurt dat de tijd, waarin dit gebeurt, naar nul nadert dan worden er op dat moment wervelstromen opgewekt zodanig dat in de plaat geldt:

$$d\vec{A}^w = -d\vec{A}^o \tag{32}$$

waarin $d\vec{A}^o$ de verandering van $\vec{A}^o$ is. Dit moet gelden omdat het rechterlid van (27) eindig is.

Met andere woorden als er zich een $d\vec{A}^o$ voordoet dan wordt daarop door de plaat gereageerd met een $d\vec{A}^w$ die in eerste instantie $d\vec{A}^o$ volledig compenseert in de plaat, daarna sterft deze $d\vec{A}^w$ volgens (31) uit.
Dit kan in rekening worden gebracht door een bron $d\vec{A}^w$, die tegengesteld is aan $d\vec{A}^o$, en die zich op het moment van de verandering van $\vec{A}^o$ op de plaats van $d\vec{A}^o$ bevindt of op de plaats van het spiegelbeeld van $d\vec{A}^o$ ten opzichte van de plaat. Deze bron $d\vec{A}^w$ beweegt zich na de verandering van $\vec{A}^o$ van de plaat af volgens (31).
Een volgende plotselinge verandering van $\vec{A}^o$ brengt weer een nieuw stelsel wervelstromen met een bijbehorende $d\vec{A}^w$ met zich mee, die zich ook van de plaat afbeweegt enz...
Het momentane veld $\vec{A}^w$ is een sommatie van alle voorgaande uitstervende $d\vec{A}^w$'s.

Stel dat de bron van $d\vec{A}^o$ zich boven de plaat bevindt in het punt (x_o, y_o, z_o), dan zal op het moment dat $d\vec{A}^o$ zich voordoet de bron $d\vec{A}^w$ zich bevinden in (x_o, y_o, z_o) of in $(x_o, y_o, -z_o)$, daarna zal $d\vec{A}^w$ zich van de plaat afbewegen met een snelheid w. Na een tijdsduur τ zal de bron $d\vec{A}^w$ zich in $(x_o, y_o, z_o + w\tau)$ of in $(x_o, y_o, -z_o - w\tau)$bevinden.
Als de bron van $d\vec{A}^o$ zich onder de plaat in (x_o, y_o, z_o) bevindt, dan zal op het moment dat $d\vec{A}^o$zich voordoet $d\vec{A}^w$ zich in $(x_o, y_o, -z_o)$ of in (x_o, y_o, z_o) bevinden. Na een tijdsduur τ zal $d\vec{A}^w$ zich in $(x_o, y_o, -z_o + w\tau)$ of in $(x_o, y_o, z_o - w\tau)$ bevinden.
In het algemene geval dat de bron van $d\vec{A}^o$ zich boven of onder de plaat bevindt in (x_o, y_o, z_o) dan zal de bron van $d\vec{A}^w$ zich na een tijdsduur τ , na de verandering van $\vec{A}^o$, in $(x_o, y_o, |z_o| + w\tau)$ of in $(x_o, y_o, -|z_o| - w\tau)$ bevinden

Als de grootten van de discontinue veranderingen $d\overline{A}^{o}$ van het veld $\overline{A}^{o}$ en de intervals tussen deze veranderingen kleiner worden, dan wordt een limietbenadering verkregen van een continu veranderend veld $\overline{A}^{o}$.

De vectorpotentiaal ter plaatse (x,y,z) tengevolge van een stilstaande bron ter plaatse (x_o,y_o,z_o) is een functie van (x,y,z,x_o,y_o,z_o,t) kortheidshalve wordt dit aangeduid met

$$\overline{A}^{o} = f(x_o,y_o,z_o,t) \qquad \text{zie fig. a} \tag{33}$$

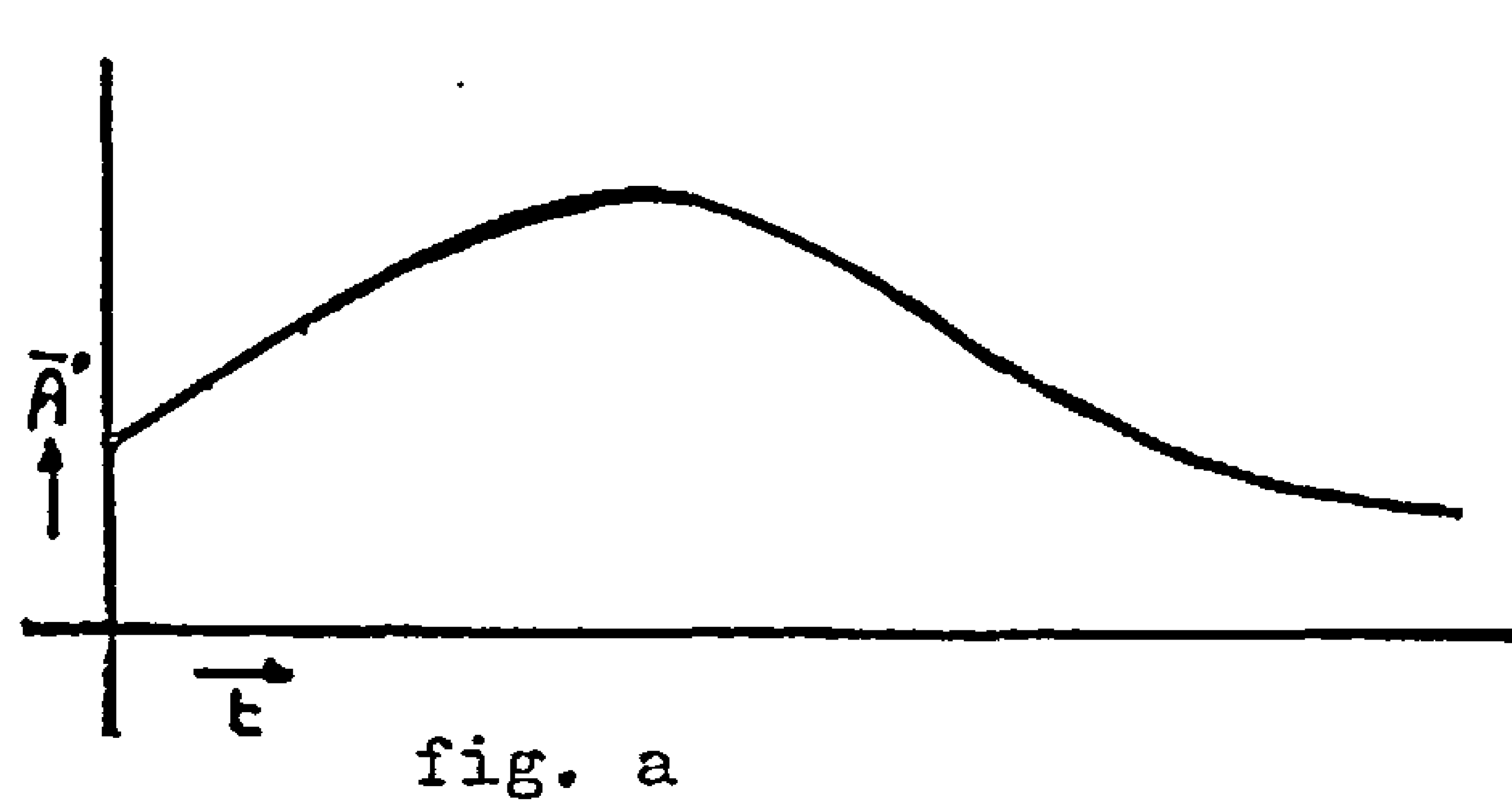


fig. a

In fig. b is het verloop van $\frac{\partial \overline{A}^{o}}{\partial t}$ geschetst.

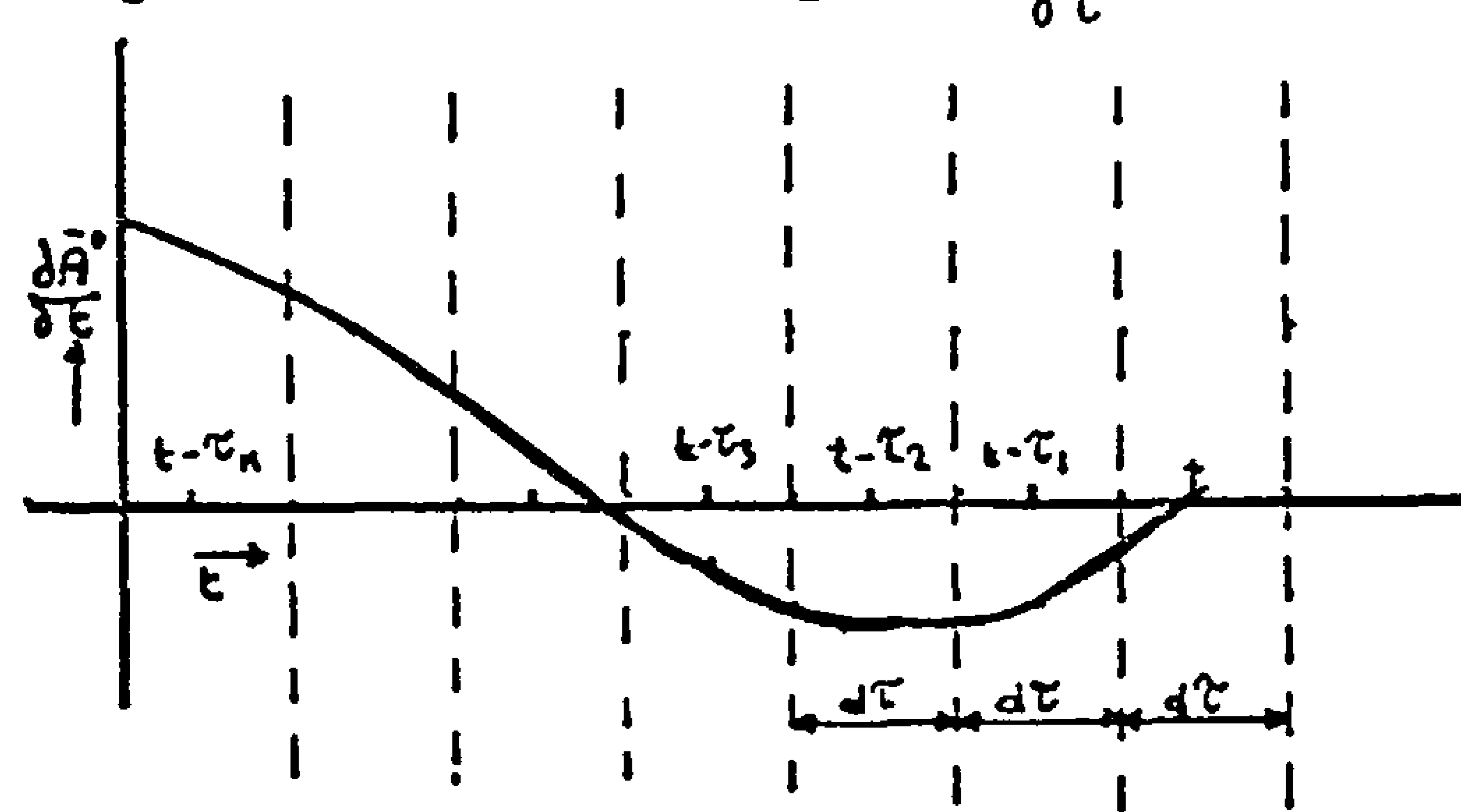


fig. b

Het verloop van dA^{o} kan benaderd worden door discontinue veranderingen op tijdstippen $t-\tau_n$. Voor $d\overline{A}^{o}$ geldt dan op zo'n tijdstip

$$d\overline{A}^{o} = \frac{\partial \overline{A}^{o}}{\partial t} d\tau \qquad \text{zie fig. c} \tag{34}$$

waarbij $d\tau$ de tijd is die verloopt tussen twee discontinue veranderingen van $d\overline{A}^{o}$

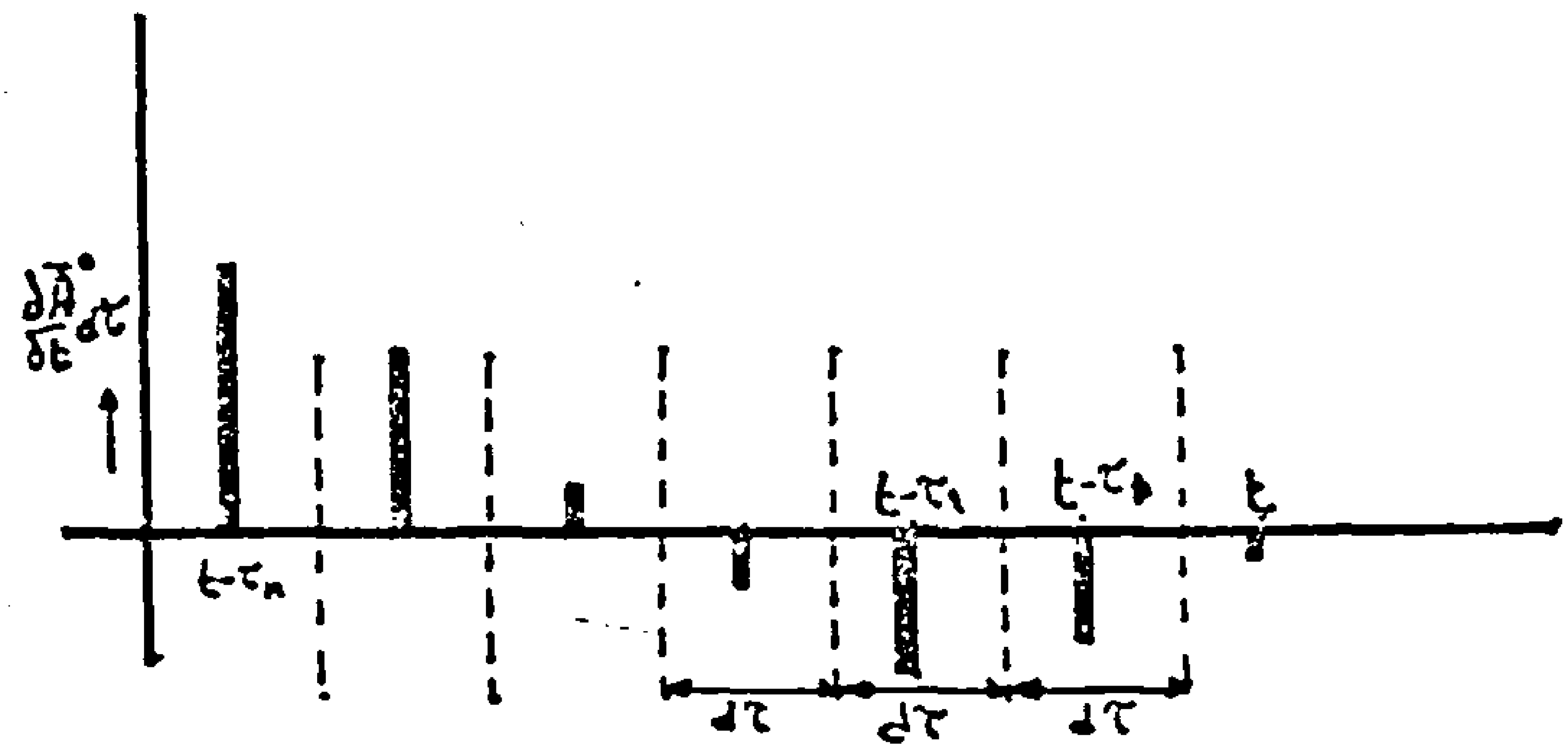


fig. c

De op het moment $t-\tau_n$ plaatsvindende verandering in de vectorpotentiaal ter grootte $d\bar{A}^o = \frac{\partial f(x_o, y_o, z_o, t-\tau_n)}{\partial t} d\tau$ (35)

levert op het moment t , ter plaatse van $z = \pm \delta/2$

$$d\bar{A}^w = - \frac{\partial f(x_o, y_o, |z| + w\tau_n, t-\tau_n)}{\partial t} d\tau \quad z = -\delta/2 \quad \text{of} \tag{36a}$$

$$d\bar{A}^w = - \frac{\partial f(x_o, y_o, -|z| - w\tau_n, t-\tau_n)}{\partial t} d\tau \quad z = +\delta/2 \tag{36b}$$

Het momentane veld $\bar{A}^w$ ten tijde t is een sommatie van alle voorgaande uitstervende $d\bar{A}^w$'s.

$$\bar{A}^w = \sum_{n=0}^{\infty} - \frac{\partial f(x_o, y_o, |z| + w\tau_n, t-\tau_n)}{\partial t} d\tau \quad z = -\delta/2 \quad \text{of} \tag{37a}$$

$$\bar{A}^w = \sum_{n=0}^{\infty} - \frac{\partial f(x_o, y_o, -|z| - w\tau_n, t-\tau_n)}{\partial t} d\tau \quad z = +\delta/2 \tag{37b}$$

Als nu de limiet wordt genomen voor $d\tau \to 0$ dan wordt het verloop van $d\bar{A}^o$ continu en gaat (37) over in de volgende integraal.

$$\bar{A}^w = - \int_0^{\infty} \frac{\partial f(x_o, y_o, \pm|z| \pm w\tau, t-\tau)}{\partial t} d\tau \quad z = \mp \delta/2 \tag{38}$$

met $\bar{A}^o = f(x_o, y_o, z_o, t)$

$$\bar{J}^w = -\sigma \left(\pm w \frac{\partial \bar{A}^w}{\partial z} \right)_{z = \pm\delta/2} = \mp \sigma w \frac{\partial}{\partial z} \int_0^{\infty} \frac{\partial f(x_o, y_o, \pm|z| \pm w\tau, t-\tau)}{\partial t} d\tau \quad z = \mp\delta/2 \tag{39}$$

Met behulp van formule (39) kan de wervelstroomverdeling $\bar{J}^w$ in een oneindig uitgebreide plaat worden berekend als de magnetische vectorpotentiaal van het inducerend veld, $\bar{A}^o$, bekend is.

N.B. Het maakt voor de berekeningen geen verschil of men van de relaties 36 t/m 39 de relaties voor $z = +\delta/2$ of $z = -\delta/2$ gebruikt.

Appendix A3: Calculation of the amplitudes of the diffraction pattern for even periodic surfaces.

The computer program listed on the pages A3.2-11 has been written to calculate the coefficients $|B_m|^2$ of the formula:

$$I(x) = \sum_{m=-84}^{84} |B_m|^2 \delta\left(X - \frac{m\lambda F}{D}\right) \quad (9.33)$$

for surface profiles expressed by

$$y(x) = \left(R_q \sum_{n=1}^{7} a_n \cos(2\pi n x/D)\right) \Big/ \sqrt{\frac{1}{2} \sum_{n=1}^{7} a_n^2} \quad (9.27)$$

The coefficients a_n are entered at line 1020

The coefficients B_m are produced at the lines 5170-5240, while line 5250 calculates $\sum_{m=-7}^{7} |B_m|^2$ to check whether or not enough diffraction orders are taken into account.

The "$\frac{1}{4}$" symbol stands for raising to the power of.

```
1000 MOVE 100,200
1010 CLG
1020 INPUT "a1,a2,a3,a4,a5,a6,a7";a1,a2,a3,a4,a5,a6,a7
1030 q=SQR(a1*a1+a2*a2+a3*a3+a4*a4+a5*a5+a6*a6+a7*a7)
1040 INPUT "Rq/lambda";Rq
1050 b0=(a1*a1+a2*a2+a3*a3+a4*a4+a5*a5+a6*a6+a7*a7)/2
1060 b1=a1*a2+a2*a3+a3*a4+a4*a5+a5*a6+a6*a7
1070 b2=a1*a3+a2*a4+a3*a5+a4*a6+a5*a7+a1*a1/2
1080 b3=a1*a2+a1*a4+a2*a5+a3*a6+a4*a7
1090 b4=a1*a3+a1*a5+a2*a6+a3*a7+a2*a2/2
1100 b5=a1*a4+a1*a6+a2*a3+a2*a7
1110 b6=a1*a5+a1*a7+a2*a4+a3*a3/2
1120 b7=a1*a6+a2*a5+a3*a4
1130 b8=a1*a7+a2*a6+a3*a5+a4*a4/2
1140 b9=a2*a7+a3*a6+a4*a5
1150 b10=a3*a7+a4*a6+a5*a5/2
1160 b11=a4*a7+a5*a6
1170 b12=a5*a7+a6*a6/2
1180 b13=a6*a7
1190 b14=a7*a7/2
1200 c0=(a1*b1+a2*b2+a3*b3+a4*b4+a5*b5+a6*b6+a7*b7)/2
1210 c1=(a1*(2*b0+b2)+a2*(b1+b3)+a3*(b2+b4)+a4*(b3+b5)+a5*(b4+b6)+a6*(b5+b7)+a7*(b6+b8))/2
1220 c2=(a1*(b1+b3)+a2*(2*b0+b4)+a3*(b1+b5)+a4*(b2+b6)+a5*(b3+b7)+a6*(b4+b8)+a7*(b5+b9))/2
1230 c3=(a1*(b2+b4)+a2*(b1+b5)+a3*(2*b0+b6)+a4*(b1+b7)+a5*(b2+b8)+a6*(b3+b9)+a7*(b4+b10))/2
1240 c4=(a1*(b3+b5)+a2*(b2+b6)+a3*(b1+b7)+a4*(2*b0+b8)+a5*(b1+b9)+a6*(b2+b10)+a7*(b3+b11))/2
1250 c5=(a1*(b4+b6)+a2*(b3+b7)+a3*(b2+b8)+a4*(b1+b9)+a5*(2*b0+b10)+a6*(b1+b11)+a7*(b2+b12))/2
1260 c6=(a1*(b5+b7)+a2*(b4+b8)+a3*(b3+b9)+a4*(b2+b10)+a5*(b1+b11)+a6*(2*b0+b12)+a7*(b1+b13))/2
1270 c7=(a1*(b6+b8)+a2*(b5+b9)+a3*(b4+b10)+a4*(b3+b11)+a5*(b2+b12)+a6*(b1+b13)+a7*(2*b0+b14))/2
1280 c8=(a1*(b7+b9)+a2*(b6+b10)+a3*(b5+b11)+a4*(b4+b12)+a5*(b3+b13)+a6*(b2+b14)+a7*b1)/2
1290 c9=(a1*(b8+b10)+a2*(b7+b11)+a3*(b6+b12)+a4*(b5+b13)+a5*(b4+b14)+a6*b3+a7*b2)/2
1300 c10=(a1*(b9+b11)+a2*(b8+b12)+a3*(b7+b13)+a4*(b6+b14)+a5*b5+a6*b4+a7*b3)/2
1310 c11=(a1*(b10+b12)+a2*(b9+b13)+a3*(b8+b14)+a4*b7+a5*b6+a6*b5+a7*b4)/2
1320 c12=(a1*(b11+b13)+a2*(b10+b14)+a3*b9+a4*b8+a5*b7+a6*b6+a7*b5)/2
1330 c13=(a1*(b12+b14)+a2*b11+a3*b10+a4*b9+a5*b8+a6*b7+a7*b6)/2
```

```
1340 c14=(a1*b13+a2*b12+a3*b11+a4*b10+a5*b9+a6*b8+a7*b7)/2
1350 c15=(a1*b14+a2*b13+a3*b12+a4*b11+a5*b10+a6*b9+a7*b8)/2
1360 c16=(a2*b14+a3*b13+a4*b12+a5*b11+a6*b10+a7*b9)/2
1370 c17=(a3*b14+a4*b13+a5*b12+a6*b11+a7*b10)/2
1380 c18=(a4*b14+a5*b13+a6*b12+a7*b11)/2
1390 c19=(a5*b14+a6*b13+a7*b12)/2
1400 c20=(a6*b14+a7*b13)/2
1410 c21=a7*b14/2
1420 d0=b0*b0+(b1*b1+b2*b2+b3*b3+b4*b4+b5*b5+b6*b6+b7*b7+b8*b8+b9*b9+b10*b10+b11*b11+b12*b12+b13*b13+b14*b14)/2
1430 d1=2*b0*b1+b2*(b1+b3)+b4*(b3+b5)+b6*(b5+b7)+b8*(b7+b9)+b10*(b9+b11)+b12*(b11+b13)+b13*b14
1440 d2=2*b0*b2+b1*b1/2+b3*(b1+b5)+b4*(b2+b6)+b7*(b5+b9)+b8*(b6+b10)+b11*(b9+b13)+b12*(b10+b14)
1450 d3=2*b0*b3+b1*b2+b4*(b1+b7)+b5*(b2+b8)+b6*(b3+b9)+b10*(b7+b13)+b11*(b8+b14)+b9*b12
1460 d4=2*b0*b4+b2*b2/2+b1*b3+b5*(b1+b9)+b6*(b2+b10)+b7*(b3+b11)+b8*(b4+b12)+b9*b13+b10*b14
1470 d5=2*b0*b5+b1*b4+b2*b3+b6*(b1+b11)+b7*(b2+b12)+b8*(b3+b13)+b9*(b4+b14)+b5*b10
1480 d6=2*b0*b6+b3*b3/2+b7*(b1+b13)+b8*(b2+b14)+b3*b9+b4*(b10+b2)+b5*(b11+b1)+b6*b12
1490 d7=2*b0*b7+b1*(b6+b8)+b2*(b5+b9)+b3*(b4+b10)+b4*b11+b5*b12+b6*b13+b7*b14
1500 d8=2*b0*b8+b4*b4/2+b1*(b7+b9)+b2*(b6+b10)+b3*(b5+b11)+b4*b12+b5*b13+b6*b14
1510 d9=2*b0*b9+b1*(b8+b10)+b2*(b7+b11)+b3*(b6+b12)+b4*(b5+b13)+b5*b14
1520 d10=2*b0*b10+b5*b5/2+b1*(b9+b11)+b2*(b8+b12)+b3*(b7+b13)+b4*(b6+b14)
1530 d11=2*b0*b11+b1*(b10+b12)+b2*(b9+b13)+b3*(b8+b14)+b4*b7+b5*b6
1540 d12=2*b0*b12+b6*b6/2+b1*(b11+b13)+b2*(b10+b14)+b3*b9+b4*b8+b5*b7
1550 d13=2*b0*b13+b1*(b12+b14)+b2*b11+b3*b10+b4*b9+b5*b8+b6*b7
1560 d14=2*b0*b14+b7*b7/2+b1*b13+b2*b12+b3*b11+b4*b10+b5*b9+b6*b8
1570 d15=b1*b14+b2*b13+b3*b12+b4*b11+b5*b10+b6*b9+b7*b8
1580 d16=b2*b14+b3*b13+b4*b12+b5*b11+b6*b10+b7*b9+b8*b8/2
1590 d17=b3*b14+b4*b13+b5*b12+b6*b11+b7*b10+b8*b9
1600 d18=b4*b14+b5*b13+b6*b12+b7*b11+b8*b10+b9*b9/2
1610 d19=b5*b14+b6*b13+b7*b12+b8*b11+b9*b10
1620 d20=b6*b14+b7*b13+b8*b12+b9*b11+b10*b10/2
1630 d21=b7*b14+b8*b13+b9*b12+b10*b11
1640 d22=b8*b14+b9*b13+b10*b12+b11*b11/2
1650 d23=b9*b14+b10*b13+b11*b12
1660 d24=b10*b14+b11*b13+b12*b12/2
```

```
1670 d25=b11*b14+b12*b13
1680 d26=b12*b14+b13*b13/2
1690 d27=b13*b14
1700 d28=b14*b14/2
1710 e0=(a1*d1+a2*d2+a3*d3+a4*d4+a5*d5+a6*d6+a7*d7)/2
1720 e1=a1*d0+(a1*d2+a2*(d1+d3)+a3*(d2+d4)+a4*(d3+d5)+a5*(d4+d6)+a6*(d5+d7)+a7*(d6+d8))/2
1730 e2=a2*d0+(a1*(d1+d3)+a2*d4+a3*(d1+d5)+a4*(d2+d6)+a5*(d3+d7)+a6*(d4+d8)+a7*(d5+d9))/2
1740 e3=a3*d0+(a1*(d2+d4)+a2*(d1+d5)+a3*d6+a4*(d1+d7)+a5*(d2+d8)+a6*(d3+d9)+a7*(d4+d10))/2
1750 e4=a4*d0+(a1*(d3+d5)+a2*(d2+d6)+a3*(d1+d7)+a4*d8+a5*(d1+d9)+a6*(d2+d10)+a7*(d3+d11))/2
1760 e5=a5*d0+(a1*(d4+d6)+a2*(d3+d7)+a3*(d2+d8)+a4*(d1+d9)+a5*d10+a6*(d1+d11)+a7*(d2+d12))/2
1770 e6=a6*d0+(a1*(d5+d7)+a2*(d4+d8)+a3*(d3+d9)+a4*(d2+d10)+a5*(d1+d11)+a6*d12+a7*(d1+d13))/2
1780 e7=a7*d0+(a1*(d6+d8)+a2*(d5+d9)+a3*(d4+d10)+a4*(d3+d11)+a5*(d2+d12)+a6*(d1+d13)+a7*d14)/2
1790 e8=(a1*(d7+d9)+a2*(d6+d10)+a3*(d5+d11)+a4*(d4+d12)+a5*(d3+d13)+a6*(d2+d14)+a7*(d1+d15))/2
1800 e9=(a1*(d8+d10)+a2*(d7+d11)+a3*(d6+d12)+a4*(d5+d13)+a5*(d4+d14)+a6*(d3+d15)+a7*(d2+d16))/2
1810 e10=(a1*(d9+d11)+a2*(d8+d12)+a3*(d7+d13)+a4*(d6+d14)+a5*(d5+d15)+a6*(d4+d16)+a7*(d3+d17))/2
1820 e11=(a1*(d10+d12)+a2*(d9+d13)+a3*(d8+d14)+a4*(d7+d15)+a5*(d6+d16)+a6*(d5+d17)+a7*(d4+d18))/2
1830 e12=(a1*(d11+d13)+a2*(d10+d14)+a3*(d9+d15)+a4*(d8+d16)+a5*(d7+d17)+a6*(d6+d18)+a7*(d5+d19))/2
1840 e13=(a1*(d12+d14)+a2*(d11+d15)+a3*(d10+d16)+a4*(d9+d17)+a5*(d8+d18)+a6*(d7+d19)+a7*(d6+d20))/2
1850 e14=(a1*(d13+d15)+a2*(d12+d16)+a3*(d11+d17)+a4*(d10+d18)+a5*(d9+d19)+a6*(d8+d20)+a7*(d7+d21))/2
1860 e15=(a1*(d14+d16)+a2*(d13+d17)+a3*(d12+d18)+a4*(d11+d19)+a5*(d10+d20)+a6*(d9+d21)+a7*(d8+d22))/2
1870 e16=(a1*(d15+d17)+a2*(d14+d18)+a3*(d13+d19)+a4*(d12+d20)+a5*(d11+d21)+a6*(d10+d22)+a7*(d9+d23))/2
1880 e17=(a1*(d16+d18)+a2*(d15+d19)+a3*(d14+d20)+a4*(d13+d21)+a5*(d12+d22)+a6*(d11+d23)+a7*(d10+d24))/2
1890 e18=(a1*(d17+d19)+a2*(d16+d20)+a3*(d15+d21)+a4*(d14+d22)+a5*(d13+d23)+a6*(d12+d24)+a7*(d11+d25))/2
1900 e19=(a1*(d18+d20)+a2*(d17+d21)+a3*(d16+d22)+a4*(d15+d23)+a5*(d14+d24)+a6*(d13+d25)+a7*(d12+d26))/2
1910 e20=(a1*(d19+d21)+a2*(d18+d22)+a3*(d17+d23)+a4*(d16+d24)+a5*(d15+d25)+a6*(d14+d26)+a7*(d13+d27))/2
1920 e21=(a1*(d20+d22)+a2*(d19+d23)+a3*(d18+d24)+a4*(d17+d25)+a5*(d16+d26)+a6*(d15+d27)+a7*(d14+d28))/2
1930 e22=(a1*(d21+d23)+a2*(d20+d24)+a3*(d19+d25)+a4*(d18+d26)+a5*(d17+d27)+a6*(d16+d28)+a7*d15)/2
1940 e23=(a1*(d22+d24)+a2*(d21+d25)+a3*(d20+d26)+a4*(d19+d27)+a5*(d18+d28)+a6*d17+a7*d16)/2
1950 e24=(a1*(d23+d25)+a2*(d22+d26)+a3*(d21+d27)+a4*(d20+d28)+a5*d19+a6*d18+a7*d17)/2
1960 e25=(a1*(d24+d26)+a2*(d23+d27)+a3*(d22+d28)+a4*d21+a5*d20+a6*d19+a7*d18)/2
1970 e26=(a1*(d25+d27)+a2*(d24+d28)+a3*d23+a4*d22+a5*d21+a6*d20+a7*d19)/2
1980 e27=(a1*(d26+d28)+a2*d25+a3*d24+a4*d23+a5*d22+a6*d21+a7*d20)/2
1990 e28=(a1*d27+a2*d26+a3*d25+a4*d24+a5*d23+a6*d22+a7*d21)/2
```

```
2000 e29=(a1*d28+a2*d27+a3*d26+a4*d25+a5*d24+a6*d23+a7*d22)/2
2010 e30=(a2*d28+a3*d27+a4*d26+a5*d25+a6*d24+a7*d23)/2
2020 e31=(a3*d28+a4*d27+a5*d26+a6*d25+a7*d24)/2
2030 e32=(a4*d28+a5*d27+a6*d26+a7*d25)/2
2040 e33=(a5*d28+a6*d27+a7*d26)/2
2050 e34=(a6*d28+a7*d27)/2
2060 e35=(a7*d28)/2
2070 f0=(a1*e1+a2*e2+a3*e3+a4*e4+a5*e5+a6*e6+a7*e7)/2
2080 f1=(a1*(2*e0+e2)+a2*(e1+e3)+a3*(e2+e4)+a4*(e3+e5)+a5*(e4+e6)+a6*(e5+e7)+a7*(e6+e8))/2
2090 f2=(a1*(e1+e3)+a2*(2*e0+e4)+a3*(e1+e5)+a4*(e2+e6)+a5*(e3+e7)+a6*(e4+e8)+a7*(e5+e9))/2
2100 f3=(a1*(e2+e4)+a2*(e1+e5)+a3*(2*e0+e6)+a4*(e1+e7)+a5*(e2+e8)+a6*(e3+e9)+a7*(e4+e10))/2
2110 f4=(a1*(e3+e5)+a2*(e2+e6)+a3*(e1+e7)+a4*(2*e0+e8)+a5*(e1+e9)+a6*(e2+e10)+a7*(e3+e11))/2
2120 f5=(a1*(e4+e6)+a2*(e3+e7)+a3*(e2+e8)+a4*(e1+e9)+a5*(2*e0+e10)+a6*(e1+e11)+a7*(e2+e12))/2
2130 f6=(a1*(e5+e7)+a2*(e4+e8)+a3*(e3+e9)+a4*(e2+e10)+a5*(e1+e11)+a6*(2*e0+e12)+a7*(e1+e13))/2
2140 f7=(a1*(e6+e8)+a2*(e5+e9)+a3*(e4+e10)+a4*(e3+e11)+a5*(e2+e12)+a6*(e1+e13)+a7*(2*e0+e14))/2
2150 f8=(a1*(e7+e9)+a2*(e6+e10)+a3*(e5+e11)+a4*(e4+e12)+a5*(e3+e13)+a6*(e2+e14)+a7*(e1+e15))/2
2160 f9=(a1*(e8+e10)+a2*(e7+e11)+a3*(e6+e12)+a4*(e5+e13)+a5*(e4+e14)+a6*(e3+e15)+a7*(e2+e16))/2
2170 f10=(a1*(e9+e11)+a2*(e8+e12)+a3*(e7+e13)+a4*(e6+e14)+a5*(e5+e15)+a6*(e4+e16)+a7*(e3+e17))/2
2180 f11=(a1*(e10+e12)+a2*(e9+e13)+a3*(e8+e14)+a4*(e7+e15)+a5*(e6+e16)+a6*(e5+e17)+a7*(e4+e18))/2
2190 f12=(a1*(e11+e13)+a2*(e10+e14)+a3*(e9+e15)+a4*(e8+e16)+a5*(e7+e17)+a6*(e6+e18)+a7*(e5+e19))/2
2200 f13=(a1*(e12+e14)+a2*(e11+e15)+a3*(e10+e16)+a4*(e9+e17)+a5*(e8+e18)+a6*(e7+e19)+a7*(e6+e20))/2
2210 f14=(a1*(e13+e15)+a2*(e12+e16)+a3*(e11+e17)+a4*(e10+e18)+a5*(e9+e19)+a6*(e8+e20)+a7*(e7+e21))/2
2220 f15=(a1*(e14+e16)+a2*(e13+e17)+a3*(e12+e18)+a4*(e11+e19)+a5*(e10+e20)+a6*(e9+e21)+a7*(e8+e22))/2
2230 f16=(a1*(e15+e17)+a2*(e14+e18)+a3*(e13+e19)+a4*(e12+e20)+a5*(e11+e21)+a6*(e10+e22)+a7*(e9+e23))/2
2240 f17=(a1*(e16+e18)+a2*(e15+e19)+a3*(e14+e20)+a4*(e13+e21)+a5*(e12+e22)+a6*(e11+e23)+a7*(e10+e24))/2
2250 f18=(a1*(e17+e19)+a2*(e16+e20)+a3*(e15+e21)+a4*(e14+e22)+a5*(e13+e23)+a6*(e12+e24)+a7*(e11+e25))/2
2260 f19=(a1*(e18+e20)+a2*(e17+e21)+a3*(e16+e22)+a4*(e15+e23)+a5*(e14+e24)+a6*(e13+e25)+a7*(e12+e26))/2
2270 f20=(a1*(e19+e21)+a2*(e18+e22)+a3*(e17+e23)+a4*(e16+e24)+a5*(e15+e25)+a6*(e14+e26)+a7*(e13+e27))/2
2280 f21=(a1*(e20+e22)+a2*(e19+e23)+a3*(e18+e24)+a4*(e17+e25)+a5*(e16+e26)+a6*(e15+e27)+a7*(e14+e28))/2
2290 f22=(a1*(e21+e23)+a2*(e20+e24)+a3*(e19+e25)+a4*(e18+e26)+a5*(e17+e27)+a6*(e16+e28)+a7*e15)/2
2300 f23=(a1*(e22+e24)+a2*(e21+e25)+a3*(e20+e26)+a4*(e19+e27)+a5*(e18+e28)+a6*e17+a7*e16)/2
2310 f24=(a1*(e23+e25)+a2*(e22+e26)+a3*(e21+e27)+a4*(e20+e28)+a5*e19+a6*e18+a7*e17)/2
2320 f25=(a1*(e24+e26)+a2*(e23+e27)+a3*(e22+e28)+a4*e21+a5*e20+a6*e19+a7*e18)/2
```

```
2330 f26=(a1*(e25+e27)+a2*(e24+e28)+a3*e23+a4*e22+a5*e21+a6*e20+a7*e19)/2
2340 f27=(a1*(e26+e28)+a2*e25+a3*e24+a4*e23+a5*e22+a6*e21+a7*e20)/2
2350 f28=(a1*e27+a2*e26+a3*e25+a4*e24+a5*e23+a6*e22+a7*e21)/2
2360 g0=(a1*f1+a2*f2+a3*f3+a4*f4+a5*f5+a6*f6+a7*f7)/2
2370 g1=(a1*(2*f0+f2)+a2*(f1+f3)+a3*(f2+f4)+a4*(f3+f5)+a5*(f4+f6)+a6*(f5+f7)+a7*(f6+f8))/2
2380 g2=(a1*(f1+f3)+a2*(2*f0+f4)+a3*(f1+f5)+a4*(f2+f6)+a5*(f3+f7)+a6*(f4+f8)+a7*(f5+f9))/2
2390 g3=(a1*(f2+f4)+a2*(f1+f5)+a3*(2*f0+f6)+a4*(f1+f7)+a5*(f2+f8)+a6*(f3+f9)+a7*(f4+f10))/2
2400 g4=(a1*(f3+f5)+a2*(f2+f6)+a3*(f1+f7)+a4*(2*f0+f8)+a5*(f1+f9)+a6*(f2+f10)+a7*(f3+f11))/2
2410 g5=(a1*(f4+f6)+a2*(f3+f7)+a3*(f2+f8)+a4*(f1+f9)+a5*(2*f0+f10)+a6*(f1+f11)+a7*(f2+f12))/2
2420 g6=(a1*(f5+f7)+a2*(f4+f8)+a3*(f3+f9)+a4*(f2+f10)+a5*(f1+f11)+a6*(2*f0+f12)+a7*(f1+f13))/2
2430 g7=(a1*(f6+f8)+a2*(f5+f9)+a3*(f4+f10)+a4*(f3+f11)+a5*(f2+f12)+a6*(f1+f13)+a7*(2*f0+f14))/2
2440 g8=(a1*(f7+f9)+a2*(f6+f10)+a3*(f5+f11)+a4*(f4+f12)+a5*(f3+f13)+a6*(f2+f14)+a7*(f1+f15))/2
2450 g9=(a1*(f8+f10)+a2*(f7+f11)+a3*(f6+f12)+a4*(f5+f13)+a5*(f4+f14)+a6*(f3+f15)+a7*(f2+f16))/2
2460 g10=(a1*(f9+f11)+a2*(f8+f12)+a3*(f7+f13)+a4*(f6+f14)+a5*(f5+f15)+a6*(f4+f16)+a7*(f3+f17))/2
2470 g11=(a1*(f10+f12)+a2*(f9+f13)+a3*(f8+f14)+a4*(f7+f15)+a5*(f6+f16)+a6*(f5+f17)+a7*(f4+f18))/2
2480 g12=(a1*(f11+f13)+a2*(f10+f14)+a3*(f9+f15)+a4*(f8+f16)+a5*(f7+f17)+a6*(f6+f18)+a7*(f5+f19))/2
2490 g13=(a1*(f12+f14)+a2*(f11+f15)+a3*(f10+f16)+a4*(f9+f17)+a5*(f8+f18)+a6*(f7+f19)+a7*(f6+f20))/2
2500 g14=(a1*(f13+f15)+a2*(f12+f16)+a3*(f11+f17)+a4*(f10+f18)+a5*(f9+f19)+a6*(f8+f20)+a7*(f7+f21))/2
2510 g15=(a1*(f14+f16)+a2*(f13+f17)+a3*(f12+f18)+a4*(f11+f19)+a5*(f10+f20)+a6*(f9+f21)+a7*(f8+f22))/2
2520 g16=(a1*(f15+f17)+a2*(f14+f18)+a3*(f13+f19)+a4*(f12+f20)+a5*(f11+f21)+a6*(f10+f22)+a7*(f9+f23))/2
2530 g17=(a1*(f16+f18)+a2*(f15+f19)+a3*(f14+f20)+a4*(f13+f21)+a5*(f12+f22)+a6*(f11+f23)+a7*(f10+f24))/2
2540 g18=(a1*(f17+f19)+a2*(f16+f20)+a3*(f15+f21)+a4*(f14+f22)+a5*(f13+f23)+a6*(f12+f24)+a7*(f11+f25))/2
2550 g19=(a1*(f18+f20)+a2*(f17+f21)+a3*(f16+f22)+a4*(f15+f23)+a5*(f14+f24)+a6*(f13+f25)+a7*(f12+f26))/2
2560 g20=(a1*(f19+f21)+a2*(f18+f22)+a3*(f17+f23)+a4*(f16+f24)+a5*(f15+f25)+a6*(f14+f26)+a7*(f13+f27))/2
2570 g21=(a1*(f20+f22)+a2*(f19+f23)+a3*(f18+f24)+a4*(f17+f25)+a5*(f16+f26)+a6*(f15+f27)+a7*(f14+f28))/2
2580 h0=(a1*g1+a2*g2+a3*g3+a4*g4+a5*g5+a6*g6+a7*g7)/2
2590 h1=(a1*(2*g0+g2)+a2*(g1+g3)+a3*(g2+g4)+a4*(g3+g5)+a5*(g4+g6)+a6*(g5+g7)+a7*(g6+g8))/2
2600 h2=(a1*(g1+g3)+a2*(2*g0+g4)+a3*(g1+g5)+a4*(g2+g6)+a5*(g3+g7)+a6*(g4+g8)+a7*(g5+g9))/2
2610 h3=(a1*(g2+g4)+a2*(g1+g5)+a3*(2*g0+g6)+a4*(g1+g7)+a5*(g2+g8)+a6*(g3+g9)+a7*(g4+g10))/2
2620 h4=(a1*(g3+g5)+a2*(g2+g6)+a3*(g1+g7)+a4*(2*g0+g8)+a5*(g1+g9)+a6*(g2+g10)+a7*(g3+g11))/2
2630 h5=(a1*(g4+g6)+a2*(g3+g7)+a3*(g2+g8)+a4*(g1+g9)+a5*(2*g0+g10)+a6*(g1+g11)+a7*(g2+g12))/2
2640 h6=(a1*(g5+g7)+a2*(g4+g8)+a3*(g3+g9)+a4*(g2+g10)+a5*(g1+g11)+a6*(2*g0+g12)+a7*(g1+g13))/2
2650 h7=(a1*(g6+g8)+a2*(g5+g9)+a3*(g4+g10)+a4*(g3+g11)+a5*(g2+g12)+a6*(g1+g13)+a7*(2*g0+g14))/2
```

```
2660 h8=(a1*(g7+g9)+a2*(g6+g10)+a3*(g5+g11)+a4*(g4+g12)+a5*(g3+g13)+a6*(g2+g14)+a7*(g1+g15))/2
2670 h9=(a1*(g8+g10)+a2*(g7+g11)+a3*(g6+g12)+a4*(g5+g13)+a5*(g4+g14)+a6*(g3+g15)+a7*(g2+g16))/2
2680 h10=(a1*(g9+g11)+a2*(g8+g12)+a3*(g7+g13)+a4*(g6+g14)+a5*(g5+g15)+a6*(g4+g16)+a7*(g3+g17))/2
2690 h11=(a1*(g10+g12)+a2*(g9+g13)+a3*(g8+g14)+a4*(g7+g15)+a5*(g6+g16)+a6*(g5+g17)+a7*(g4+g18))/2
2700 h12=(a1*(g11+g13)+a2*(g10+g14)+a3*(g9+g15)+a4*(g8+g16)+a5*(g7+g17)+a6*(g6+g18)+a7*(g5+g19))/2
2710 h13=(a1*(g12+g14)+a2*(g11+g15)+a3*(g10+g16)+a4*(g9+g17)+a5*(g8+g18)+a6*(g7+g19)+a7*(g6+g20))/2
2720 h14=(a1*(g13+g15)+a2*(g12+g16)+a3*(g11+g17)+a4*(g10+g18)+a5*(g9+g19)+a6*(g8+g20)+a7*(g7+g21))/2
2730 i0=(a1*h1+a2*h2+a3*h3+a4*h4+a5*h5+a6*h6+a7*h7)/2
2740 i1=(a1*(2*h0+h2)+a2*(h1+h3)+a3*(h2+h4)+a4*(h3+h5)+a5*(h4+h6)+a6*(h5+h7)+a7*(h6+h8))/2
2750 i2=(a1*(h1+h3)+a2*(2*h0+h4)+a3*(h1+h5)+a4*(h2+h6)+a5*(h3+h7)+a6*(h4+h8)+a7*(h5+h9))/2
2760 i3=(a1*(h2+h4)+a2*(h1+h5)+a3*(2*h0+h6)+a4*(h1+h7)+a5*(h2+h8)+a6*(h3+h9)+a7*(h4+h10))/2
2770 i4=(a1*(h3+h5)+a2*(h2+h6)+a3*(h1+h7)+a4*(2*h0+h8)+a5*(h1+h9)+a6*(h2+h10)+a7*(h3+h11))/2
2780 i5=(a1*(h4+h6)+a2*(h3+h7)+a3*(h2+h8)+a4*(h1+h9)+a5*(2*h0+h10)+a6*(h1+h11)+a7*(h2+h12))/2
2790 i6=(a1*(h5+h7)+a2*(h4+h8)+a3*(h3+h9)+a4*(h2+h10)+a5*(h1+h11)+a6*(2*h0+h12)+a7*(h1+h13))/2
2800 i7=(a1*(h6+h8)+a2*(h5+h9)+a3*(h4+h10)+a4*(h3+h11)+a5*(h2+h12)+a6*(h1+h13)+a7*(2*h0+h14))/2
2805 |GOTO,1,2810
2810 j0a=e0*e0+(e1^2+e2^2+e3^2+e4^2+e5^2+e6^2+e7^2+e8^2+e9^2+e10^2+e11^2+e12^2+e13^2+e14^2+e15^2)/2
2820 j0=j0a+(e16^2+e17^2+e18^2+e19^2+e20^2+e21^2+e22^2+e23^2+e24^2+e25^2+e26^2+e27^2+e28^2)/2
2830 j1a=2*e0*e2+e2*(e1+e3)+e4*(e3+e5)+e6*(e5+e7)+e8*(e7+e9)+e10*(e9+e11)+e12*(e11+e13)+e14*(e13+15)+e16*(e15+e17)+e18*(e17+e19)
2840 j1=j1a+e20*(e19+e21)+e22*(e21+e23)+e24*(e23+e25)+e26*(e25+e27)+e28*e27+e28*e29/2
2850 j2a=2*e0*e2+e3*(e1+e5)+e4*(e2+e6)+e7*(e5+e9)+e8*(e6+e10)+e11*(e9+e13)+e12*(e10+e14)+e15*(e13+e17)+e16*(e14+e18)+e19*(e17+e21)+e
20*(e18+e22)
2860 j2=j2a+e23*(e21+e25)+e24*(e22+e26)+e25*e27+e26*e28+(e29*e27+e30*e28+e1^2)/2
2870 j3a=2*e0*e3+e1*e2+e4*(e1+e7)+e5*(e2+e8)+e6*(e3+e9)+e10*(e7+e13)+e11*(e8+e14)+e12*(e9+e15)+e16*(e13+e19)+e17*(e14+e20)+e18*(e15+
e21)+e22*(e19+e25)+e23*(e20+e26)+e24*(e21+e27)
2880 j3=j3a+e28*e25+(e26*e29+e27*e30+e28*e31)/2
2890 j4a=2*e0*e4+e1*e3+e5*(e1+e9)+e6*(e2+e10)+e7*(e3+e11)+e8*(e4+e12)+e13*(e9+e17)+e14*(e10+e18)+e15*(e11+e19)+e16*(e12+e20)+e21*(e1
7+e25)+e22*(e18+e26)+e23*(e19+e27)
2900 j4=j4a+e24*(e20+e28)+(e25*e29+e26*e30+e27*e31+e28*e32+e2^2)/2
2910 j5a=2*e0*e5+e1*e4+e2*e3+e6*(e1+e11)+e7*(e2+e12)+e8*(e3+e13)+e9*(e4+e25)+e10*(e5+e15)+e16*(e11+e21)+e17*(e12+e22)+e18*(e13+e23)+
e19*(e14+e24)+e20*(e15+e25)
2920 j5=j5a+(e24*e29+e25*e30+e26*e31+e27*e32+e28*e33)/2+e26*e21+e27*e22+e28*e23
```

```
2930 j6a=2*e0*e6+e1*e5+e2*e4+e7*(e1+13)+e8*(e2+e14)+e9*(e3+e15)+e10*(e4+e16)+e11*(e5+e17)+e12*(e6+e18)+e19*(e13+e25)+e20*(e14+e26)+e
21*(e15+e27)+e22*(e16+e28)+e23*e17+e24*e18
2940 j6=j6a+(e23*e29+e24*e30+e25*e31+e26*e32+e27*e33+e28*e34+e3↑2)/2
2950 j7a=2*e0*e7+e1*e6+e2*e5+e3*e4+e8*(e1+e15)+e9*(e2+e16)+e10*(e3+e17)+e11*(e4+e18)+e12*(e5+e19)+e13*(e6+e20)+e14*(e7+e21)+e22*e15+
e23*e16
2960 j7=j7a+e24*e17+e25*e18+e26*e19+e27*e20+e28*e21+(e22*e29+e23*e30+e24*e31+e25*e32+e26*e33+e27*e34+e28*e35)/2
2970 j8a=2*e0*e8+e1*(e7+e9)+e2*(e6+e10)+e3*(e5+e11)+e4*e12+e13*(e5+e21)+e14*(e6+e22)+e15*(e7+e23)+e16*(e8+e24)+e17*(e9+e25)+e18*(e10
+e26)
2980 j8=j8a+e19*(e11+e27)+e20*(e12+e28)+(e21*e29+e22*e30+e23*e31+e24*e32+e25*e33+e26*e34+e27*e35+e4↑2)/2
2990 j9a=2*e0*e9+e1*(e8+e10)+e2*(e7+e11)+e3*(e6+e12)+e4*(e5+e13)+e14*(e5+e23)+e15*(e6+e24)+e16*(e7+e25)+e17*(e8+e26)+e18*(e9+e27)+e1
9*(e10+e28)
3000 j9=j9a+e20*e11+e21*e12+e22*e13+(e20*e29+e21*e30+e22*e31+e23*e32+e24*e33+e25*e34+e26*e35)/2
3010 j10a=2*e0*e10+e1*(e9+e11)+e2*(e8+e12)+e3*(e7+e13)+e4*(e6+e14)+e5*e15+e16*(e6+e26)+e17*(e7+e27)+e18*(e8+e28)+e19*e9+e20*e10+e21*
e11+e22*e12
3020 j10=j10a+e23*e13+e24*e14+e25*e15+(e19*e29+e20*e30+e21*e31+e22*e32+e23*e33+e24*e34+e25*e35+e5↑2)/2
3030 j11a=2*e0*e11+e1*(e10+e12)+e2*(e9+e13)+e3*(e8+e14)+e4*(e7+e15)+e5*(e6+e16)+e14*e25+e15*e26+e16*e27+e17*(e6+e28)+e18*e7+e19*e8+e
20*e9+e21*e10+e22*e11
3040 j11=j11a+e23*e12+e24*e13+(e18*e29+e19*e30+e20*e31+e21*32+e22*e33+e23*e34+e24*e35)/2
3050 j12a=2*e0*e12+e1*(e11+e13)+e2*(e10+e14)+e3*(e9+e15)+e4*(e8+e16)+e5*e17+e6*e18+e7*(e5+e19)+e20*e8+e21*e9+e22*e10+e23*e11+e24*e12
+e25*e13+e26*e14+e27*e15+e28*e16
3060 j12=j12a+(e17*e29+e18*e30+e19*e31+e20*e32+e21*e33+e22*e34+e23*e35+e6↑2)/2
3070 j13a=2*e0*e13+e1*(e12+e14)+e2*(e11+e15)+e3*(e10+e16)+e4*(e9+e17)+e5*(e8+e18)+e6*(e7+e19)+e20*e7+e21*e8+e22*e9+e23*e10+e24*e11+e
25*e12+e26*e13+e27*e14+e28*e15
3080 j13=j13a+(e16*e29+e17*e30+e18*e31+e19*e32+e20*e33+e21*e34+e22*e35)/2
3090 j14a=2*e0*e14+e1*(e13+e15)+e2*(e12+e16)+e3*(e11+e17)+e4*(e10+e18)+e5*(e9+e19)+e6*(e8+e20)+e7*e21+e22*e8+e23*e9+e24*e10+e25*e11+
e26*e12+e27*e13+e28*e14
3100 j14=j14a+(e15*e29+e16*e30+e17*e31+e18*e32+e19*e33+e20*e34+e21*e35+e7↑2)/2
3110 j15a=2*e0*e15+e1*(e14+e16)+e2*(e13+e17)+e3*(e12+e18)+e4*(e11+e19)+e5*(e10+e20)+e6*(e9+e21)+e7*(e8+e22)+e23*e8+e24*e9+e25*e10+e2
6*e11+e27*e12+e28*e13
3120 j15=j15a+(e14*e29+e15*e30+e16*e31+e17*e32+e18*e33+e19*e34+e20*e35)/2
3130 j16a=2*e0*e16+e1*(e15+e17)+e2*(e14+e18)+e3*(e13+e19)+e4*(e12+e20)+e5*(e11+e21)+e6*(e10+e22)+e7*(e9+e23)+e8*e24+e25*e9+e26*e10+e
27*e11+e28*e12
3140 j16=j16a+(e13*e29+e14*e30+e15*e31+e16*e32+e17*e33+e18*e34+e19*e35+e8↑2)/2
```

```
3150 j17a=2*e0*e17+e1*(e16+e18)+e2*(e15+e19)+e3*(e14+e20)+e4*(e13+e21)+e5*(e12+e22)+e6*(e11+e23)+e7*(e10+e24)+e8*(e9+e25)+e27*e10+e2
8*e11
3160 j17=j17a+e26*e9+(e12*e29+e13*e30+e14*e31+e15*e32+e16*e33+e17*e34+e18*e35)/2
3170 j18a=2*e0*e18+e1*(e17+e19)+e2*(e16+e20)+e3*(e15+e21)+e4*(e14+e22)+e5*(e13+e23)+e6*(e12+e24)+e7*(e11+e25)+e8*(e10+e26)+e9*e27+e2
8*e10
3180 j18=j18a+(e11*e29+e12*e30+e13*e31+e14*e32+e15*e33+e16*e34+e17*e35+e9^2)/2
3190 j19a=2*e0*e19+e1*(e18+e20)+e2*(e17+e21)+e3*(e16+e22)+e4*(e15+e23)+e5*(e14+e24)+e6*(e13+e25)+e7*(e12+e26)+e8*(e11+e27)+e9*(e10+e
28)
3200 j19=j19a+(e10*e29+e11*e30+e12*e31+e13*e32+e14*e33+e15*e34+e16*e35)/2
3210 j20a=2*e0*e20+e1*(e19+e21)+e2*(e18+e22)+e3*(e17+e23)+e4*(e16+e24)+e5*(e15+e25)+e6*(e14+e26)+e7*(e13+e27)+e8*(e12+e28)+e9*e11+(e
9*e29+e10*e30+e11*e31+e12*e32+e13*e33)/2
3220 j20=j20a+(e14*e34+e15*e35+e10^2)/2
3230 j21a=2*e0*e21+e1*(e20+e22)+e2*(e19+e23)+e3*(e18+e24)+e4*(e17+e25)+e5*(e16+e26)+e6*(e15+e27)+e7*(e14+e28)+e8*e13+e9*e12+e10*e11
3240 j21=j21a+(e8*e29+e9*e30+e10*e31+e11*e32+e12*e33+e13*e34+e14*e35)/2
3250 j22a=2*e0*e22+e1*(e21+e23)+e2*(e20+e24)+e3*(e19+e25)+e4*(e18+e26)+e5*(e17+e27)+e6*(e16+e28)+e7*e15+e8*e14+e9*e13+e10*e12+(e7*e2
9+e8*e30+e9*e31+e10*e32+e11*e33)/2
3260 j22=j22a+(e12*e34+e13*e35+e11^2)/2
3270 j23a=2*e0*e23+e1*(e22+e24)+e2*(e21+e25)+e3*(e20+e26)+e4*(e19+e27)+e5*(e18+e28)+e6*e17+e7*e16+e8*e15+e9*e14+e10*e13+e11*e12
3280 j23=j23a+(e6*e29+e7*e30+e8*e31+e9*e32+e10*e33+e11*e34+e12*e35)/2
3290 j24a=2*e0*e24+e1*(e23+e25)+e2*(e22+e26)+e3*(e21+e27)+e4*(e20+e28)+e5*e19+e6*e18+e7*e17+e8*e16+e9*e15+e10*e14+e11*e13
3300 j24=j24a+(e5*e29+e6*e30+e7*e31+e8*e32+e9*e33+e10*e34+e11*e35+e12^2)/2
3310 j25a=2*e0*e25+e1*(e24+e26)+e2*(e23+e27)+e3*(e22+e28)+e4*e21+e5*e20+e6*e19+e7*e18+e8*e17+e9*e16+e10*e15+e11*14+e12*e13
3320 j25=j25a+(e4*e29+e5*e30+e6*e31+e7*e32+e8*e33+e9*e34+e10*e35)/2
3330 j26a=2*e0*e26+e1*(e25+e27)+e2*(e24+e28)+e3*e23+e4*e22+e5*e21+e6*e20+e7*e19+e8*e18+e9*e17+e10*e16+e11*e15+e12*e14
3340 j26=j26a+(e3*e29+e4*e30+e5*e31+e6*e32+e7*e33+e8*e34+e9*e35+e13^2)/2
3350 j27a=2*e0*e27+e1*(e26+e28)+e2*e25+e3*e24+e4*e23+e5*e22+e6*e21+e7*e20+e8*e19+e9*e18+e10*e17+e11*e16+e12*e15+e13*e14
3360 j27=j27a+(e2*e29+e3*e30+e4*e31+e5*e32+e6*e33+e7*e34+e8*e35)/2
3370 j28=2*e0*e28+e1*e27+e2*e26+e3*e25+e4*e24+e5*e23+e6*e22+e7*e21+e8*e20+e9*e19+e10*e18+e11*e17+e12*e16+e13*e15+(e1*e29+e2*e30+e3*e
31+e4*e32+e5*e33+e6*e34+e7*e35+e14^2)/2
3380 k0=(a1*j1+a2*j2+a3*j3+a4*j4+a5*j5+a6*j6+a7*j7)/2
3390 k1=a1*j0+(a1*j2+a2*(j1+j3)+a3*(j2+j4)+a4*(j3+j5)+a5*(j4+j6)+a6*(j5+j7)+a7*(j6+j8))/2
3400 k2=a2*j0+(a1*(j1+j3)+a2*j4+a3*(j1+j5)+a4*(j2+j6)+a5*(j3+j7)+a6*(j4+j8)+a7*(j5+j9))/2
3410 k3=a3*j0+(a1*(j2+j4)+a2*(j1+j5)+a3*j6+a4*(j1+j7)+a5*(j2+j8)+a6*(j3+j9)+a7*(j4+j10))/2
```

```
3470 k9=(a1*(j8+j10)+a2*(j7+j11)+a3*(j6+j12)+a4*(j5+j13)+a5*(j4+j14)+a6*(j3+j15)+a7*(j2+j16))/2
3480 k10=(a1*(j9+j11)+a2*(j8+j12)+a3*(j7+j13)+a4*(j6+j14)+a5*(j5+j15)+a6*(j4+j16)+a7*(j3+j17))/2
3490 k11=(a1*(j10+j12)+a2*(j9+j13)+a3*(j8+j14)+a4*(j7+j15)+a5*(j6+j16)+a6*(j5+j17)+a7*(j4+j18))/2
3500 k12=(a1*(j11+j13)+a2*(j10+j14)+a3*(j9+j15)+a4*(j8+j16)+a5*(j7+j17)+a6*(j6+j18)+a7*(j5+j19))/2
3510 k13=(a1*(j12+j14)+a2*(j11+j15)+a3*(j10+j16)+a4*(j9+j17)+a5*(j8+j18)+a6*(j7+j19)+a7*(j6+j20))/2
3520 k14=(a1*(j13+j15)+a2*(j12+j16)+a3*(j11+j17)+a4*(j10+j18)+a5*(j9+j19)+a6*(j8+j20)+a7*(j7+j21))/2
3530 k15=(a1*(j14+j16)+a2*(j13+j17)+a3*(j12+j18)+a4*(j11+j19)+a5*(j10+j20)+a6*(j9+j21)+a7*(j8+j22))/2
3540 k16=(a1*(j15+j17)+a2*(j14+j18)+a3*(j13+j19)+a4*(j12+j20)+a5*(j11+j21)+a6*(j10+j22)+a7*(j9+j23))/2
3550 k17=(a1*(j16+j18)+a2*(j15+j19)+a3*(j14+j20)+a4*(j13+j21)+a5*(j12+j22)+a6*(j11+j23)+a7*(j10+j24))/2
3560 k18=(a1*(j17+j19)+a2*(j16+j20)+a3*(j15+j21)+a4*(j14+j22)+a5*(j13+j23)+a6*(j12+j24)+a7*(j11+j25))/2
3570 k19=(a1*(j18+j20)+a2*(j17+j21)+a3*(j16+j22)+a4*(j15+j23)+a5*(j14+j24)+a6*(j13+j25)+a7*(j12+j26))/2
3580 k20=(a1*(j19+j21)+a2*(j18+j22)+a3*(j17+j23)+a4*(j16+j24)+a5*(j15+j25)+a6*(j14+j26)+a7*(j13+j27))/2
3590 k21=(a1*(j20+j22)+a2*(j19+j23)+a3*(j18+j24)+a4*(j17+j25)+a5*(j16+j26)+a6*(j15+j27)+a7*(j14+j28))/2
3600 l0=(a1*k1+a2*k2+a3*k3+a4*k4+a5*k5+a6*k6+a7*k7)/2
3610 l1=a1*k0+(a1*k2+a2*(k1+k3)+a3*(k2+k4)+a4*(k3+k5)+a5*(k4+k6)+a6*(k5+k7)+a7*(k6+k8))/2
3620 l2=a2*k0+(a1*(k1+k3)+a2*k4+a3*(k1+k5)+a4*(k2+k6)+a5*(k3+k7)+a6*(k4+k8)+a7*(k5+k9))/2
3630 l3=a3*k0+(a1*(k2+k4)+a2*(k1+k5)+a3*k6+a4*(k1+k7)+a5*(k2+k8)+a6*(k3+k9)+a7*(k4+k10))/2
3640 l4=a4*k0+(a1*(k3+k5)+a2*(k2+k6)+a3*(k1+k7)+a4*k8+a5*(k1+k9)+a6*(k2+k10)+a7*(k3+k11))/2
3650 l5=a5*k0+(a1*(k4+k6)+a2*(k3+k7)+a3*(k2+k8)+a4*(k1+k9)+a5*k10+a6*(k1+k11)+a7*(k2+k12))/2
3660 l6=a6*k0+(a1*(k5+k7)+a2*(k4+k8)+a3*(k3+k9)+a4*(k2+k10)+a5*(k1+k11)+a6*k12+a7*(k1+k13))/2
3670 l7=a7*k0+(a1*(k6+k8)+a2*(k5+k9)+a3*(k4+k10)+a4*(k3+k11)+a5*(k2+k12)+a6*(k1+k13)+a7*k14)/2
3680 l8=(a1*(k7+k9)+a2*(k6+k10)+a3*(k5+k11)+a4*(k4+k12)+a5*(k3+k13)+a6*(k2+k14)+a7*(k1+k15))/2
3690 l9=(a1*(k8+k10)+a2*(k7+k11)+a3*(k6+k12)+a4*(k5+k13)+a5*(k4+k14)+a6*(k3+k15)+a7*(k2+k16))/2
3700 l10=(a1*(k9+k11)+a2*(k8+k12)+a3*(k7+k13)+a4*(k6+k14)+a5*(k5+k15)+a6*(k4+k16)+a7*(k3+k17))/2
3710 l11=(a1*(k10+k12)+a2*(k9+k13)+a3*(k8+k14)+a4*(k7+k15)+a5*(k6+k16)+a6*(k5+k17)+a7*(k4+k18))/2
3720 l12=(a1*(k11+k13)+a2*(k10+k14)+a3*(k9+k15)+a4*(k8+k16)+a5*(k7+k17)+a6*(k6+k18)+a7*(k5+k19))/2
3730 l13=(a1*(k12+k14)+a2*(k11+k15)+a3*(k10+k16)+a4*(k9+k17)+a5*(k8+k18)+a6*(k7+k19)+a7*(k6+k20))/2
3740 l14=(a1*(k13+k15)+a2*(k12+k16)+a3*(k11+k17)+a4*(k10+k18)+a5*(k9+k19)+a6*(k8+k20)+a7*(k7+k21))/2
3750 m0=(a1*l1+a2*l2+a3*l3+a4*l4+a5*l5+a6*l6+a7*l7)/2
3760 m1=(a1*(2*l0+l2)+a2*(l1+l3)+a3*(l2+l4)+a4*(l3+l5)+a5*(l4+l6)+a6*(l5+l7)+a7*(l6+l8))/2
3770 m2=(a1*(l1+l3)+a2*(2*l0+l4)+a3*(l1+l5)+a4*(l2+l6)+a5*(l3+l7)+a6*(l4+l8)+a7*(l5+l9))/2
3780 m3=(a1*(l2+l4)+a2*(l1+l5)+a3*(2*l0+l6)+a4*(l1+l7)+a5*(l2+l8)+a6*(l3+l9)+a7*(l4+l10))/2
3790 m4=(a1*(l3+l5)+a2*(l2+l6)+a3*(l1+l7)+a4*(2*l0+l8)+a5*(l1+l9)+a6*(l2+l10)+a7*(l3+l11))/2
3800 m5=(a1*(l4+l6)+a2*(l3+l7)+a3*(l2+l8)+a4*(l1+l9)+a5*(2*l0+l10)+a6*(l1+l11)+a7*(l2+l12))/2
3810 m6=(a1*(l5+l7)+a2*(l4+l8)+a3*(l3+l9)+a4*(l2+l10)+a5*(l1+l11)+a6*(2*l0+l12)+a7*(l1+l13))/2
3820 m7=(a1*(l6+l8)+a2*(l5+l9)+a3*(l4+l10)+a4*(l3+l11)+a5*(l2+l12)+a6*(l1+l13)+a7*(2*l0+l14))/2
```

```
5000 k=2*PI*SQR(2)*Rq/q
5010 r0=1-2*k¼2*b0+2*k¼4*d0/3-4*k¼6*f0/45+2*k¼8*h0/315-4*k¼10*j0/14075+4*k¼12*l0/464475
5020 r1=-2*k¼2*b1+2*k¼4*d1/3-4*k¼6*f1/45+2*k¼8*h1/315-4*k¼10*j1/14175+4*k¼12*l1/467775
5030 r2=-2*k¼2*b2+2*k¼4*d2/3-4*k¼6*f2/45+2*k¼8*h2/315-4*k¼10*j2/14175+4*k¼12*l2/467775
5040 r3=-2*k¼2*b3+2*k¼4*d3/3-4*k¼6*f3/45+2*k¼8*h3/315-4*k¼10*j3/14175+4*k¼12*l3/467775
5050 r4=-2*k¼2*b4+2*k¼4*d4/3-4*k¼6*f4/45+2*k¼8*h4/315-4*k¼10*j4/14175+4*k¼12*l4/467775
5060 r5=-2*k¼2*b5+2*k¼4*d5/3-4*k¼6*f5/45+2*k¼8*h5/315-4*k¼10*j5/14175+4*k¼12*l5/467775
5070 r6=-2*k¼2*b6+2*k¼4*d6/3-4*k¼6*f6/45+2*k¼8*h6/315-4*k¼10*j6/14175+4*k¼12*l6/467775
5080 r7=-2*k¼2*b7+2*k¼4*d7/3-4*k¼6*f7/45+2*k¼8*h7/315-4*k¼10*j7/14175+4*k¼12*l7/467775
5090 t0=-2*k*a0+4*(k¼3*c0/3-k¼5*e0/15+2*k¼7*g0/315-k¼9*i0/2835+k¼11*k0/(55*2835)-k¼13*m0/(55*39*2835))
5100 t1=-2*k*a1+4*(k¼3*c1/3-k¼5*e1/15+2*k¼7*g1/315-k¼9*i1/2835+k¼11*k1/(55*2835)-k¼13*m1/(55*39*2835))
5110 t2=-2*k*a2+4*(k¼3*c2/3-k¼5*e2/15+2*k¼7*g2/315-k¼9*i2/2835+k¼11*k2/(55*2835)-k¼13*m2/(55*39*2835))
5120 t3=-2*k*a3+4*(k¼3*c3/3-k¼5*e3/15+2*k¼7*g3/315-k¼9*i3/2835+k¼11*k3/(55*2835)-k¼13*m3/(55*39*2835))
5130 t4=-2*k*a4+4*(k¼3*c4/3-k¼5*e4/15+2*k¼7*g4/315-k¼9*i4/2835+k¼11*k4/(55*2835)-k¼13*m4/(55*39*2835))
5140 t5=-2*k*a5+4*(k¼3*c5/3-k¼5*e5/15+2*k¼7*g5/315-k¼9*i5/2835+k¼11*k5/(55*2835)-k¼13*m5/(55*39*2835))
5150 t6=-2*k*a6+4*(k¼3*c6/3-k¼5*e6/15+2*k¼7*g6/315-k¼9*i6/2835+k¼11*k6/(55*2835)-k¼13*m6/(55*39*2835))
5160 t7=-2*k*a7+4*(k¼3*c7/3-k¼5*e7/15+2*k¼7*g7/315-k¼9*i7/2835+k¼11*k7/(55*2835)-k¼13*m7/(55*39*2835))
5170 o0=r0¼2+t0¼2
5180 o1=(r1¼2+t1¼2)/4
5190 o2=(r2¼2+t2¼2)/4
5200 o3=(r3¼2+t3¼2)/4
5210 o5=(r5¼2+t5¼2)/4
5220 o4=(r4¼2+t4¼2)/4
5230 o6=(r6¼2+t6¼2)/4
5240 o7=(r7¼2+t7¼2)/4
5250 t=o0+2*(o1+o2+o3+o4+o5+o6+o7)
5260 Rq=Rq+0.005
5270 PLOT 100+1000*rq,100+100*o0
5280 PLOT 100+1000*rq,100+100*o1
5290 PLOT 100+1000*rq,100+100*o2
5295 PLOT 100+1000*rq,100+100*o3
5340 PLOT 100+1000*rq,100+100
5350 PLOT 100+1000*rq,100+100*t
5360 GOTO 5000
```

Appendix A4. Derivation of the algorithms to extract R_q/λ from the diffraction pattern.

$J_o(x)$ can be expressed by the series expansion:

$$J_o(x) = \sum_{n=0}^{\infty} \left(-\left(\frac{x}{2}\right)^2\right)^n / (n!)^2 \quad \text{(A4.1)}$$

Depending on the value of x, only a few terms have to be used to calculate $J_o(x)$. This leads to the following three equations with their validity ranges.

$$0.0<x<0.4 \text{ or } 0.96<J_o<1 \quad ; J_o(x)=1-(x/2)^2 \quad \text{(A4.2)}$$

$$0.4<x<1.0 \text{ or } 0.77<J_o<0.96 \quad ; J_o(x)=1-(x/2)^2+(x/2)^4/4 \quad \text{(A4.3)}$$

$$1.0<x<1.7 \text{ or } 0.40<J_o<0.77 \quad ; J_o(x)=1-(x/2)^2+(x/2)^4/4-(x/2)^6/36 \quad \text{(A4.4)}$$

with $x=4\pi\sqrt{2}\, R_q/\lambda$ and $J_o=\sqrt{\dfrac{B_o^2}{\sum_{m=-\infty}^{\infty} B_m^2}}$

Equation (A4.2) leads to: $x=2\sqrt{(1-J_o)}$ or

$$R_q/\lambda=\sqrt{2(1-J_o)}\,/4\pi \quad \text{(A4.5)}$$

Equation (A4.3) leads to: $(x/2)^4-4(x/2)^2+4(1-J_o)=0$, which is a quadratic equation in $(x/2)^2$ and has as solutions:

$$(x/2)^2=(4\pm\sqrt{16-16(1-J_o)}\,)/2=2(1\pm\sqrt{J_o})$$

When J_o=1 then $(x/2)^2$ has to be zero, so the solution is

$$(x/2)^2=2(1-\sqrt{J_o}) \text{ or } x=2\sqrt{2(1-\sqrt{J_o})} \text{ or}$$

$$R_q/\lambda=\sqrt{(1-\sqrt{J_o})}/2\pi \qquad (A4.6)$$

Equation (A4.4) leads to a cubic equation of the form:

$$X^3+a_1X^2+a_2X+a_3=0 \qquad (A4.7)$$

with a_1=-9, a_2=36 and $a_3=-36(1-J_o)$. This eqation can be solved as follows:

Let $Q=(3a_1-a_1^2)/9=3$

$R=(9a_1a_2-27a_3-2a_1^3)/54=-9(1+2J_o)$

$D=Q^3+R^2=108(1+3J_o+3J_o^2)$

$S=\sqrt[3]{R+\sqrt{D}}$ and $T=\sqrt[3]{R-\sqrt{D}}$

This cubic equation has one real root since D>0. This root is:

$$(x/2)^2=S+T-a1/3=S+T+3 \text{ or}$$

$$R_q/\lambda=\sqrt{2(S+T+3)}/4\pi \qquad (A4.8)$$

Appendix A5. Accuracy of the algorithms for the determination of R_q/λ from the value of the zeroth order of the diffraction pattern.

The accuracy with which R_q/λ can be calculated by the three algorithms (9.42a–c) depends on the accuracy with which the normalised zeroth order intensity I_o can be obtained, since $R_q/\lambda = f(J_o)$ and $J_o = \sqrt{I_o}$. This accuracy can be estimated by calculating the fractional error $\Delta R_q/R_q$ as a function of $\Delta I_o/I_o$, i.e. the accuracy with which I_o can be determined. This gives the following results for the three algorithms.

Algorithm 1.

$$R_q = c \cdot \sqrt{1 - \sqrt{I_o}}$$

and

$$\Delta R_q = \frac{d}{dI_o}(R_q) \cdot \Delta I_o = c \cdot \frac{1}{2} \cdot \frac{1}{\sqrt{1-\sqrt{I_o}}} \cdot \left(\frac{-1}{2}\right) \cdot \frac{1}{\sqrt{I_o}} \, \Delta I_o$$

from which the relationship between the fractional errors follows:

$$\frac{\Delta R_q}{R_q} = \frac{-1}{4} \, \frac{\sqrt{I_o}}{(1-\sqrt{I_o})} \, \frac{\Delta I_o}{I_o} = -\frac{1}{4} \cdot \frac{J_o}{(1-J_o)} \, \frac{\Delta I_o}{I_o}$$

Algorithm 2.

$$R_q = c \cdot \sqrt{1 - \sqrt[4]{I_o}}$$

and

$$\Delta R_q = c \cdot \frac{1}{2} \, \frac{1}{\sqrt{1-\sqrt[4]{I_o}}} \cdot \left(\frac{-1}{4}\right) \, \frac{1}{\sqrt[4]{I_o^3}} \cdot \Delta I_o$$

and the relationship between the fractional errors is:

$$\frac{\Delta R_q}{R_q} = -\frac{1}{8}\,\frac{\sqrt{I_0}}{(1-\sqrt[4]{I_0})}\,\frac{\Delta I_0}{I_0} = -\frac{1}{8}\,\frac{J_0}{(1-\sqrt{J_0})}\,\frac{\Delta I_0}{I_0}$$

Algorithm 3.

$$R_q = c\sqrt{S+T+3}$$

and

$$\Delta R_q = c\cdot\frac{1}{2}\,\frac{1}{\sqrt{S+T+3}}\left(\frac{dS}{dI_0}+\frac{dT}{dI_0}\right)\cdot\Delta I_0$$

The fractional errors can then be related by:

$$\frac{\Delta R_q}{R_q} = \frac{1}{2}\,\frac{I_0}{(S+T+3)}\left(\frac{dS}{dI_0}+\frac{dT}{dI_0}\right)\frac{\Delta I_0}{I_0}$$

with $D = 108\left(1+3J_0+3J_0^2\right)$, $R = -9\left(1+2J_0\right)$

$$S = \sqrt[3]{R+\sqrt{D}} \quad , T = \sqrt[3]{R-\sqrt{D}}$$

$$\frac{dS}{dI_0} = \frac{1}{\sqrt[3]{(R+\sqrt{D})^2}}\cdot\left(\frac{-3}{\sqrt{I_0}} + \frac{27}{\sqrt{D}}\left(\frac{1}{\sqrt{I_0}}+2\right)\right)$$

$$\frac{dT}{dI_0} = \frac{1}{\sqrt[3]{(R+\sqrt{D})^2}}\left(\frac{-3}{\sqrt{I_0}} - \frac{27}{\sqrt{D}}\left(\frac{1}{\sqrt{I_0}}+2\right)\right)$$

The results from these expressions are depicted in figure 9.6.

Appendix A6. Calculation of the second and fourth moments of the diffraction pattern for a sinusoidal surface profile.

In order to calculate the second and fourth moments, various relationships are derived with the aid of Neumann's addition theorem.

$$J_n(u \mp v) = \sum_{k=-\infty}^{\infty} J_{n \mp k}(u) \cdot J_k(v) \qquad \text{(A6.1)}$$

From this the following equalities can be derived for the various combination of values of n, k, u and v.

(a) u-v=0, n=0, k=m : $J_0(0) = 1 = \sum_{m=-\infty}^{\infty} J_m^2(u)$

(b) u-v=0, n=-2, k=m+1 : $J_{-2}(0) = 0 = \sum_{m=-\infty}^{\infty} J_{m-1}(u) \cdot J_{m+1}(u)$

(c) u-v=0, n=-4, k=m+2 : $J_{-4}(0) = 0 = \sum_{m=-\infty}^{\infty} J_{m-2}(u) \cdot J_{m+2}(u)$

(d) u-v=0, n=-2, k=m : $J_{-2}(0) = 0 = \sum_{m=-\infty}^{\infty} J_{m-2}(u) \cdot J_m(u)$

(e) u-v=0, n=2, k=m : $J_2(0) = 0 = \sum_{m=-\infty}^{\infty} J_{m+2}(u) \cdot J_m(u)$

Furthermore, the recurrence relationship is used:

(f) : $J_{n-1}(u) + J_{n+1}(u) = \frac{2n}{u} J_n(u) \qquad \text{(A6.2)}$

The second moment M_2 for a surface profile $y(x)=h\cos(2\pi x/D)$ is:

$$M_2 = \sum_{n=-\infty}^{\infty} \left(\frac{n\lambda}{D}\right)^2 J_n^2\left(\frac{4\pi h}{\lambda}\right) = \left(\frac{\lambda}{D}\right)^2 \sum_{n=-\infty}^{\infty} n^2 J_n^2(u) \tag{A6.3}$$

with $u=4\pi h/\lambda$. Using property (f) leads to:

$$M_2 = \left(\frac{\lambda}{D}\right)^2 \sum_{n=-\infty}^{\infty} \left(\frac{u}{2}\right)^2 \left(\frac{2n}{u} J_n(u)\right)^2 = \left(\frac{\lambda}{D}\right)^2 \left(\frac{u}{2}\right)^2 \sum_{n=-\infty}^{\infty} \left(J_{n-1}(u) + J_{n+1}(u)\right)^2$$

or

$$M_2 = \left(\frac{\lambda}{D}\right)^2 \left(\frac{u}{2}\right)^2 \sum_{n=-\infty}^{\infty} \left(J_{n-1}^2(u) + J_{n+1}^2(u) + 2 J_{n-1}(u) \cdot J_{n+1}(u)\right)$$

Now using (a) and (f) yields:

$$M_2 = 2 \left(\frac{\lambda}{D}\right)^2 \left(\frac{u}{2}\right)^2$$

or

$$M_2 = 2 \left(\frac{2\pi h}{D}\right)^2 \tag{A6.4}$$

The fourth moment M_4 is defined by:

$$M_4 = \sum_{n=-\infty}^{\infty} \left(\frac{n\lambda}{D}\right)^4 J_n^2(u) = \left(\frac{\lambda}{D}\right)^4 \left(\frac{u}{2}\right)^2 \sum_{n=-\infty}^{\infty} n^2 \left(\frac{2n}{u} J_n(u)\right)^2 \tag{A6.5}$$

With property (f) this can be rewritten as:

$$M_4 = \left(\frac{\lambda}{D}\right)^4 \left(\frac{u}{2}\right)^2 \sum_{n=-\infty}^{\infty} n^2 \left(J_{n-1}(u) + J_{n+1}(u)\right)^2 = \left(\frac{\lambda}{D}\right)^4 \left(\frac{u}{2}\right)^2 \sum_{n=-\infty}^{\infty} n^2 \left(J_{n-1}^2(u) + J_{n+1}^2(u) + 2 J_{n-1}(u) \cdot J_{n+1}(u)\right)$$

rewriting yields:

$$M_4 = \left(\frac{\lambda}{D}\right)^4 \left(\frac{u}{2}\right)^2 \sum_{n=-\infty}^{\infty} \Big[(n-1)^2 J_{n-1}^2(u) + (n+1)^2 J_{n+1}^2(u) + 2(n-1) J_{n-1}(u)(n+1) J_{n+1}(u) + $$
$$ + (2n-1) J_{n-1}^2(u) - (2n+1) J_{n+1}^2(u) - 2 J_{n-1}(u) \cdot J_{n+1}(u) \Big]$$

Here the last term equals zero because of (b), rearranging gives:

$$M_4 = \left(\frac{\lambda}{D}\right)^4 \left(\frac{u}{2}\right)^2 \sum_{n=-\infty}^{\infty} \Big[\left(\frac{u}{2}\right)^2 \Big[\left(\frac{2(n-1)}{u} J_{n-1}\right)^2 + \left(\frac{2(n+1)}{u} J_{n+1}(u)\right)^2 + $$
$$ + 2 \cdot \frac{2(n-1)}{u} J_{n-1}(u) \cdot \frac{2(n+1)}{u} J_{n+1}(u) \Big] + 2(n-1) J_{n-1}^2(u) + J_{n-1}^2(u) + J_{n+1}^2(u) - 2(n+1) J_{n+1}^2(u) \Big]$$

Since $\sum_{n=-\infty}^{\infty} 2(n-1) J_{n-1}^2(u) = \sum_{n=-\infty}^{\infty} 2(n+1) J_{n+1}^2(u)$ and with property (a)

$$M_4 = \left(\frac{\lambda}{D}\right)^4 \left(\frac{u}{2}\right)^2 \sum_{n=-\infty}^{\infty} \Big[\left(\frac{u}{2}\right)^2 \Big[\big(J_{n-2}(u) + J_n(u) \big)^2 + \big(J_n(u) + J_{n+2}(u) \big)^2 + $$
$$ + 2\big(J_{n-2}(u) + J_n(u) \big)\big(J_n(u) + J_{n+2}(u) \big) \Big] + 2 \Big]$$

or

$$M_4 = \left(\frac{\lambda}{D}\right)^4 \left(\frac{u}{2}\right)^2 \sum \left(\frac{u}{2}\right)^2 \Big[J_{n-2}^2(u) + J_n^2(u) + 2 J_{n-2}(u) \cdot J_n(u) + J_n^2(u) + J_{n+2}^2(u) + $$
$$ + 2 J_n(u) \cdot J_{n+2}(u) + 2\big(J_{n-2}(u) \cdot J_n(u) + J_{n-2}(u) \cdot J_{n+2}(u) + J_n^2(u) + J_n(u) J_{n+2}(u) \big) \Big] + 2 \Big]$$

Now with the properties (a), (c), (d) and (e) M_4 becomes:

$$M_4 = 2\left(\frac{\lambda}{D}\right)^2 \left(\frac{u}{2}\right)^2 \left(3 \left(\frac{\lambda}{D}\right)^2 \left(\frac{u}{2}\right)^2 + \left(\frac{\lambda}{D}\right)^2 \right) = M_2 \left(\frac{3}{2} M_2 + \left(\frac{\lambda}{D}\right)^2 \right) \qquad \text{(A6.6)}$$

Appendix A7 Data sheets on the collimated laser diode and the photo diode array.

COLLIMATED LASER DIODE

The CQL13A has been designed for reading applications in optical data recording systems, security systems or other applications requiring a low level collimated source of near infrared emission.
The device is mounted in a non-hermetic encapsulation, specifically designed for simple alignment in an optical read/write system. The internal lens system collimates the diverging laser emission. The wavefront quality is diffraction limited. A cylindrical lens is incorporated for correction of the inherent astigmatism of a semiconductor laser.
The cylindrical housing has a closely controlled diameter with a tolerance of +0 and −11 μm and a socket is supplied with each device for connection to the drive circuit.

QUICK REFERENCE DATA

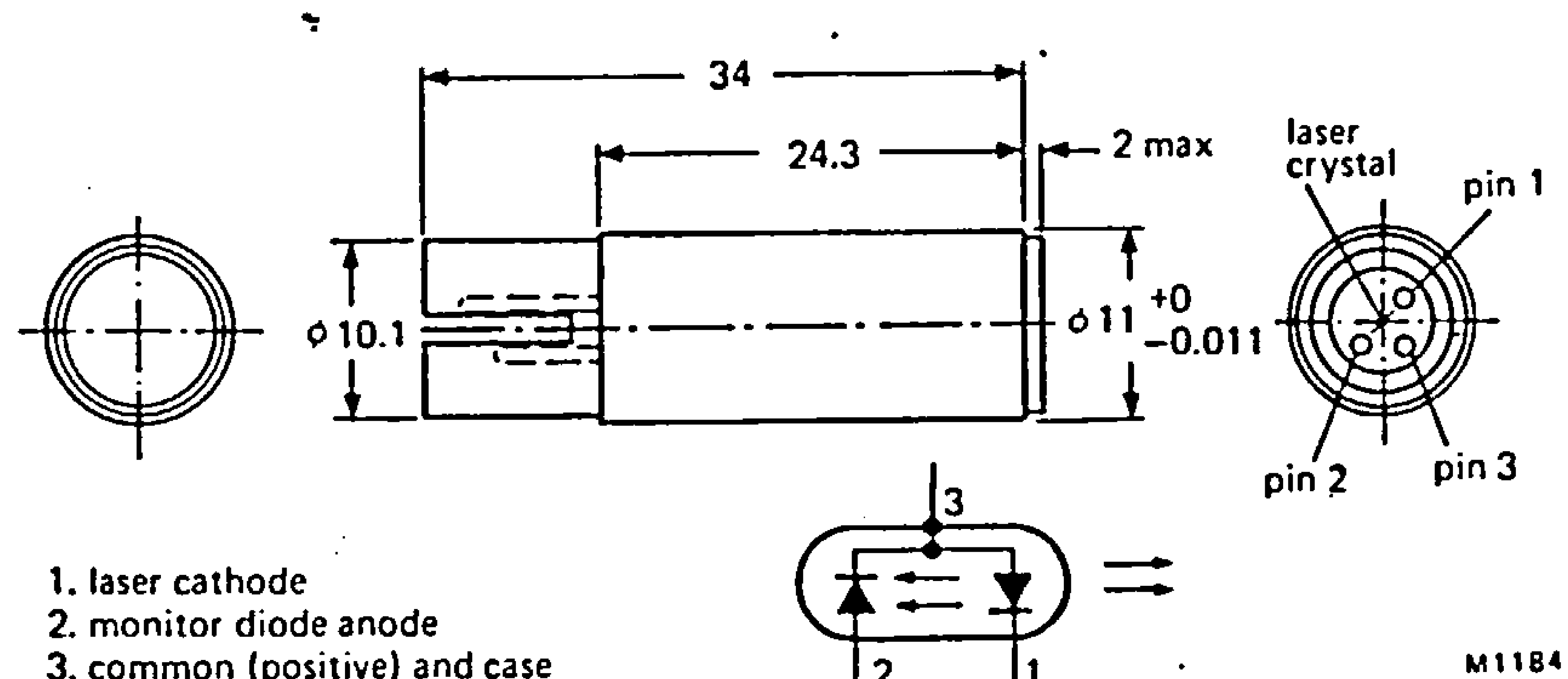
Output power	ϕ_e		2	mW
Current at output power ϕ_e = 2 mW and temperature of 60 °C		<	175	mA
Wavelength at peak emission	λ_{pk}	typ.	820	nm
Wavefront form of bundle (non-convergent) divergence		<	0.3	mrad

This data must be read in conjunction with GENERAL SAFETY RECOMMENDATIONS – OPTOELECTRONIC DEVICES

blue binder, tab 11

MECHANICAL DATA

Dimensions in mm



Mass	max. 8 g
Coaxial alignment	angle between mechanical and optical axis ⩽ 10 mrad

OPERATIONAL HAZARD – SEMICONDUCTOR LASER

This laser emits radiation which is invisible to the human eye. When in use, do not look directly into the device. Direct viewing of laser emission at close range may cause eye damage, especially in conjunction with collimating lenses.
The device falls into product class 3B as defined in BS4803: part 2: 1983, 'Radiation safety of laser products and systems'.
Users must observe the procedures and requirements appropriate to this class of laser as defined in BS4803.

PRODUCT SAFETY

Modern high technology materials have been used in the manufacture of this device to ensure high performance. Some of these materials are toxic in certain circumstances. Mechanical or electrical damage is unlikely to give rise to any hazard, but toxic vapours may be generated if the device is heated to destruction. In the United Kingdom, disposal of large quantities should therefore be carried out in accordance with the Deposit of Poisonous Waste Act 1972 and the control of Pollution Act 1974, or with the latest legislation.

LASER DIODE CHARACTERISTICS

Measuring conditions

Climatic	atmospheric pressure at a relative humidity of 5 to 90%. Case temperature 5 to 60 °C.
Electrical	d.c. operation, optical feedback drive and protection against transients, (see fig.2)
Optical	only the radiation within a diameter of 5.4 mm is relevant.

Optical data

Output power	ϕ_e	max. 2 mW
Spectral		
wavelength	λ	820 ± 10 nm
bandwidth	$\Delta\lambda$	< 4 nm (see note)
wavelength temperature coefficient	$\Delta\lambda/\Delta T$	approx. 0.25 nm/K
Bundle properties		
diameter		5.4 mm
wavefront form		plane (non-convergent), divergence < 0.3 mrad
aberrations		variance of wavefront with respect to the Gaussian reference sphere is less than $\lambda^2/300$
polarization		plane
polarization ratio		typ. 35 : 1
Intensity distribution		
transverse mode		TE_{00}-fundamental
longitudinal mode		see note
symmetry		variation of optical power in the four quadrants (see Fig.1) typ. 20%
ripple		mean intensity < 15% of the smooth value
filling ratio I_{rim}/I_{max}		> 0.17

Note: the number of longitudinal modes is related to the spectral width.

Electrical data

Current at ϕ_e = 2 mW and T_h = 60 °C	≤ 175 mA
Drive voltage V_d	≤ 5 V
Resistance between laser diode and connector at 10 mA, max. 20 mV_{pp} and 1 kHz	< 12 mΩ

Fig. 1

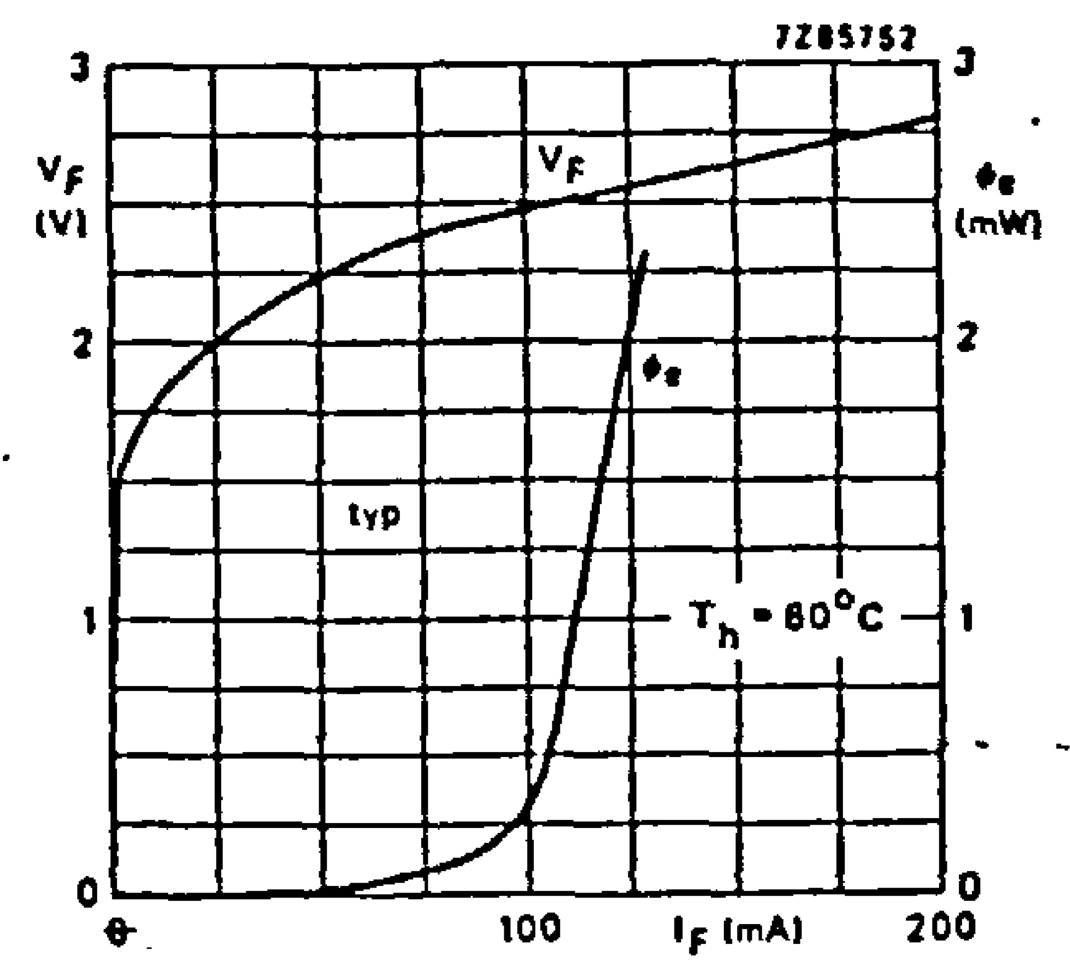
MONITOR DIODE

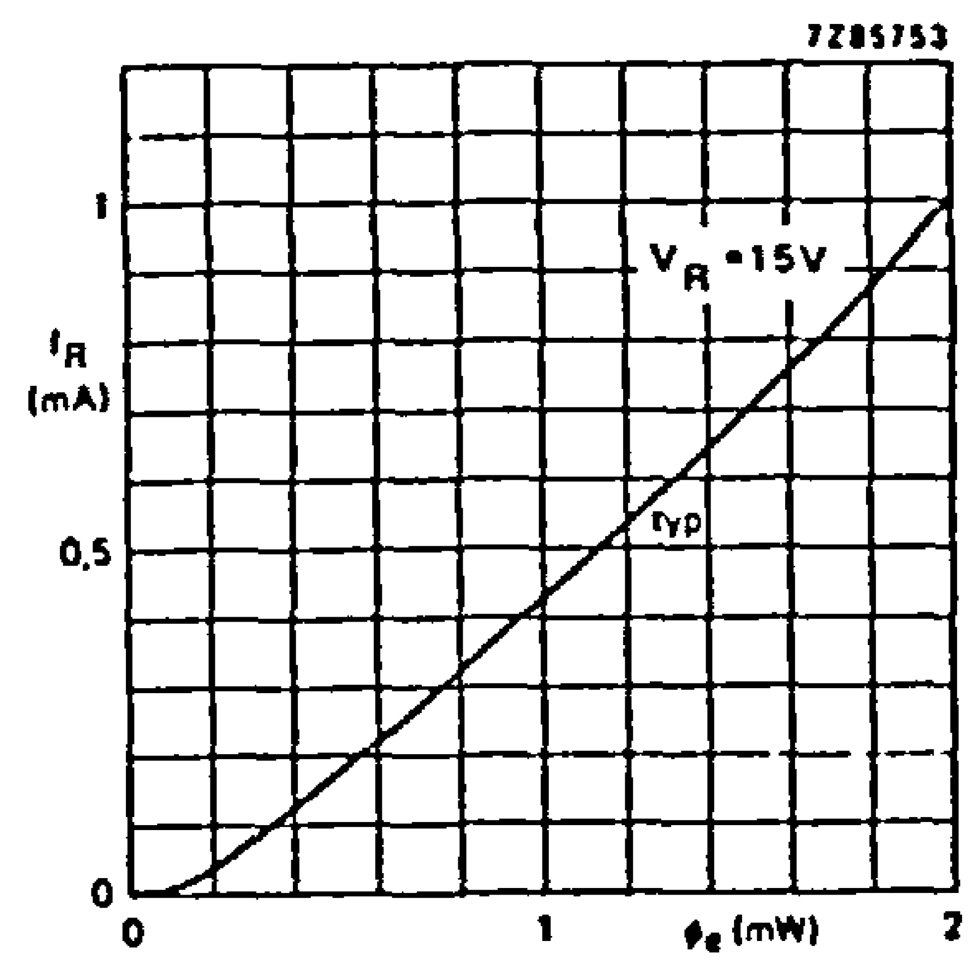
RATING

Reverse voltage	V_R	≤	30	V

CHARACTERISTICS

Sensitivity (ratio of monitor diode current to optical power emitted by laser) at ϕ_e = 2 mW			0.1	A/W
Dark reverse current at V_R = 15 V	$I_{R(D)}$	<	10	nA
Capacitance at V_R = 0	C_d	<	5	pF





ENVIRONMENTAL TESTS

Devices are subjected to the following non-operational tests; electrical testing is carried out two hours after removal from the test equipment.

test	in accordance with	conditions
Rapid change of temperature	IEC68-2-14, test Na	−25 °C to 25 °C to 70° to 25 °C; duration of each exposure 30 min and 10 cycles in total
Dry heat	IEC 68-2-2, test Bc	Temperature: 70 °C Duration: 7 days
Cold	IEC 68-2-1 test Aa	Temperature: −25 °C Duration: 7 days
Damp heat, steady state	IEC 68-2-3 test Ca	Temperature: 40 °C Relative humidity: 90 to 95% Duration: 42 days
Damp heat, cyclic	IEC 68-2-30	Temperature: 45 °C and relative humidity: 100% (12 h) to 25 °C and 85% (12 h) Duration of one cycle: 24 hours Number of cycles: 42

ENVIRONMENTAL TESTS (continued)

test	in accordance with	conditions
Vibration	IEC 68-2-6 test Db	Frequency range: 10 to 55 Hz Amplitude: 0.75 mm Duration: 6 hours (2 h in each of 3 directions)
Shock	IEC 68-2-27 test Ea	Pulse shape: half sine wave Pulse duration: 11 ms Peak acceleration: 981 m/s^2 Number of shocks: 10 in each of 3 directions
Bump	IEC 68-2-29 test Eb	Pulse duration: 6 ms Peak acceleration: 390 m/s^2 Number of bumps: 1000 Axes of direction: 6

Cleaning test — The exposed surface of the cylindrical lens is cleaned 100x with a small piece of cotton wool, attached to the end of a small stick, soaked in a 45 to 95% solution of ethyl alcohol in water.

OPERATING PRECAUTIONS

Semiconductor lasers in general are easily damaged by overdriving and transients. Electrically, the laser diode is a very reliable device and can easily withstand current surges of several amperes. Optically, however, the laser diode is more susceptible to damage because an extremely high optical flux density passes through both facets while in operation. The optical flux density can rise to unacceptable values, (10 to 100 MW/cm^2), if the laser diode is overdriven or subjected to transients. This will result in gradual or catastrophic degradation of the laser facets. Current transients should therefore be avoided; they decrease the life of the laser.

RECOMMENDED CONTROL CIRCUIT (available as accessory 56389)

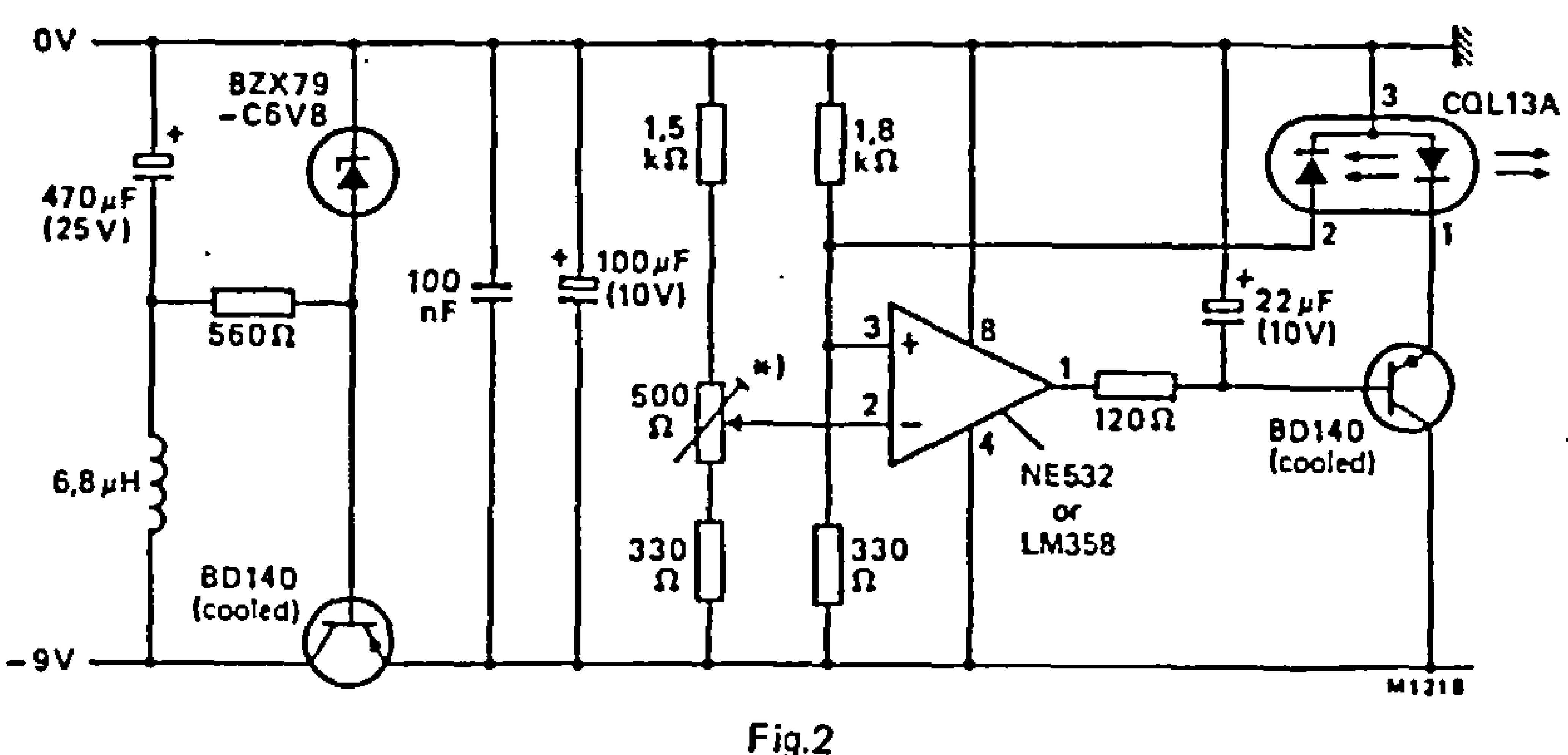
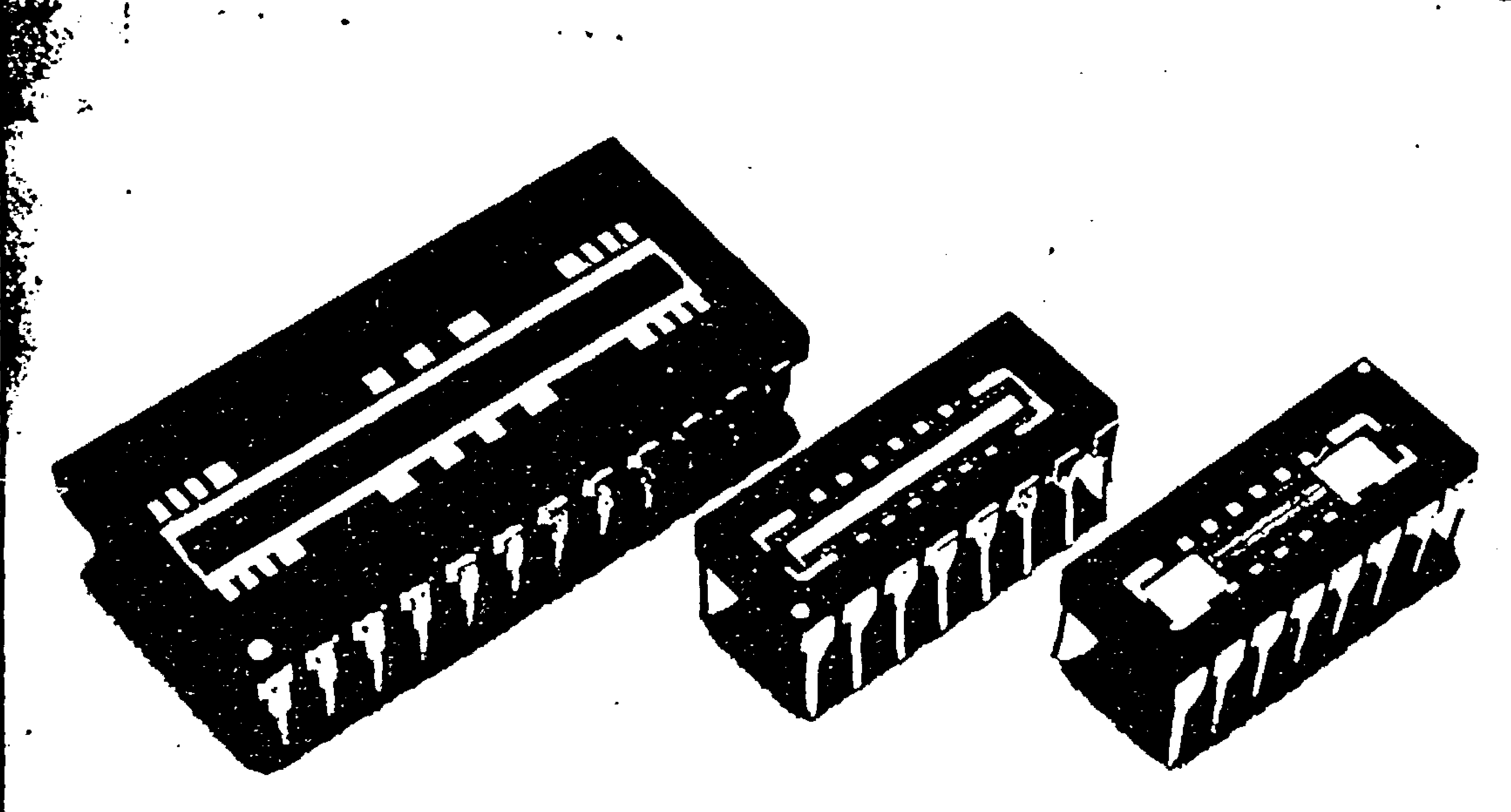


Fig.2

*Ten-turn. Zero position is at 0.58 revolution. Each revolution is equivalent to 500 µA monitor diode current.

IPL M Series Array Data

information sheet



Features

- Arrays of 256, 512, or 1024 on .001" centres
- Scan speeds to 10 MHz
- Recharge mode signal output.

Introduction

The IPL M-series of arrays are silicon self-scanned linear photodiode arrays with integrated scanning circuitry included on the same chip. The arrays are manufactured in three different sizes, 256 elements, 512 elements and 1024 elements. The diode pitch is .001". Two sizes of diode aperture are available to enable users to select the aperture width best suited to their application.

General Description

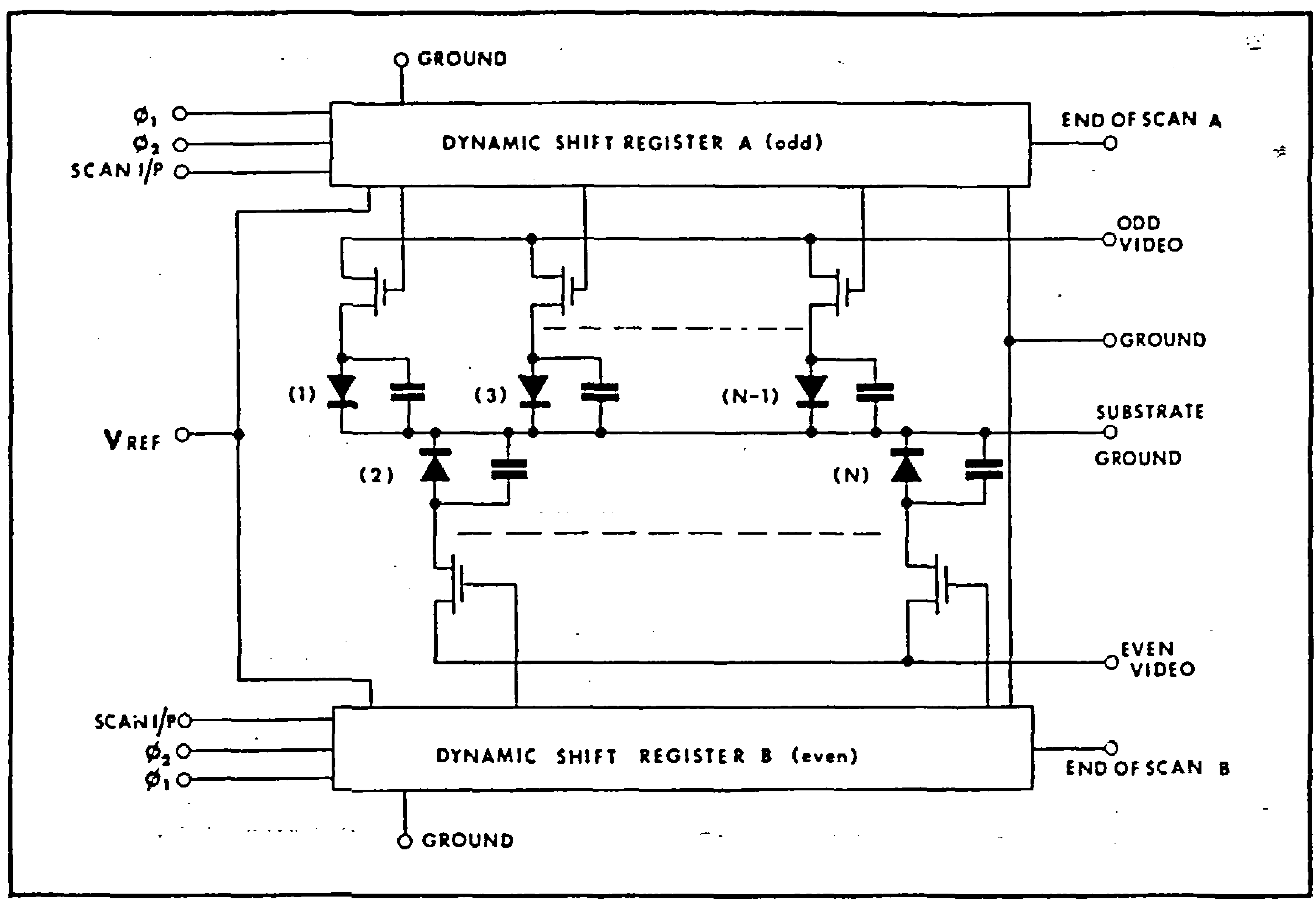


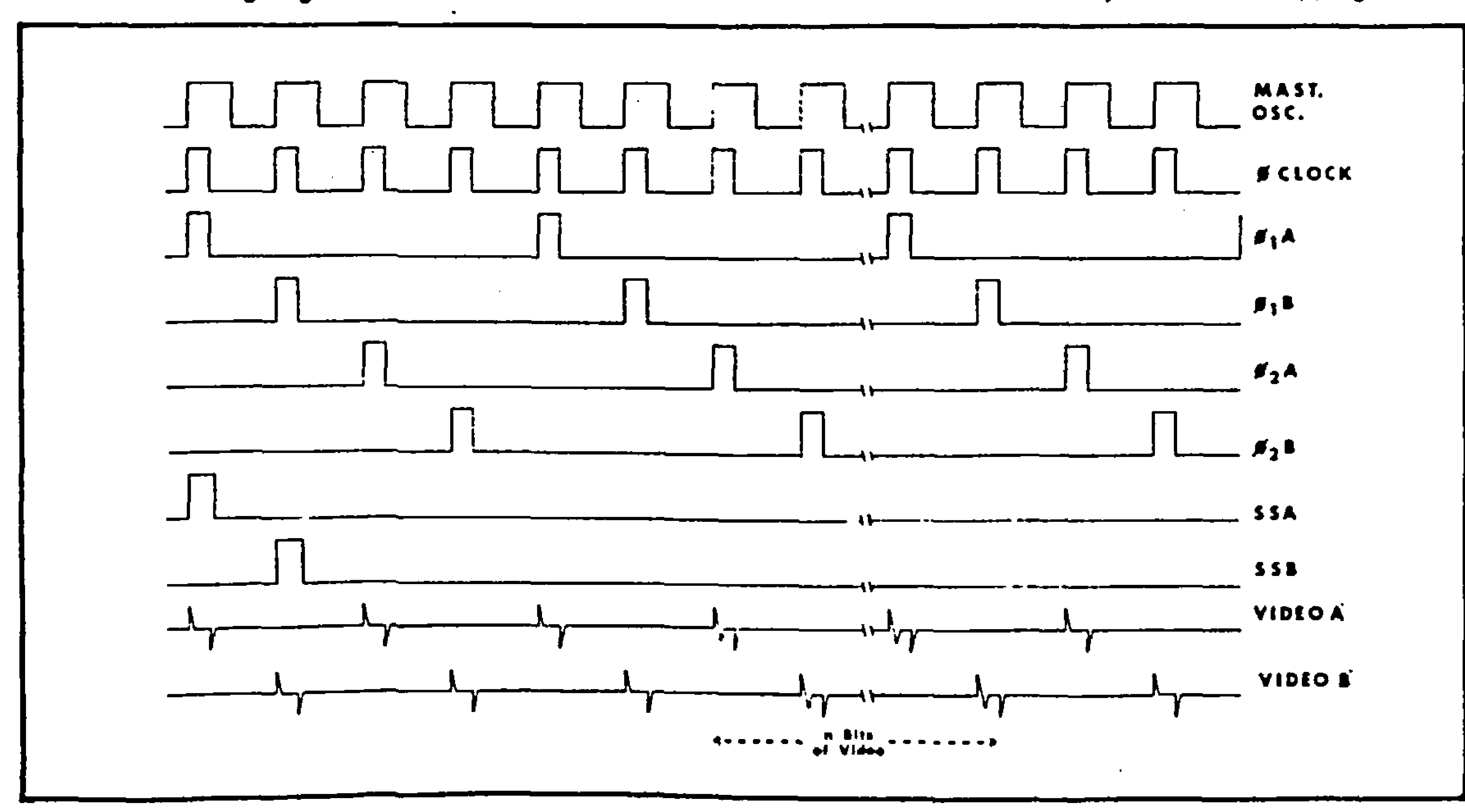
FIG 1

A simplified circuit diagram of the M-series array is shown in Fig. 1.

The Shift Register

The Photodiodes with their associated parallel storage capacitors are connected through MOS transistor switches to a common video output line. The switches are turned on and off in sequence by a shift register. Each device contains two shift registers, each register accessing alternate diodes. Each shift register is driven by two non-overlapping clock

FIG 2 TTL Timing Logic



pulse trains Ø1 and Ø2. The array scan is initiated when a scan start pulse is applied at the scan start input terminal The scan pulse is propogated through the register alternatively by Ø1 and Ø2 A TTL timing diagram of the clock pulses is shown in Fig. 2.

The shift registers can be operated in several different ways. By clocking the two registers alternately and connecting the video outputs in parallel a serial stream of video information is obtained representing every diode in the array. Alternatively, by clocking the two registers in parallel, and processing the separate video outputs each video output will represent the odd and even diodes respectively. A third alternative is to operate the registers simultaneously and connect the video outputs in parallel. This gives the effect of a .002 inch pitch array with half the total number of diodes being accessed. A reference voltage termed V.ref. must be applied to the shift register at all times in order to operate it. This voltage is nominally - 8.5V in value.

When all the photodiodes in the array have been sampled, an output pulse appears at the end of the scan terminal which is provided for each shift register. The end of scan pulse appears two clock periods after the final diode is sampled. The pulses are referenced to the end of scan ground terminal and this needs to be connected to ground. A buffer circuit for the end of scan is shown in Fig.3.

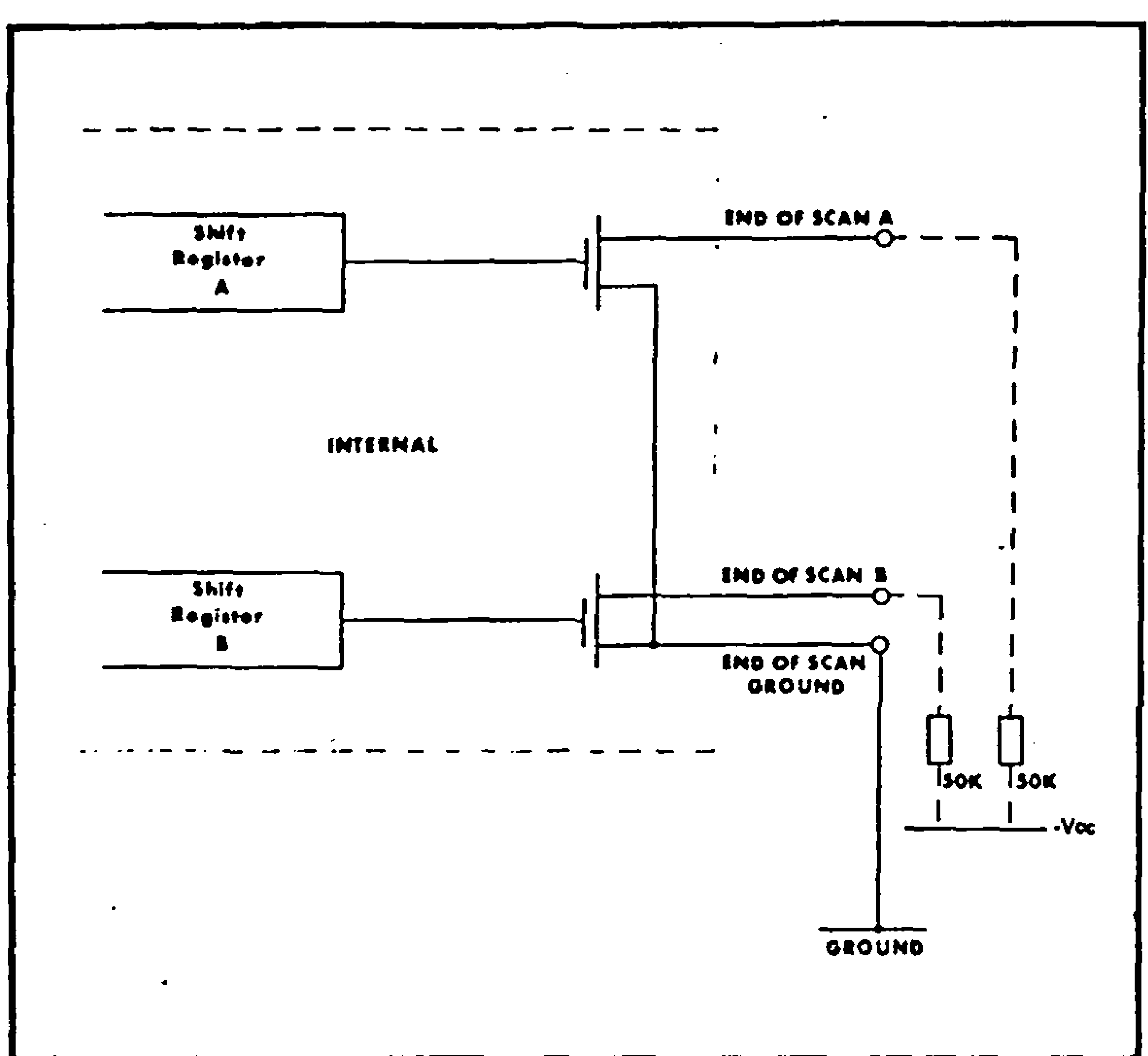
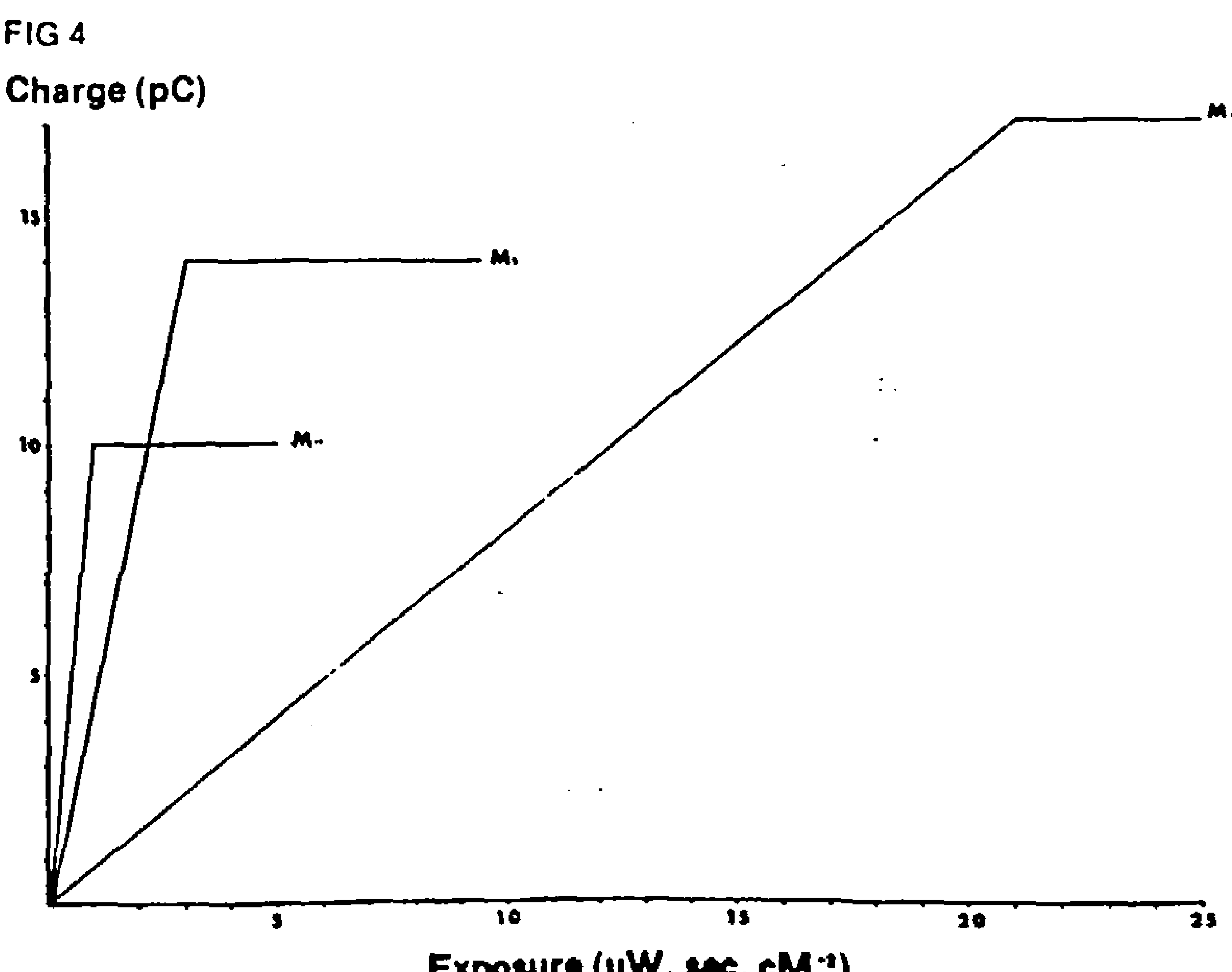


FIG 3

The Photodiode Array

FIG 4



The Photodiodes in the M array operate in a reverse bias, light integration mode. In this mode, the initial scan pulse propagated through the shift register causes each photodiode in turn to be charged to the negative potential applied to the video output terminal. During the period before the subsequent scan pulse, termed the integration time, the photodiode loses an amount of charge proportional to the total amount of light incident upon it. The subsequent scan pulse recharges the photodiode to the video output terminal potential. The amount of charge necessary to restore the photodiode to this potential represents the video signal. Hence a readout of charge proportional to light intensity is obtained.

The output charge is directly proportional to the exposure, exposure being defined as the light intensity multiplied by the integration time. A graph of the output charge against exposure is shown in Fig.4.

Video Processing

The video output of the M array is a serial train of charge pulses. to process these charge pulses IPL recommend the use of an integrator with sample and hold circuits to produce a d.c. referenced boxcar output waveform. IPL produce circuit boards to drive and process the output signal from the M arrays. The boards are sold as an assembly called the MPDA.

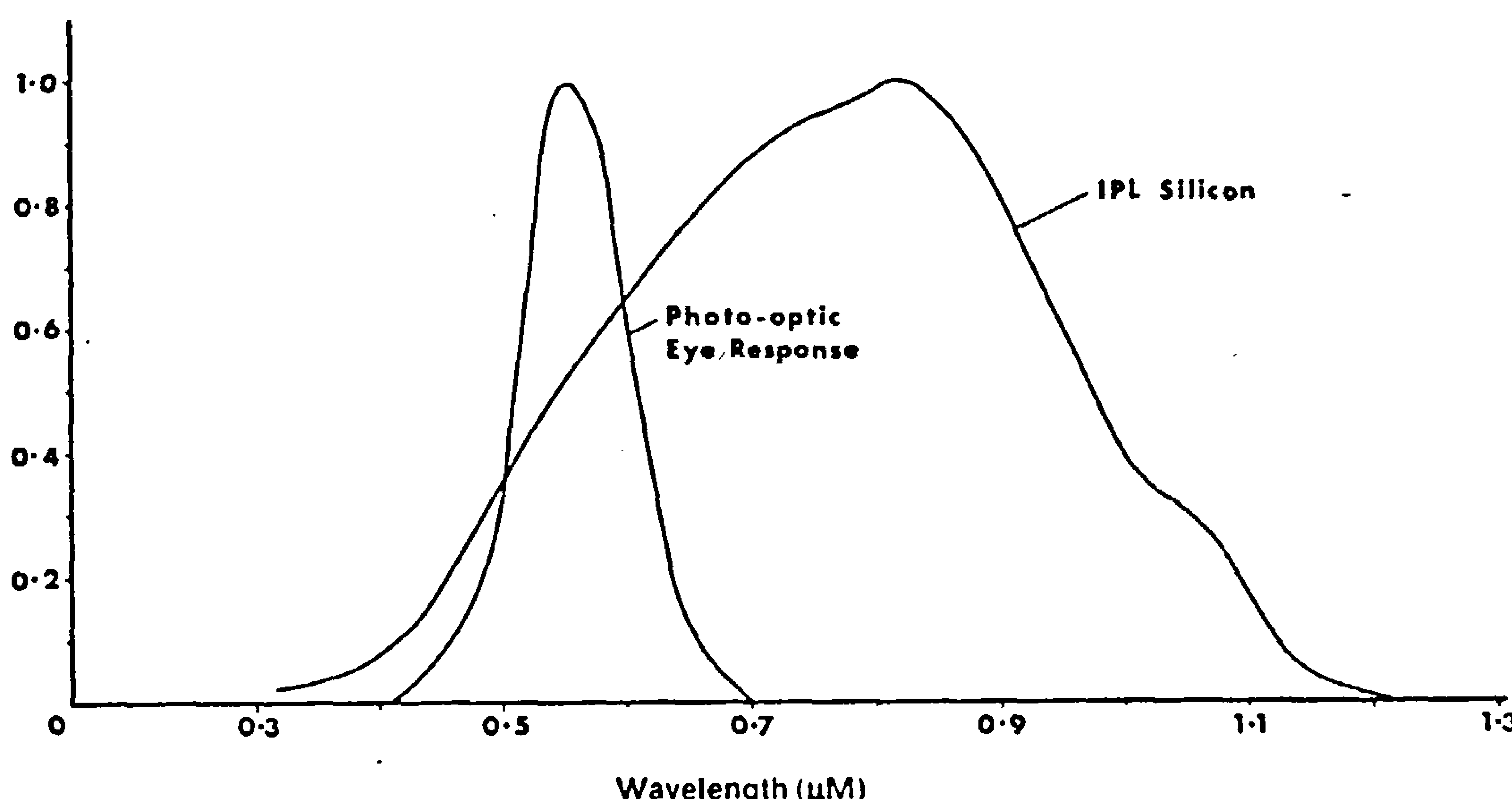
The video processor used in the MPDA allows optimum performance from the array to be obtained in terms of the signal to noise ratio.

Spectral Response and Responsitivity

The spectral response of the M arrays is shown in Fig.5. The response extends from the near ultra-violet to the infra-red region of the electro-magnetic spectrum. The peak response occurs near 820nm.

The responsivity figures given in the specification section are mean values. There is some variation of responsivity along the length of the array caused by the manufacturing processes used to fabricate the array. However, these uniformity variations also depend upon the particular light source used with the array. Tungsten light sources contain a high proportion of infra-red light to which the array is very sensitive. Uniformity variations are proportionally worse at infra-red wavelengths than visible. If light at the visible wavelength only is used, the responsivity variations are approx. 50% of the figures quoted in the technical specification.

FIG 5



Signal to Noise Considerations

The maximum signal to noise ratio that can be achieved with the M arrays depends on the output signal processing used. Using the MPDA system, a signal to noise ratio of 400 is typically obtained with a 512M array. This figure is the ratio of peak output signal level to r.m.s. noise, including Fixed-pattern noise.

This noise contains three components. Dark leakage current non-uniformity, fixed pattern noise and thermodynamic noise. Dark leakage current is temperature dependant and only becomes significant at either temperatures above 30°C or integration times in excess of 40ms. Fixed pattern noise is dependant on the shift register drive waveforms, and can be cancelled in the MPDA. The last component, thermodynamic noise, is the random non-repetitive fluctuation which is superimposed on the dark signal. This cannot be removed by the signal processing; however, it is negligible compared with the noise in the signal processing circuitry for all normal applications.

IPL reserves the right to change the products and circuits used in this information sheet in the interests of improved specifications, etc. Methods used are typical at the time of going to print but are not necessarily those most recommended.
No responsibility is assumed for the use of information contained herein, nor for any infringement of patents or rights of others which may result from such use.
No licence is granted by implication or otherwise under any patent or patent right of Integrated Photomatrix or others.

Mechanical Details and Pin Connections

Pin Connections

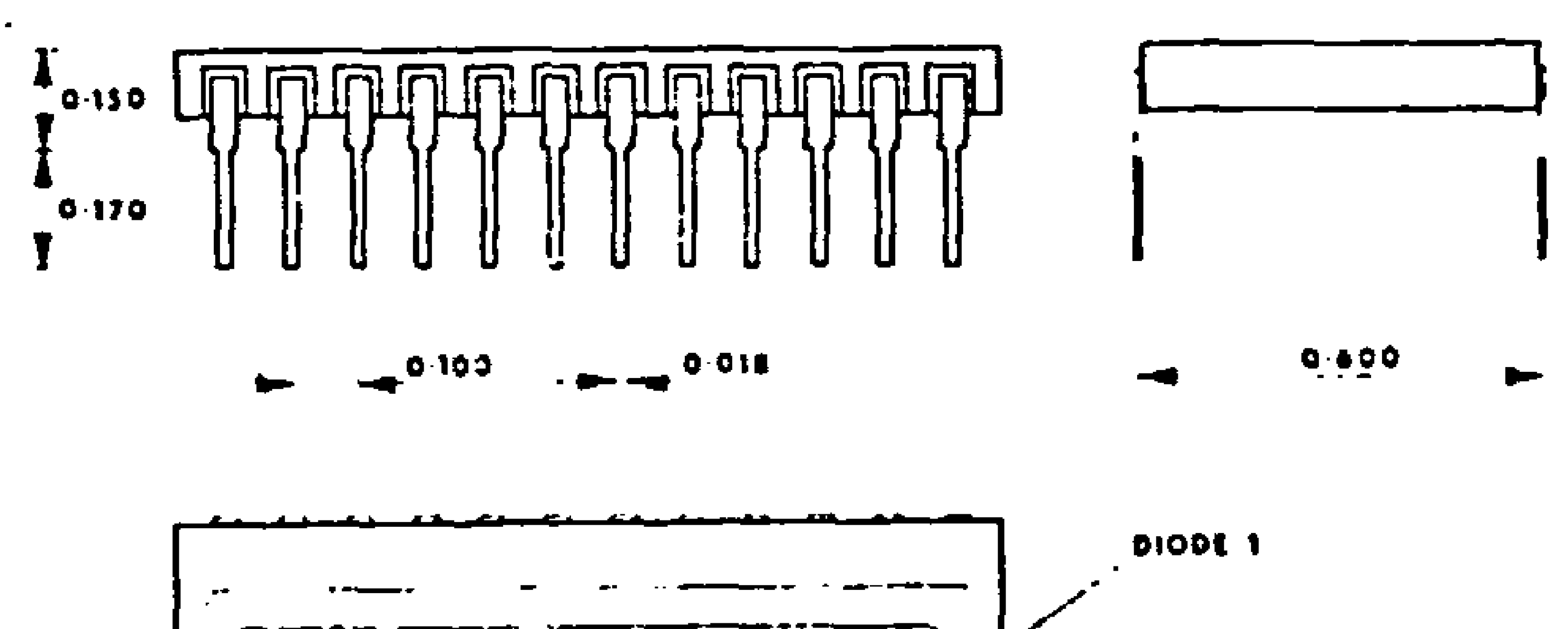
256M 512M 16 lead DIL

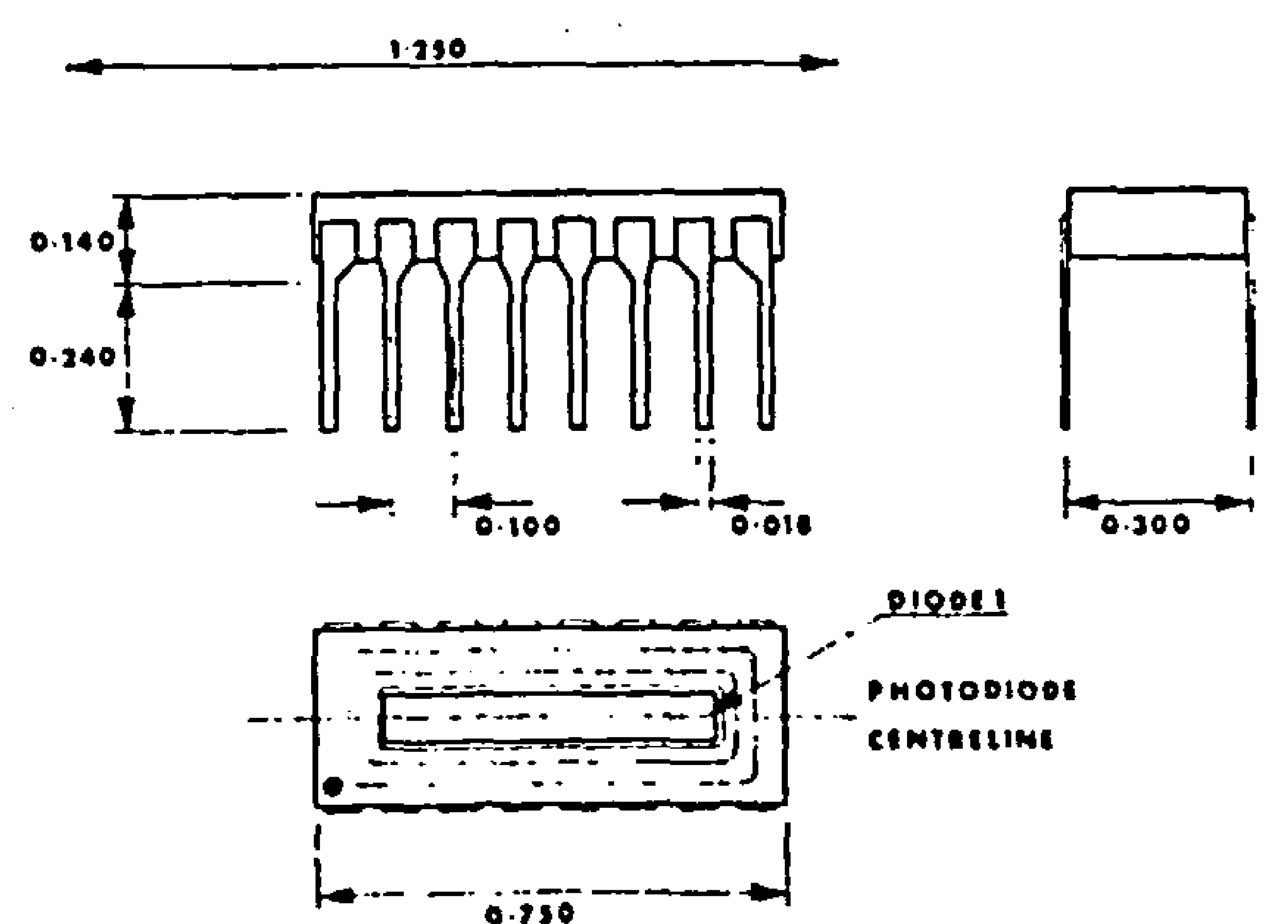
Pin No	Function
1	End of Scan A
2	Screen
3	V ref (both sides)
4	Ø2 A
5	Ø 1A
6	Scan start A
7	Video A
8	Screen and substrate
9	Screen and substrate
10	Video B
11	Scan start B
12	Ø 1B
13	Ø 2B
14	Shift register ground (both sides)
15	End of scan B
16	End of scan ground

1024M 24 lead DIL

Pin No	Function
1	End of scan A
2	Screen
3	V ref (both sides)
4	N/C
5	N/C
6	Ø2 A
7	Ø1 A
8	N/C
9	Scan start A
10	Shift reg. ground (both sides)
11	Video A
12	Screen and substrate
13	Screen and substrate
14	V ref (both sides)
15	Video B
16	Shift reg. ground (both sides)
17	Scan start B
18	N/C
19	Ø 1B
20	Ø 2B
21	N/C
22	Shift reg. ground (both sides)
23	End of scan B
24	End of scan ground

Package Details





Notes

1 Photodiode array centre line falls on the centre line of the package to a tolerance of ± 0.010"

2 Chip surface to package surface 0.051".

3 The refractive index of the package lid is 1.53

4 All the arrays scan towards Pin 1.

5 All dimensions are in inches.

Technical Specification

Performance measured at 20°C. Electro-optical parameters obtained with a tungsten filament source at 2870°K

Electrical Characteristics

Parameter	Value Min	Value Typ	Value Max	Unit
Min. shift register operating frequency	.	100	.	Hz
Max shift register operating frequency	.	5×10^{6}	.	Hz
Minimum clock width	.	200	.	ns
Clock amplitude	-23	-24	-25	V
Scan start pulse amplitude	-23	-24	-25	V
Clock crossover voltage	.	.	-2	V
Scan pulse overlap on trailing edge of Ø1	30	.	.	ns
V ref (1µA max current)	.	-8.5	.	V
Video line bias	.	-10	.	V
Clock line capacitance				
1024	.	170	.	pF
512	.	90	.	pF
256	.	50	.	pF
Video line capacitance				
1024	.	60	.	pF
512	.	31	.	pF
256	.	17	.	pF

Electro-Optical Characteristics

Parameter	Device Type M1	Device Type M5	Unit
Responsivity	0.8	4.5	pA/µW/cm² (note 1)
Saturation exposure	21	3	µW sec/cm²
Saturation charge	17	14	pC
Uniformity of response	±8	+9	% of signal (note 1)
Dark fixed pattern noise	0.15	0.15	pC peak-peak (note 2)
Dark current equivalent	8	1.4	µW/cm² at 20°C (notes 1, 2)
Centre to centre spacing	1 x 10 [illegible]	1 x 10 [illegible]	inches
Aperture width	1 x 10 [illegible]	5 x 10 [illegible]	inches (see note 3)

Absolute Maximum Ratings

	Min	Max	
Voltage applied to any pin with respect to substrate	+0.2	-30	Volts
Ambient operating temperature	-10	+70	°C
Storage temperature	-20	+85	°C

Notes

1 *Using a 2870K tungsten light source.*

2 A + 2V bias applied to pins 2, 8 and 9 (256M and 512M) or pins 2, 12, and 13 (1024M) with respect to the shift register ground pin reduces these figures to approximately 60%.

3 The wide aperture array is specified by the suffix after the letter M. The 256 M5 is the 256 element array with the 0.005" aperture.

PX303A/1282

Printed in Gt. Britain

INTEGRATED PHOTOMATRIX LIMITED

The Grove Trading Estate
Dorchester, Dorset,
DT1 1SY, U.K.
Phone 0305 63673
Telex 41166

The M Series Processor Driver Assembly
M PDA

Features

★ Complete System to Operate the IPL M Series of Arrays

★ Operating Frequency Range from 2 KHz to 2 MHz

★ Video Output. 'Boxcar' Waveform

Introduction

The M PDA has all the circuitry necessary to operate the IPL M Series of arrays, and is comprised of three printed circuit boards. The two main p.c. boards contain the logic circuitry and the video processing circuitry respectively. The smallest p.c. board is the array carrier board, which contains high level clock drivers and the video pre-amplifier. The logic board and the processor board are connected to each other by a fixed DIL plug and socket. The array carrier board is connected to the main p.c. boards by means of a flexible ribbon cable. This enables the user to operate the array in different mechanical configurations.

INTEGRATED PHOTOMATRIX LTD.

The Logic Board

The logic board uses CMOS logic and provides the four phase clocks and two scan pulses necessary to operate the IPL M Series arrays. The input voltage requirements are ± 15V typical. The actual board voltages are provided by on board regulators.

An oscillator located on the board provides the main system clock. The oscillator's frequency can be selected over four ranges from 1 KHz to 2 MHz by a DIL programming switch and precise adjustment is provided by a fine frequency potentiometer. An external oscillator can be used when the INT/EXT frequency link is removed. The external signal must be at a CMOS level (12V maximum).

The four phase clock pulses are derived from a fixed width clock pulse which operates a 4 bit shift register. The two scan start pulses are derived from programmable counters. In this way all the different array lengths can be used simply by re-programming the counters. Three DIL programmable switches located along the edge of the board control the counters and give a maximum count of 4096 bits between scan pulses.

Two potentiometers are located on the logic board. One is the fine frequency control and the other is the integrator reset control. The four phase clocks and scan pulses are accessible by means of signal test points located on the end of the board.

The clock waveforms used to operate the array are shown in the M Series array data sheet.

The logic p.c. board connects to the processor board by means of a DIL plug and socket. The separation of the two boards is maintained by spacers located at each of the four corners. Spacers are also provided to mount the array carrier board onto the logic board.

The Processor Board

The IPL M Series of arrays are operated in the Recharge Mode. In this mode of operation the output from the array is a serial train of current pulses whose amplitudes are proportional to the light intensity profile across the photodiodes. To transform these current pulses to a boxcar waveform an integrator with Sample and Hold circuits is used. A block diagram of the processor and logic boards is illustrated in Fig.1.

The clock pulses used to drive the array are of a fixed width, irrespective of the operating frequency. In this way the integrator period is fixed, and always includes the clock pulse edge disturbances. By including both of the disturbing edges in the integrator period it is possible to cancel any clock disturbances present in the video information.

A fine control for balancing all four clock disturbances is provided by four potentiometers situated on the end of the processor board. These are provided for precise applications where the minimum clock disturbance is required. Two other potentiometers are provided for the D.C. offset and the video boxcar amplitude. A D.C. offset potentiometer is also provided on the array carrier board. Since, in certain circumstances, it may be impossible to gain access to the head card, when it is mounted in a system, a duplicate is provided on the processor card.

A typical output waveform using the M PDA and an M512 array is illustrated in Fig.2.

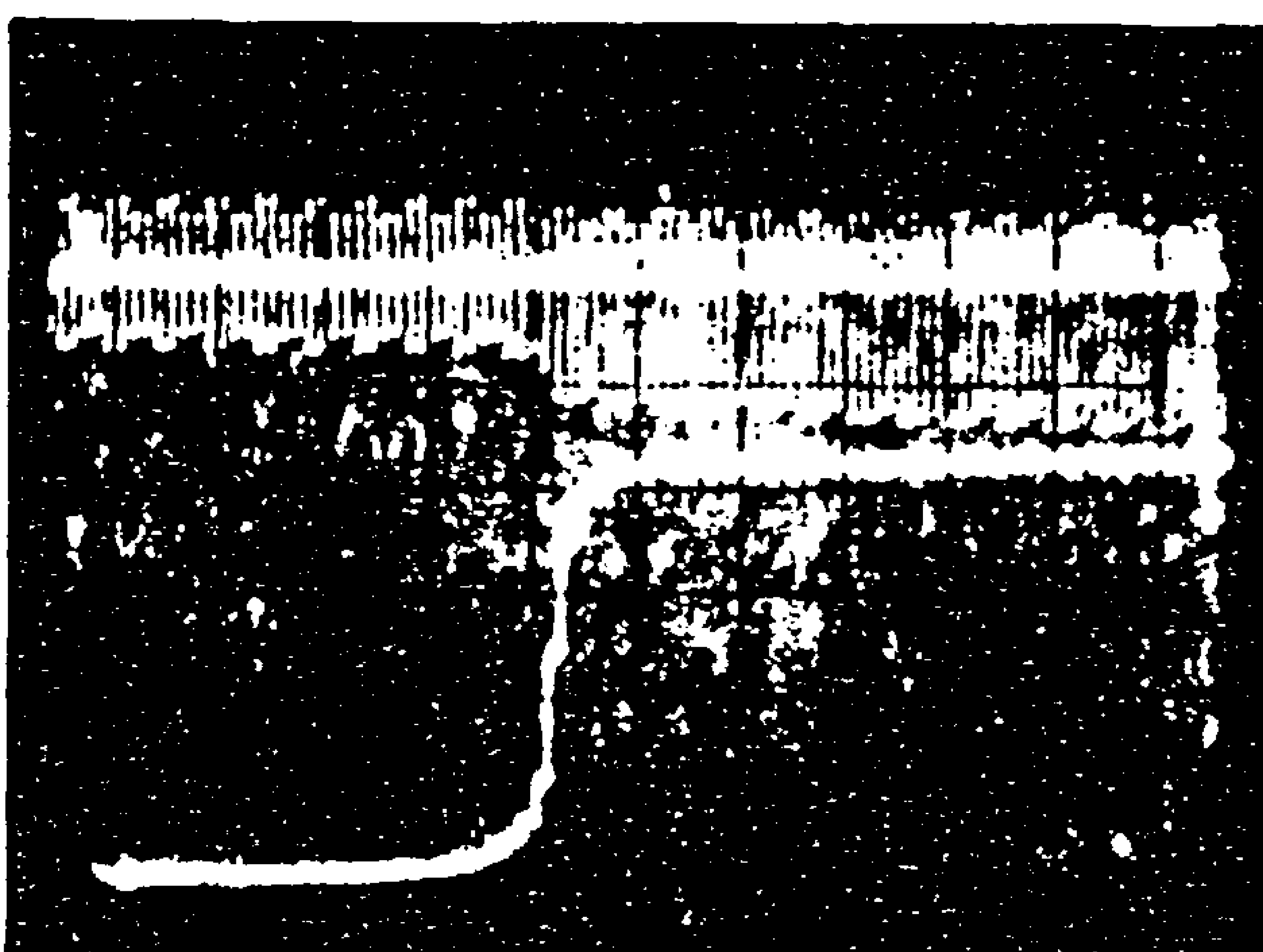


FIG 2

Output from a 512 array top trace output from pre-amplifier bottom trace output from the final stage

The Head Board

The head board contains a DIL socket in which the array is mounted. In addition the video pre-amplifier and array high level clock drivers are also included on the board. The potentiometer located at the end of the board is the D.C. offset control. This enables the array output signal in black conditions to be adjusted, so that it is co-incident with OV.

The board connects to the main logic and processor boards by means of a flexible ribbon cable. This allows the user to mount the array in several different positions, two of which are depicted in Fig.3.

It should be noted that when purchasing the M PDA system, the array is not included. *It must be purchased as a separate item.*

PRE-AMP

ARRAY

MOS CLOCK DRIVERS

ϕ_1 A
ϕ_1 B
ϕ_2 A
ϕ_2 B

HEAD BOARD (PC210-01)

SCAN A

SCAN B

MONO 3

ϕ CLOCK

SCAN COUNT-DOWN

LEVEL SHIFT

INT. EXT.

OSC 2MHz -1KHz

MONO 1

MONO 2

4 BIT SHIFT REG.

DRIVER BOARD (PC210-02)

INT. RESET

SAMPLE 1

SAMPLE 2

PICK-UP CONTROL

INTEGRATOR

SAMPLE & HOLD 1

SAMPLE & HOLD 2

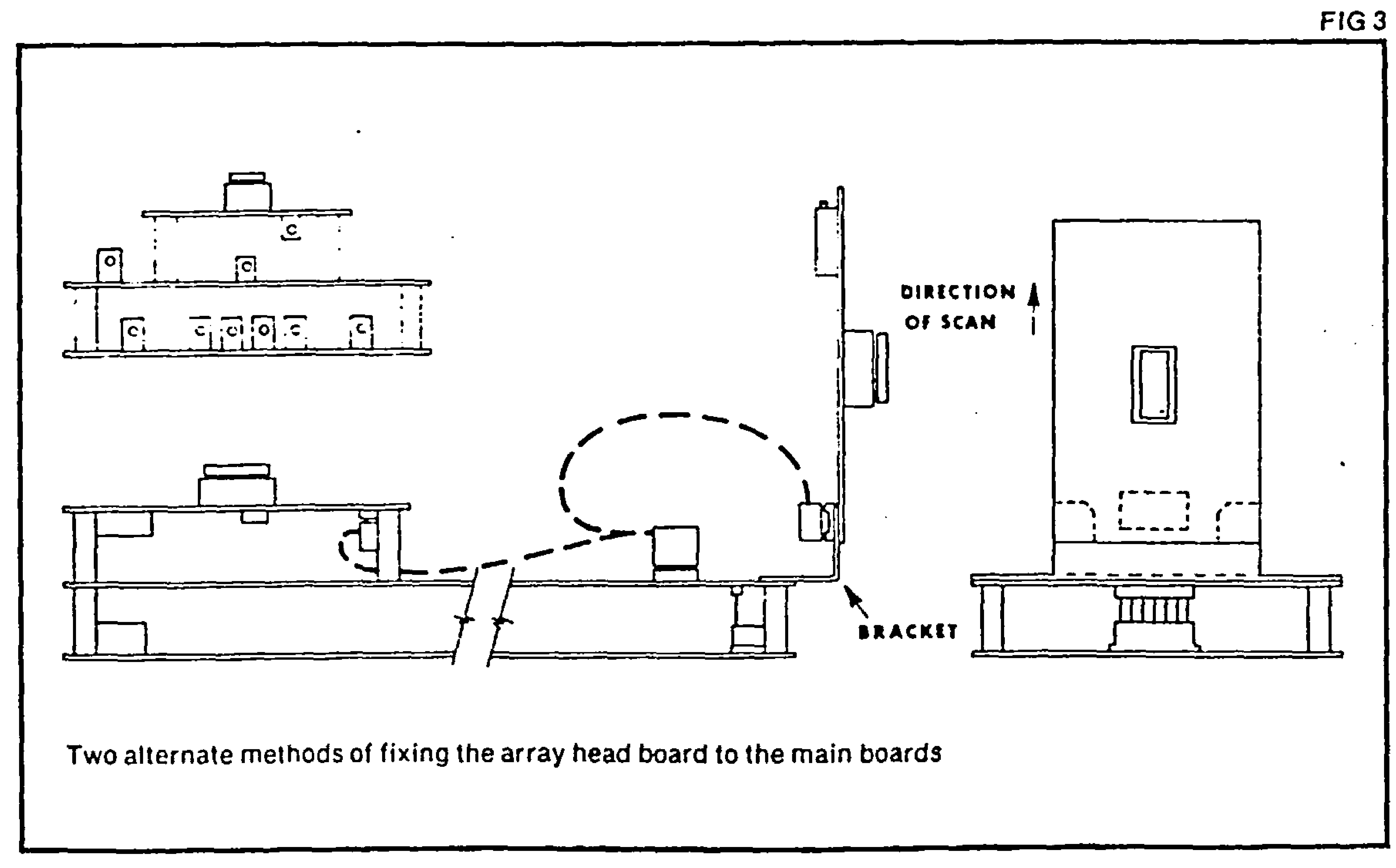
BUFFER

VIDEO O/P

PROCESSOR BOARD (PC210-03)

FIG 1

FIG 3



Two alternate methods of fixing the array head board to the main boards

Technical Specification

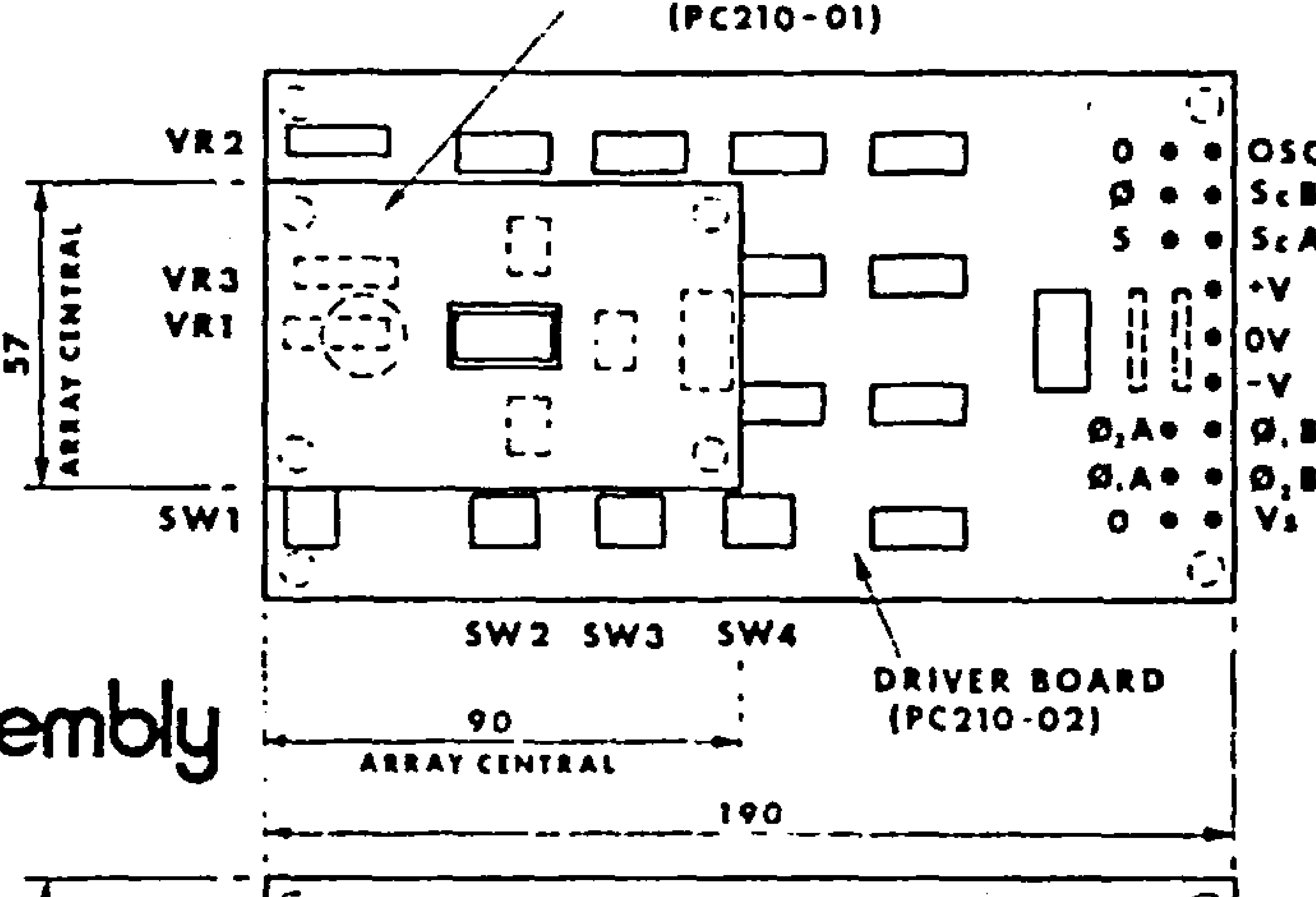
Typical performance at 22°C ± 2°C, with any M series array using a tungsten filament source at 2850°K.

Parameter	Value Min.	Value Typ.	Value Max	Unit
Operating Frequency	1×10^3	10	2×10^6	Hz
External Trigger input	9.0		12	Volts
Scan Start Countdown	1		4096	Clock periods between scan start pulses
Video Output	1	5	10	Volts (adjustable)
Signal to Noise Ratio		55		Decibels Sat Signal/ RMS Noise
Zero Adjustment Range	3		+3	Volts at max. output
Video Output Impedance		100		ohms
Supply Voltage	±14.5	± 15	± 18	volts
Supply Current			+160	mA
			-140	mA
Operating Temperature	0	20	70	°C

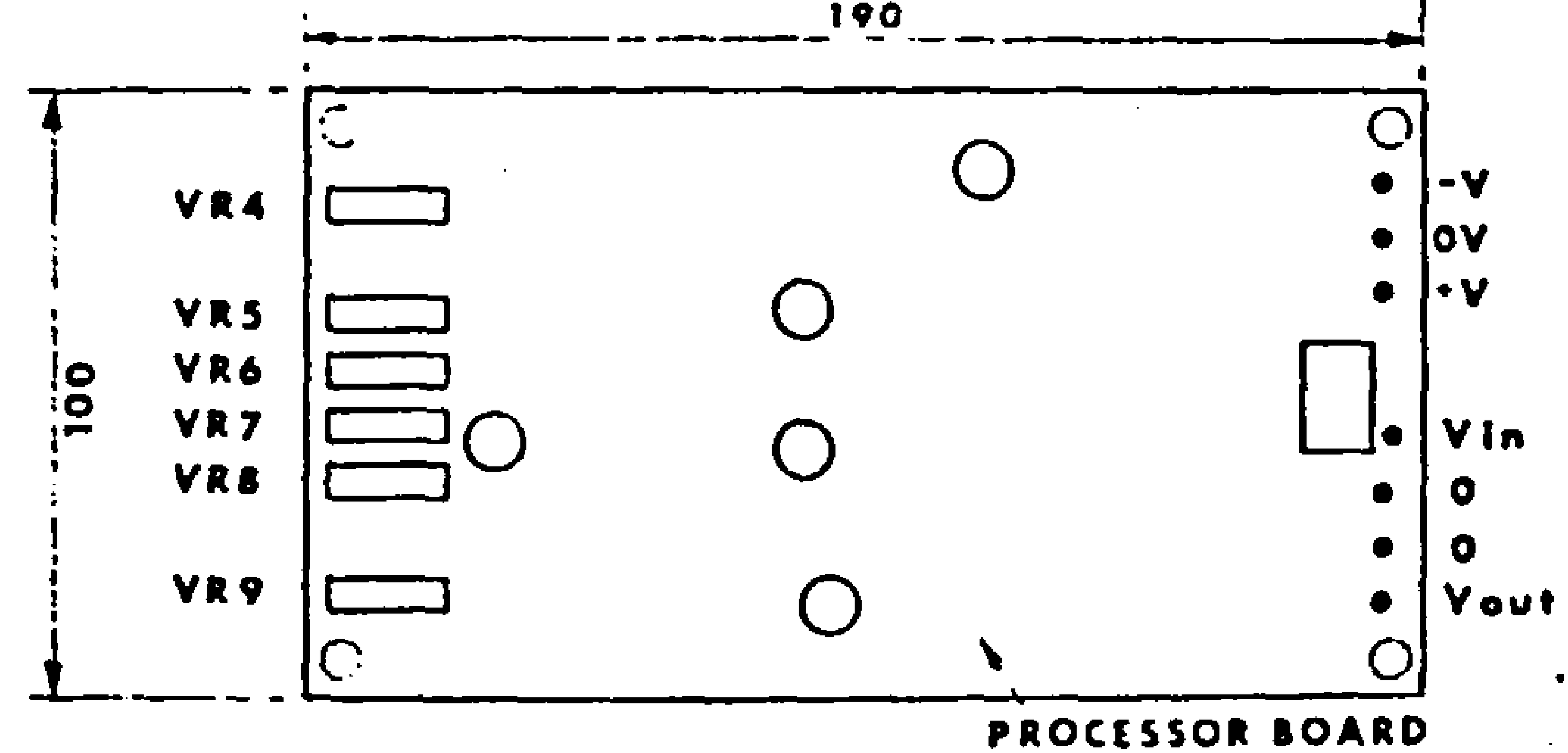
WIRING

The headcard is normally supplied with a 160 mm length of flexible ribbon cable connecting the head card to the logic board. Longer lengths can be supplied on request up to a maximum of 400 mm.

Full operating instructions are supplied with every system.



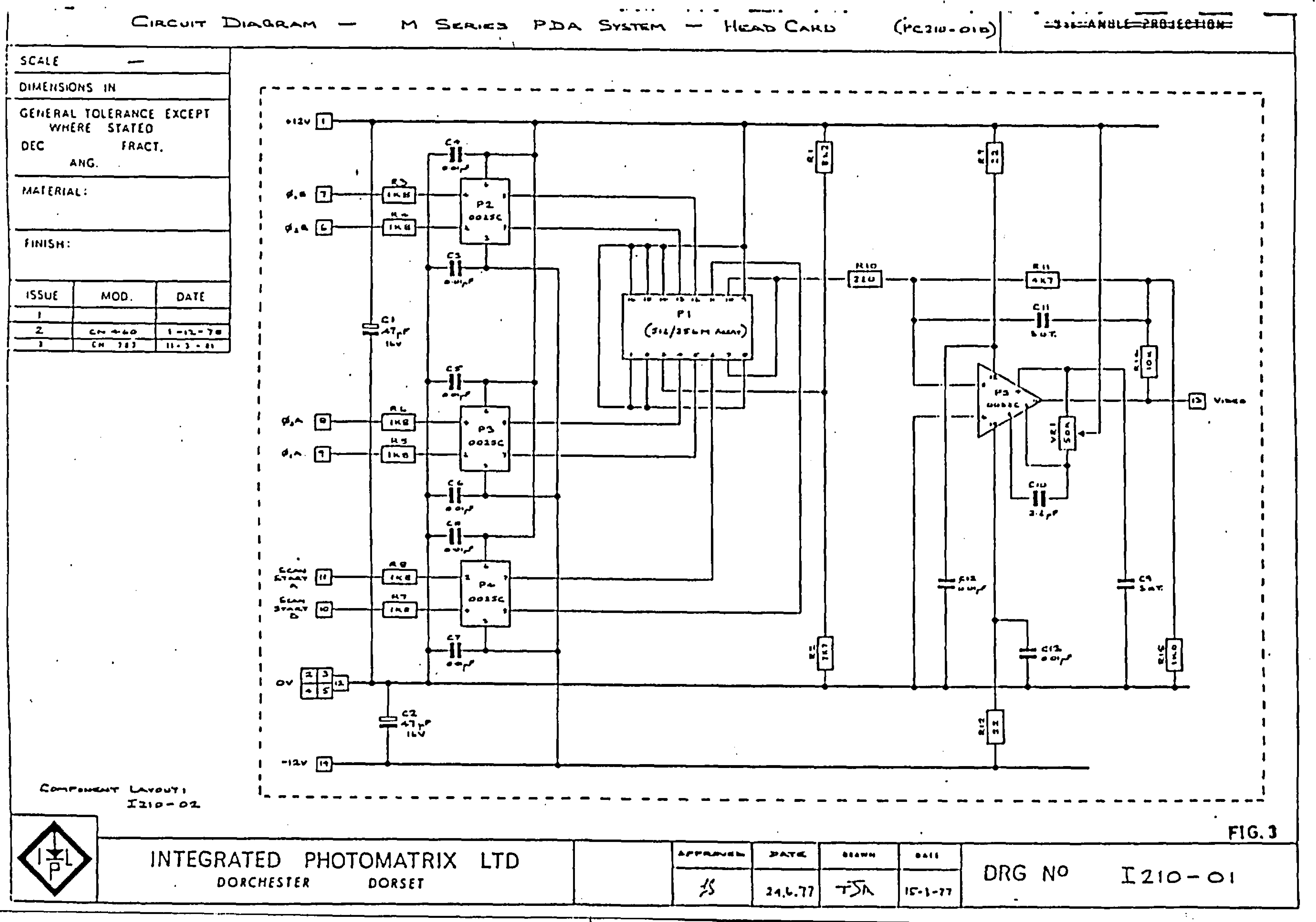
General Assembly



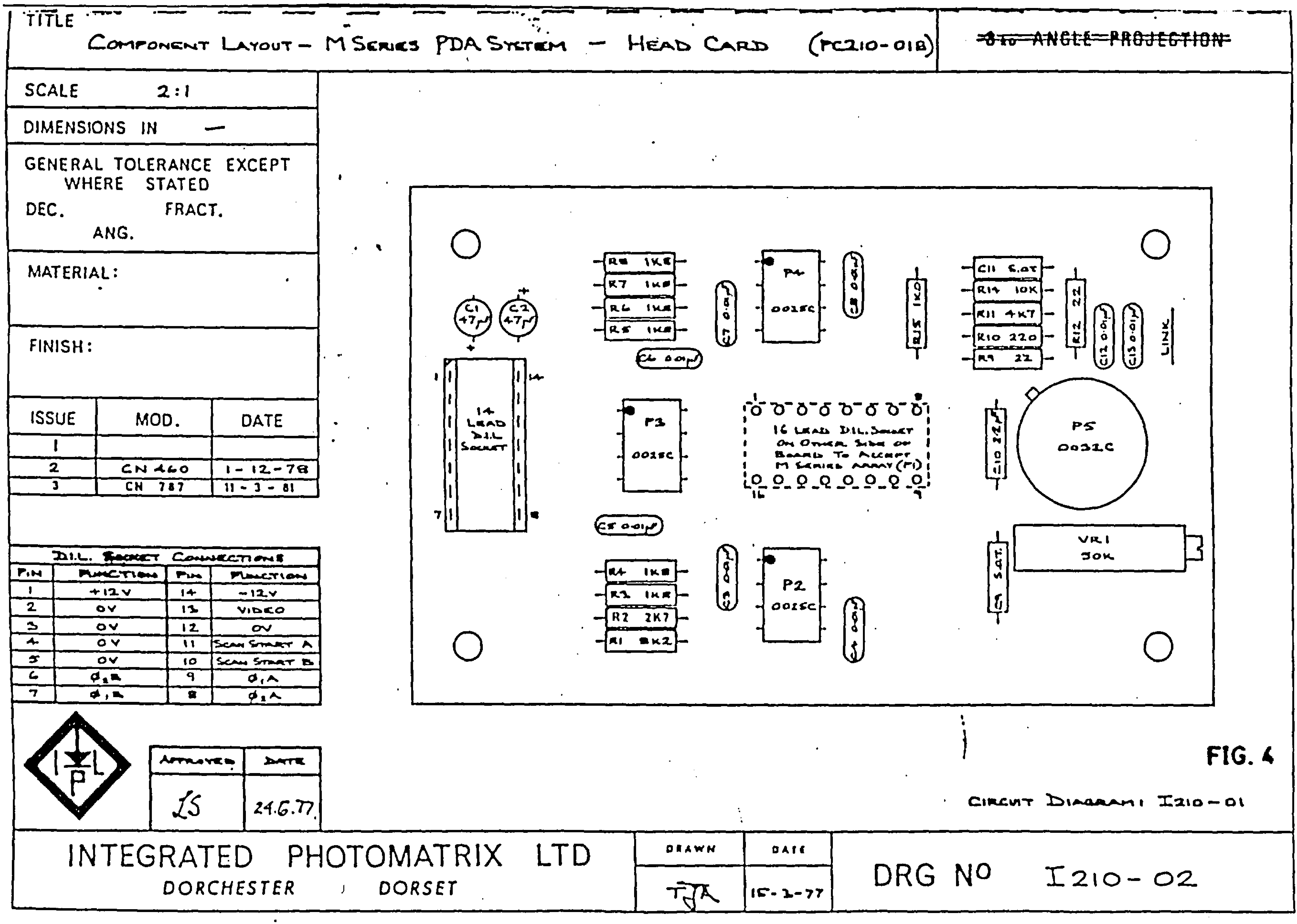
IPL reserves the right to change the products and circuits used in this information sheet in the interests of improved specifications etc
Methods used are typical at the time of going to print but are not necessarily those most recommended
No responsibility is assumed for the use of information contained herein, nor for any infringement of patents or rights of others which may result from such use
No licence is granted by implication or otherwise under any patent or patent right of Integrated Photomatrix or others

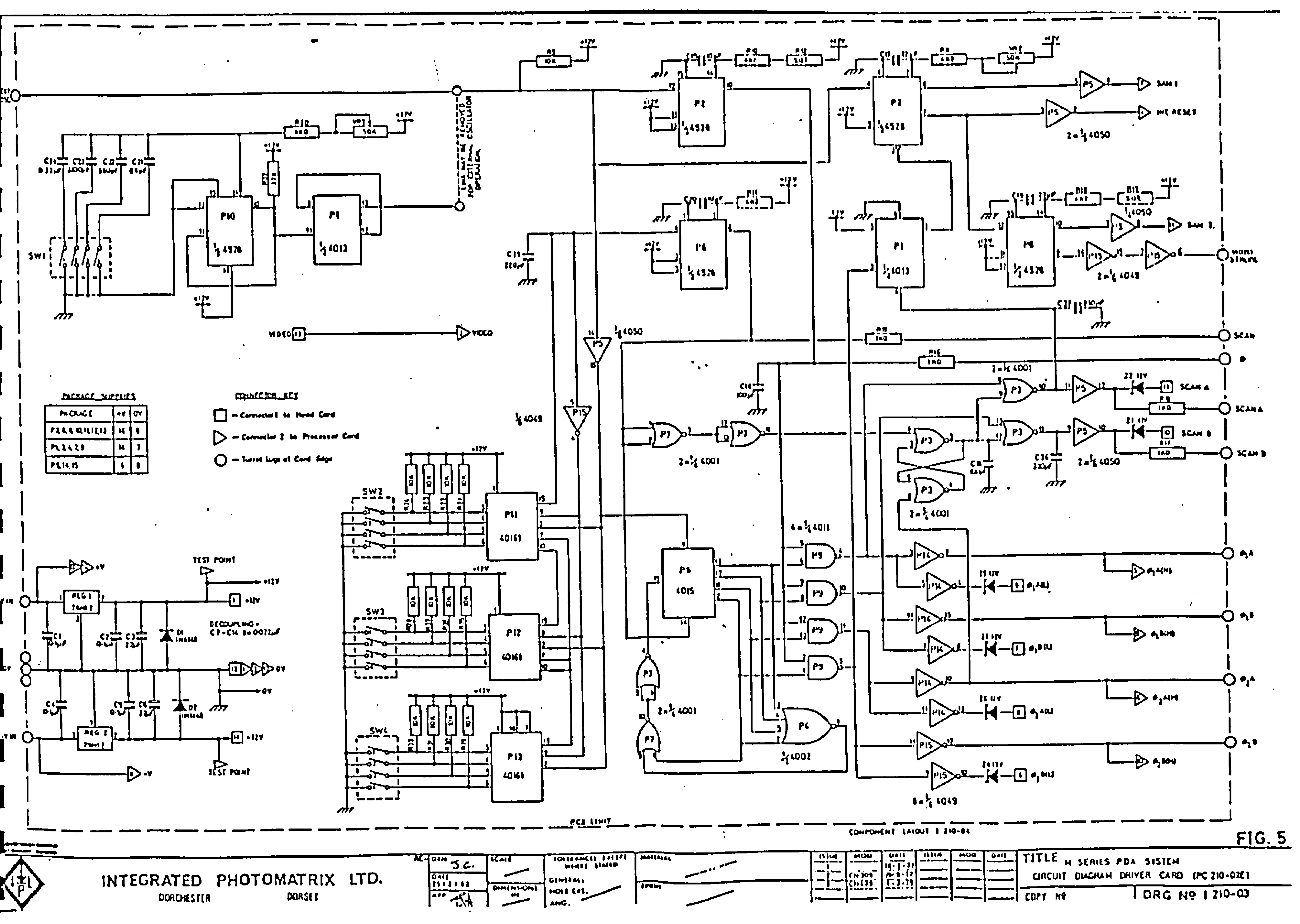
May 1977 No 30

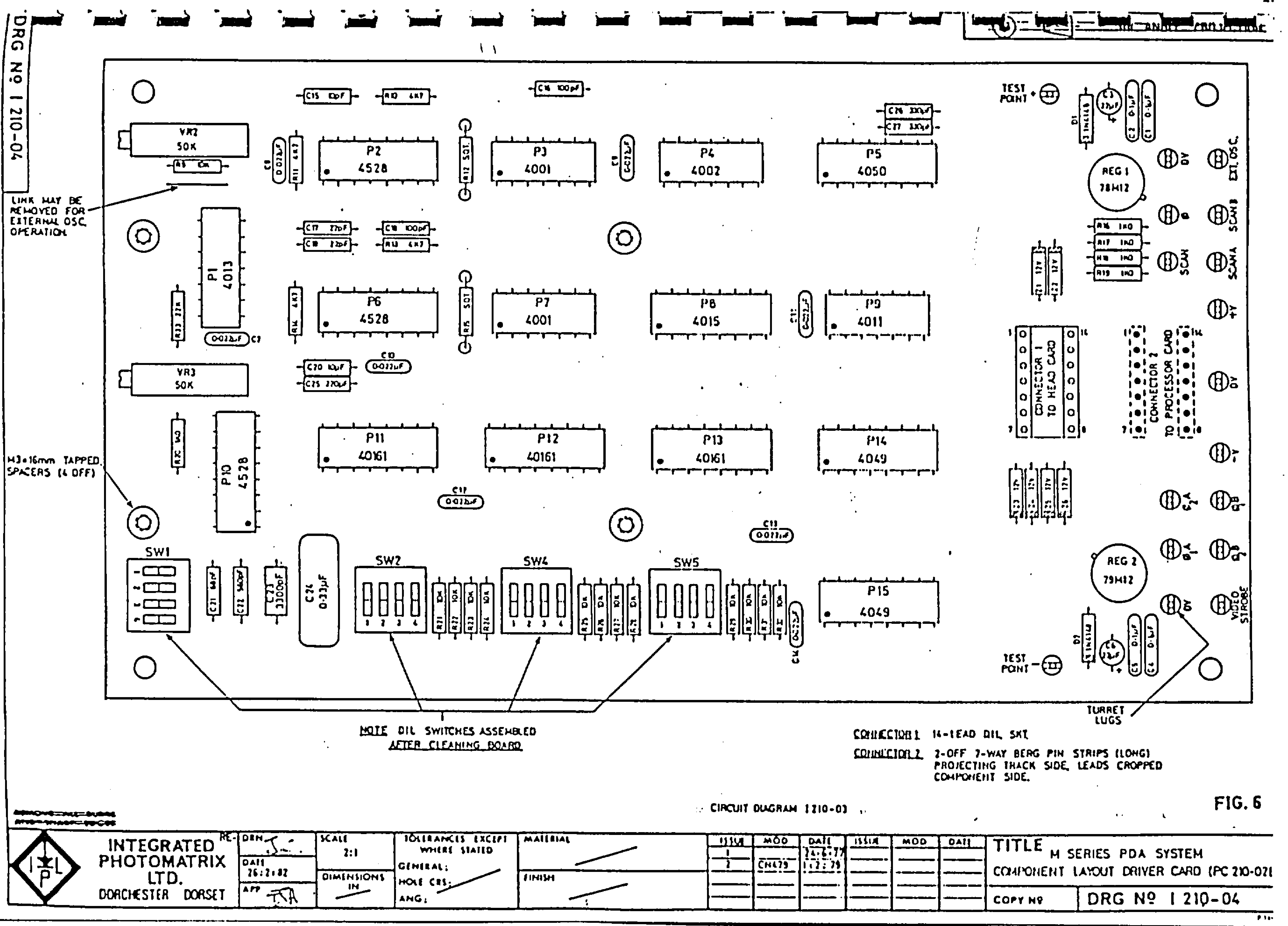
INTEGRATED PHOTOMATRIX LIMITED
The Grove Trading Estate
Dorchester, Dorset,
DT1 1SY, U.K.
Phone 0305 3673
Telex 41166

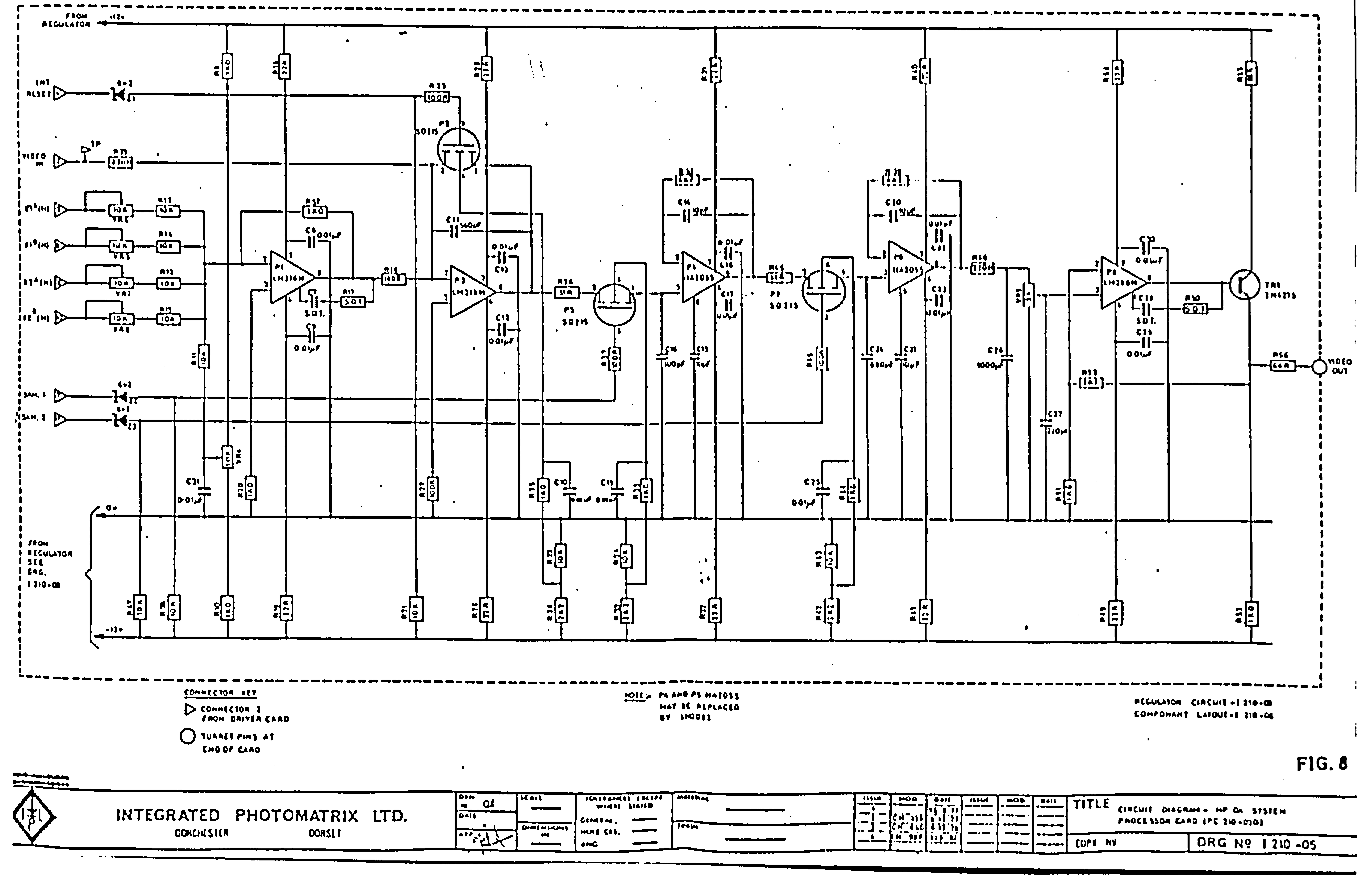


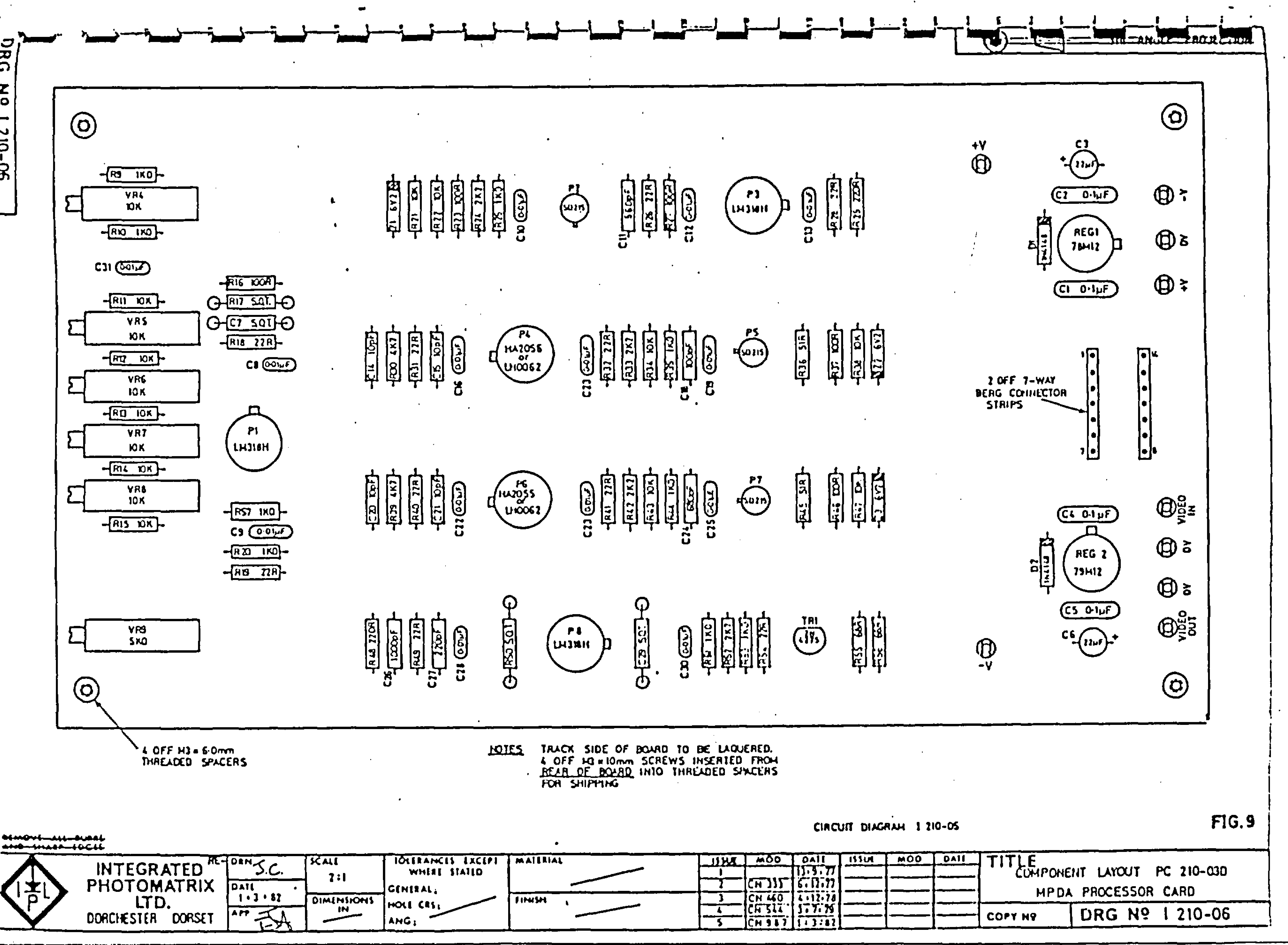
CIRCUIT DIAGRAM — M SERIES PDA SYSTEM — HEAD CARD (PC210-01D)
SCALE —
DIMENSIONS IN
GENERAL TOLERANCE EXCEPT WHERE STATED
DEC FRACT.
ANG.
MATERIAL:
FINISH:
ISSUE MOD. DATE
1
2 CN 460 1-12-78
3 CN 783 11-3-81
+12V 1
ϕ2B 7
ϕ1B 6
ϕ2A 8
ϕ1A 9
SCAN START A 11
SCAN START D 10
0V 2 3 4 5 12
-12V 14
VIDEO 13
P1 (512/256M ARRAY)
P2 0025C
P3 0025C
P4 0025C
P5
C1 47μF 16V
C2 47μF 16V
R10 220
R11 4K7
C11 S.O.T.
C10 2.2pF
VR1 50R
C9 S.O.T.
R12 2.2
C12 0.01μF
R13 10K
COMPONENT LAYOUT: I210-02
FIG. 3
INTEGRATED PHOTOMATRIX LTD
DORCHESTER DORSET
APPROVED DATE DRAWN DATE
24.6.77 15-3-77
DRG No I210-01


TITLE
COMPONENT LAYOUT – M SERIES PDA SYSTEM – HEAD CARD (PC210-01B)
ANGLE PROJECTION
SCALE 2:1
DIMENSIONS IN —
GENERAL TOLERANCE EXCEPT WHERE STATED
DEC. FRACT.
ANG.
MATERIAL:
FINISH:
ISSUE MOD. DATE
1
2 CN 460 1-12-78
3 CN 787 11-3-81
D.I.L. SOCKET CONNECTIONS
PIN FUNCTION PIN FUNCTION
1 +12V 14 -12V
2 0V 13 VIDEO
3 0V 12 0V
4 0V 11 SCAN START A
5 0V 10 SCAN START B
6 φ1B 9 φ1A
7 φ1B 8 φ1A
APPROVED DATE
LS 24.6.77
INTEGRATED PHOTOMATRIX LTD
DORCHESTER, DORSET
DRAWN DATE
TJA 15-2-77
DRG NO I210-02
FIG. 4
CIRCUIT DIAGRAM: I210-01
14 LEAD D.I.L. SOCKET
16 LEAD D.I.L. SOCKET ON OTHER SIDE OF BOARD TO ACCEPT M SERIES ARRAY (P1)
C1 47μ
C2 47μ
R8 1K8
R7 1K8
R6 1K8
R5 1K8
C6 0.01μ
P3 0025C
P4 0025C
C7 0.01μ
C8 0.01μ
R15 1K0
C11 5.6T
R14 10K
R11 4K7
R10 220
R9 22
R12 22
C12 0.01μ
C13 0.01μ
LINK
P5 0032C
C10 22μ
VR1 50K
C9 S.O.T.
C5 0.01μ
R4 1K8
R3 1K8
R2 2K7
R1 8K2
P2 0025C
C3 0.01μ
C4 0.01μ


PACKAGE SUPPLIES
CONNECTOR KEY
TEST POINT
FIG. 5
INTEGRATED PHOTOMATRIX LTD.
DORCHESTER
DORSET
TITLE H SERIES PDA SYSTEM
CIRCUIT DIAGRAM DRIVER CARD (PC 210-02E)
DRG No I 210-03


P1 4013
P2 4528
P3 4001
P4 4002
P5 4050
P6 4528
P7 4001
P8 4015
P9 4011
P10 4528
P11 40161
P12 40161
P13 40161
P14 4049
P15 4049
VR2 50K
VR3 50K
REG 1 78M12
REG 2 79M12
SW1
SW2
SW4
SW5
CONNECTOR 1 TO HEAD CARD
CONNECTOR 2 TO PROCESSOR CARD
TEST POINT +
TEST POINT −
0V
EXT. OSC.
SCAN B
SCAN A
SCAN
+V
−V
VIDEO STROBE
LINK MAY BE REMOVED FOR EXTERNAL OSC. OPERATION
M3×16mm TAPPED SPACERS (4 OFF)
NOTE DIL SWITCHES ASSEMBLED AFTER CLEANING BOARD
TURRET LUGS
CONNECTOR 1 14-LEAD DIL SKT
CONNECTOR 2 2-OFF 7-WAY BERG PIN STRIPS (LONG) PROJECTING TRACK SIDE, LEADS CROPPED COMPONENT SIDE.
CIRCUIT DIAGRAM 1210-03
FIG. 6
INTEGRATED PHOTOMATRIX LTD.
DORCHESTER DORSET
DATE 26:2:82
SCALE 2:1
DIMENSIONS IN
TOLERANCES EXCEPT WHERE STATED
GENERAL:
HOLE CRS:
ANG:
MATERIAL
FINISH
ISSUE
MOD
DATE
CN429
24:6:77
1:2:79
TITLE M SERIES PDA SYSTEM
COMPONENT LAYOUT DRIVER CARD (PC 210-021
COPY Nº
DRG Nº 1 210-04


FROM REGULATOR
VIDEO IN
VIDEO OUT
CONNECTOR KEY
CONNECTOR 2 FROM DRIVER CARD
TURRET PINS AT END OF CARD
REGULATOR CIRCUIT - I 210-08
COMPONANT LAYOUT - I 210-06
FIG. 8
INTEGRATED PHOTOMATRIX LTD.
DORCHESTER
DORSET
TITLE
DRG Nº I 210-05



DRG Nº I 210-06
+V
-V
A-
0V
A+
VIDEO IN
0V
0V
VIDEO OUT
REG1 78M12
REG 2 79M12
C1 0·1µF
C2 0·1µF
C4 0·1µF
C5 0·1µF
C3
C6
D1
D2
2 OFF 7-WAY BERG CONNECTOR STRIPS
P1 LH318H
P3 LH318H
P8 LH318H
P4 HA2056 or LH0062
P6 HA2055 or LH0062
P2
P5
P7
TR1
VR4 10K
VR5 10K
VR6 10K
VR7 10K
VR8 10K
VR9 5K0
R9 1K0
R10 1K0
R11 10K
R12 10K
R13 10K
R14 10K
R15 10K
R16 100R
R17 S.O.T.
C7 S.O.T.
R18 22R
C8 0·01µF
C31 0·01µF
R57 1K0
C9 0·01µF
R20 1K0
R19 22R
R50 S.O.T.
C29 S.O.T.
4 OFF M3 × 6·0mm THREADED SPACERS
NOTES
TRACK SIDE OF BOARD TO BE LAQUERED.
4 OFF M3 × 10mm SCREWS INSERTED FROM REAR OF BOARD INTO THREADED SPACERS FOR SHIPPING
CIRCUIT DIAGRAM I 210-05
FIG. 9
INTEGRATED PHOTOMATRIX LTD.
DORCHESTER DORSET
DRN S.C.
DATE 1·3·82
APP
SCALE 2:1
DIMENSIONS IN
TOLERANCES EXCEPT WHERE STATED
GENERAL;
HOLE CRS;
ANG;
MATERIAL
FINISH
ISSUE
MOD
DATE
1
2
3
4
5
CH 335
CH 460
CH 514
CH 987
TITLE
COMPONENT LAYOUT PC 210-030
MPDA PROCESSOR CARD
COPY Nº
DRG Nº I 210-06

Appendix 9 Listing and flowcharts of the sensor software.

The computer program is written BBC-Basic and 6502 assembly language. The pages A9.6-35 contain the assembly statements and the flowchart of each subroutine. The page A9.36 -A9.43 contain the mnemonics of the 6502 microprocessor instruction set.

Only the statements in Basic are listed in full. The position of the assembly code is indicated.

LISTING

```
10 REM SENSOR, (c) 1985 J.H.RAKELS

30 HIHEM=&498F

40 *KEY 9 "OLD ¶M 1990 VDU3 ¶M GOTO 60 ¶M "

42 *KEY 4 "OLD ¶M 1990 VDU2 ¶M GOTO 60 ¶M "

45 *KEY 2 "GOTO 1665 ¶M "

50 *KEY 1 "GOTO 1400 ¶M "

60 MODE4

65 *FX 5,1

70 VDU 24, 0;761;1200;1023

80 VDU 28,0,31,39,10

90 HIHEM=&498F:V=0

110 E0=&5300:E1=E0+256:E2=E1+256:E3=E2+256:E4=E3+256:TEMPX=E0+&F0:
    MAXA=E0+&FF:WASTE=E0+&F3:SX=E0+&FF:SY=E0+&FA:XTEMP=E0+&FD:
    MEANX=E0+&1

130 E11=E1+&1:E21=E2+&1:E12=E1+&2:E22=E2+&2;E38=E3+&8:E39=E3+&9:
```

```
    E02=E0+&2:E48=E4+&8:E49=E4+&9:E14=E1+&4:E24=E2+&4:E34=E3+&4:
    E44=E4+&4
150 E41=E4+&1:E42=E41+&1:E03=E39+&1:E31=E3+&1:E04=E49+&1:E32=E3+&2:
    E33=E3+&3:E34=E3+&4:E35=E3+&5:E36=E3+&6:E37=E3+&7:E43=E42+&1
170 E45=E44+&1:E46=E45+&1:E47=E46+&1:E411=E04+&1:E412=E411+&1:
    E341=E34+&10:E351=E35+&10:E361=E36+&10:SIGN=E361+&1:
190 FOR C=0 TO 3 STEP 3:P%=&4990:[OPT 0
200
    .SCALI
    .START
    .MAAL
    .SCALE
    .SSCAN
    .GROEP
    .MAX
    .REOV
    .FILT
    .DELING
    .A1M
    .R11
    .REDIN
    .LOC
    .AMPO
1390 ]NEXT C
1400 CLS
1410 PRINT TAB(0,0);"press ESCAPE when not in use":PRINT "use f1
     after ESCAPE":PRINT"press BREAK after system's crash":PRINT
```

```
     "use f9 after BREAK and NO printer is connected, else use f4"
1420 Z=0:PR=0
1430 CALL SCALI
1440 Z=Z+1
1450 IF Z=3 THEN VDU7:PRINT TAB(0,6);"please switch the sensor on
     or switch to external oscillator           "
1460 IF Z>3 GOTO 1420
1470 CALL SSCAN
1480 IF ?XTEMP=1 GOTO 1430
1490 PRINT TAB(0.6);"                                                  "
1495 CALL GROEP
1500 Z=0
1510 CALL START
1520 CALL SSCAN
1530 IF ?XTEMP=1 GOTO 1420
1540 IF ?SY=0 GOTO 1420
1550 CALL GROEP
1570 IF ?MAXA<40 THEN VDU7:PRINT TAB(0,8);"is the laser switched on
     check position of the object under test":GOTO 1410
1580 PRINT TAB(0.8);"                                                  "
1590 CALL A1M
1600 C=(1-2*?SIGN)*?E03+256
1610 IF ?E03>50 THEN VDU7:?E2=255:PRINT TAB(0,8);"adjust the sensor
     direction in such a way that this number [";C;"]approaches 256
     :GOTO 1410
1620 PRINT TAB(0,8);"                                                  "
1630 CALL REDIN:R1I=256*(256*?E47+?E48)+?E49
```

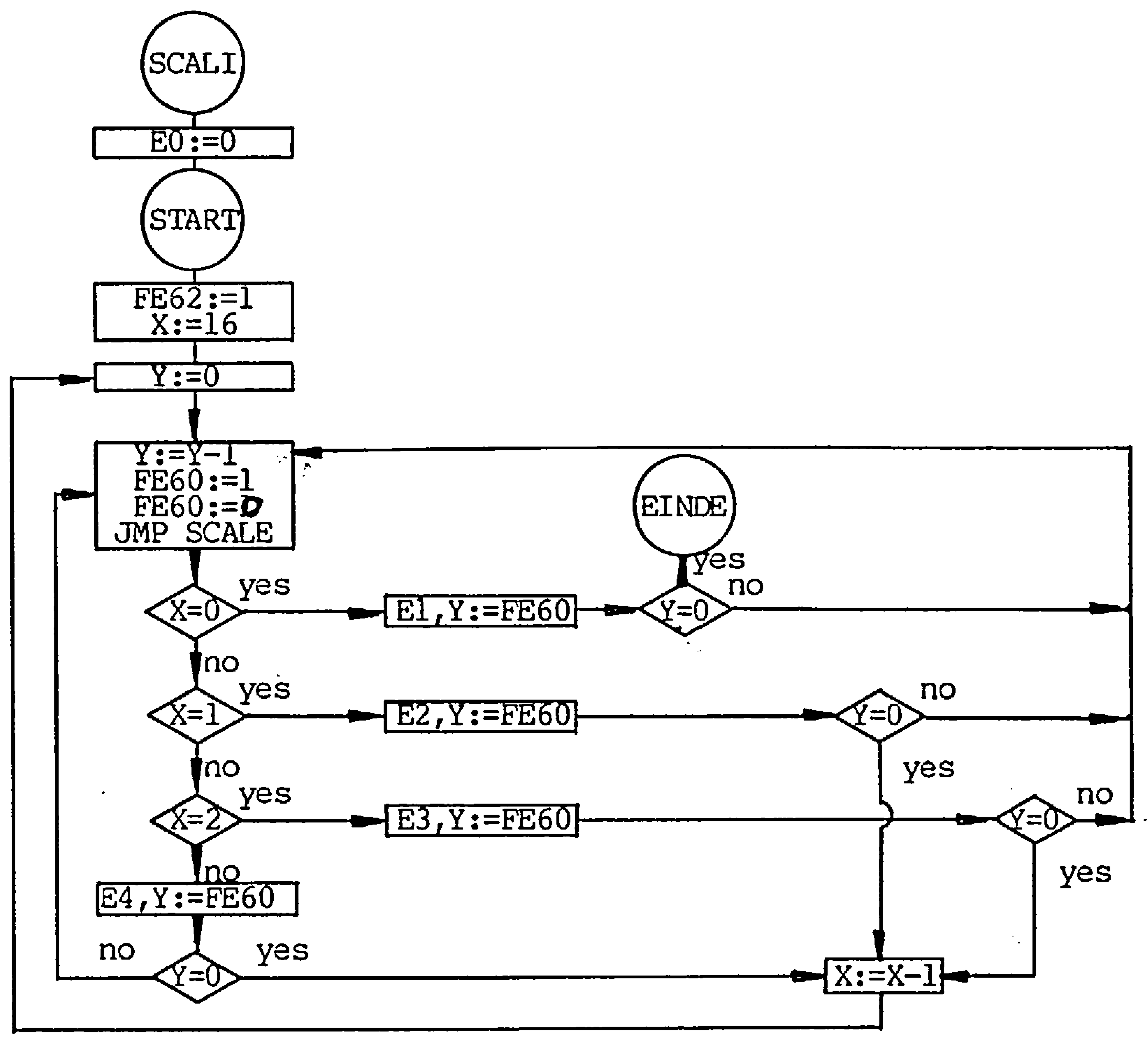
```
1640 IF R1I<10 THEN GOTO 1670

1650 PR=PR+1

1660 IF PR=3 THEN VDU7:PRINT TAB(0,8);"the spectrum is too asymmetric
     Press f2 to continue. ";R1I:END

1663 GOTO 1430

1665 print tab(0,10);"                                                  "

1670 CALL LOC:?(E1+?E43)=255:?(E2+?E44)=255:?(E1+?E41)=255:
     ?(E1+?E42)=255:?(E1+?E45)=255:?(E1+?E46)=255

1671 IF ?E49=0 THEN VDU7:PRINT TAB(0,15);"cannot locate the spectrum,
     press f1 to continue":END

1760 PW=2*(256+?E44-?E43)

1761 IF ?E45-?E44<PW THEN GOTO 1800

1770 PR=PR+1

1780 IF PR=3 THEN GOTO 1840

1790 GOTO 1430

1800 IF ?E45-?E44<PW THEN GOTO 1850

1810 PR=PR+1

1820 IF PR=3 THEN GOTO 1840

1830 GOTO 1430

1840 VDU7:PRINT TAB(0,8);"cannot locate the harmonics":GOTO 2040

1850 V=V+1:CALL AMPO:SIGMA=?E41+?E42+?E43+?E44+?E45*5*?E02:
     ANUL=(?E02+?E43)/SIGMA:AEEN=(2*?E02+?E42+?E44)/(2*SIGMA)

1890 RQ=SQR(1-ANUL)/(4*PI)

1900 IF RQ>0.03 GOTO 1920

1910 GOTO 2000

1920 RQ=.932*(1-ANUL)+.0167

1930 IF RQ>.07 GOTO 1950
```

```
1940 GOTO 2000

1950 RQ=.609*AEEN*(1-2*AEEN/(1-ANUL))+.0946

1960 IF RQ>.125 GOTO 1980 THEN VDU7:PRINT TAB(0.15):"the surface is
     too rough":GOTO 2040

1970 IF ?E47<24 THEN VDU7:PRINT TAB(0,15);"cannot determine the feed
     spacing":GOTO 2040

1990 VDU3

2000 RQ=INT(RQ*877):FE=INT((.877*273*512)/(12.7*?E47))

2010 PRINT TAB(0,15);"the roughness Rq equals ";RQ;" nanometers "

2030 PRINT TAB(0,17);"the feed spacing is ";FE;" micrometers "

2034 *FX 5,0

2035 VDU#

2040 CLG

2050 FOR N=0 TO 511

2060 MOVE 2*N,761

2070 DRAW 2*N,?(E1+N)+761

2080 NEXT N

2090 PRINT TAB(0,15);"                                                    "

2110 GOTO 1420
```

Subroutine .SCALI

SCALI scans the photo diode array 8 times. The last two scans are stored in memory locations &5400-&57FF. The scan time is governed by the value of E0. When starting from .SCALI the fastest scanning time is selected. When starting from .START, the value E0 is calculated by subroutine .MAAL. The extra scanning time is generated by subroutine .SCALE. The statements between .KLOK and .WEER produce pulses for driving the photo diode array and the A/D converter.

LABEL	OPERATION	BYTES	CYCLES
.SCALI	CLD	1	2
	LDA #&0	2	2
	STA E0	3	4
.START	CLD	1	2
	LDA #&1	2	2
	STA &FE62	3	4
	LDX #&10	2	2
.TEL	LDY #&0	2	2
.KLOK	DEY	1	2
	LDA #&1	2	2
	STA &FE60	3	4
	LDA #&0	2	2
	STA &FE60	3	4
	JMP SCALE	3	3
.WEER	CPX #&0	2	2
	BEQ BLAD1	2	2,3
	CPX #&1	2	2
	BEQ BLAD2	2	2,3
	CPX #&2	2	2
	BEQ BLAD3	2	2,3
.BLAD4	LDA &FE60	3	4
	STA E4,Y	3	5
	JMP B4	3	3
.B4	TYA	1	2
	BEQ FRAME	2	2,3
	JMP B6	3	3
.B6	JMP KLOK	3	3
.BLAD3	LDA &FE60	3	4
	STA E3,Y	3	5
	NOP	1	2
	TYA	1	2
	BEQ FRAME	2	2,3
	JMP B7	3	3
.B7	JMP KLOK	3	3
.BLAD2	LDA &FE60	3	4
	STA E2,Y	3	5
	ROR WASTE	3	6
	TYA	1	2
	BEQ FRAME	2	2,3
	JMP B8	3	3
.B8	JMP KLOK	3	3
.BLAD1	LDA &FE60	3	4
	STA E1,Y	3	5
	ROR WASTE	3	6
	NOP	1	2
	NOP	1	2
	TYA	1	2
	BEQ EINDE	2	2,3
	JMP B5	3	3
.B5	JMP KLOK	3	3
.FRAME	DEX	1	2
	JMP TEL	3	3
.EINDE	RTS	1	6


SCALI
E0:=0
START
FE62:=1
X:=16
Y:=0
Y:=Y-1
FE60:=1
FE60:=0
JMP SCALE
X=0
yes
E1,Y:=FE60
Y=0
yes
EINDE
no
no
X=1
yes
E2,Y:=FE60
Y=0
no
yes
no
X=2
yes
E3,Y:=FE60
Y=0
no
yes
no
E4,Y:=FE60
no
Y=0
yes
X:=X-1

Subroutine .MAAL

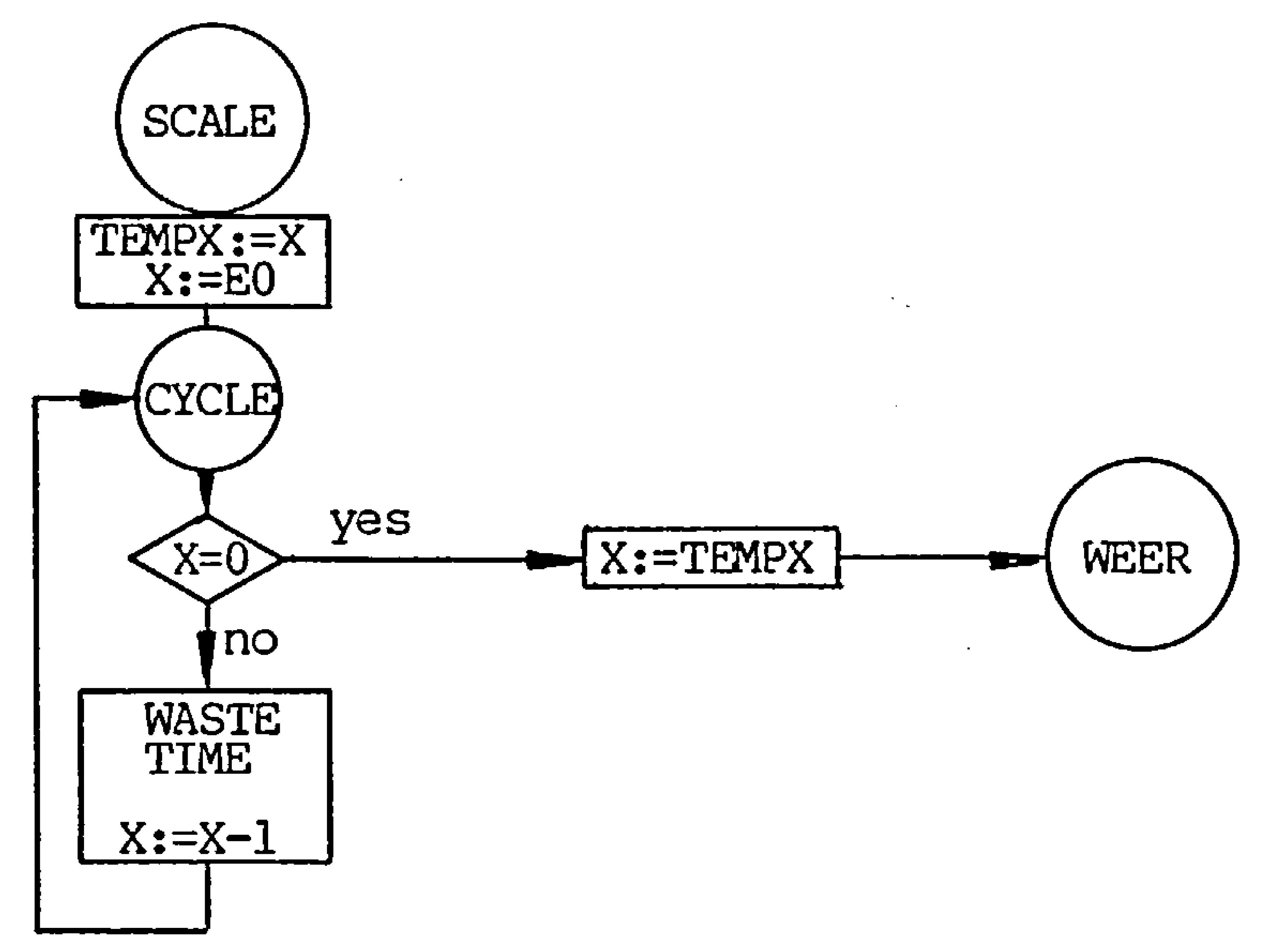
MAAL calculates the value E0, which governs the diode scanning rate. The scanning rate is adjusted for the second scan such that photo diode array is close to saturation. The input values for this subroutine are the dc offset of the diode array, E02, and the maximum intensity value MAXA, obtained during the first scan.

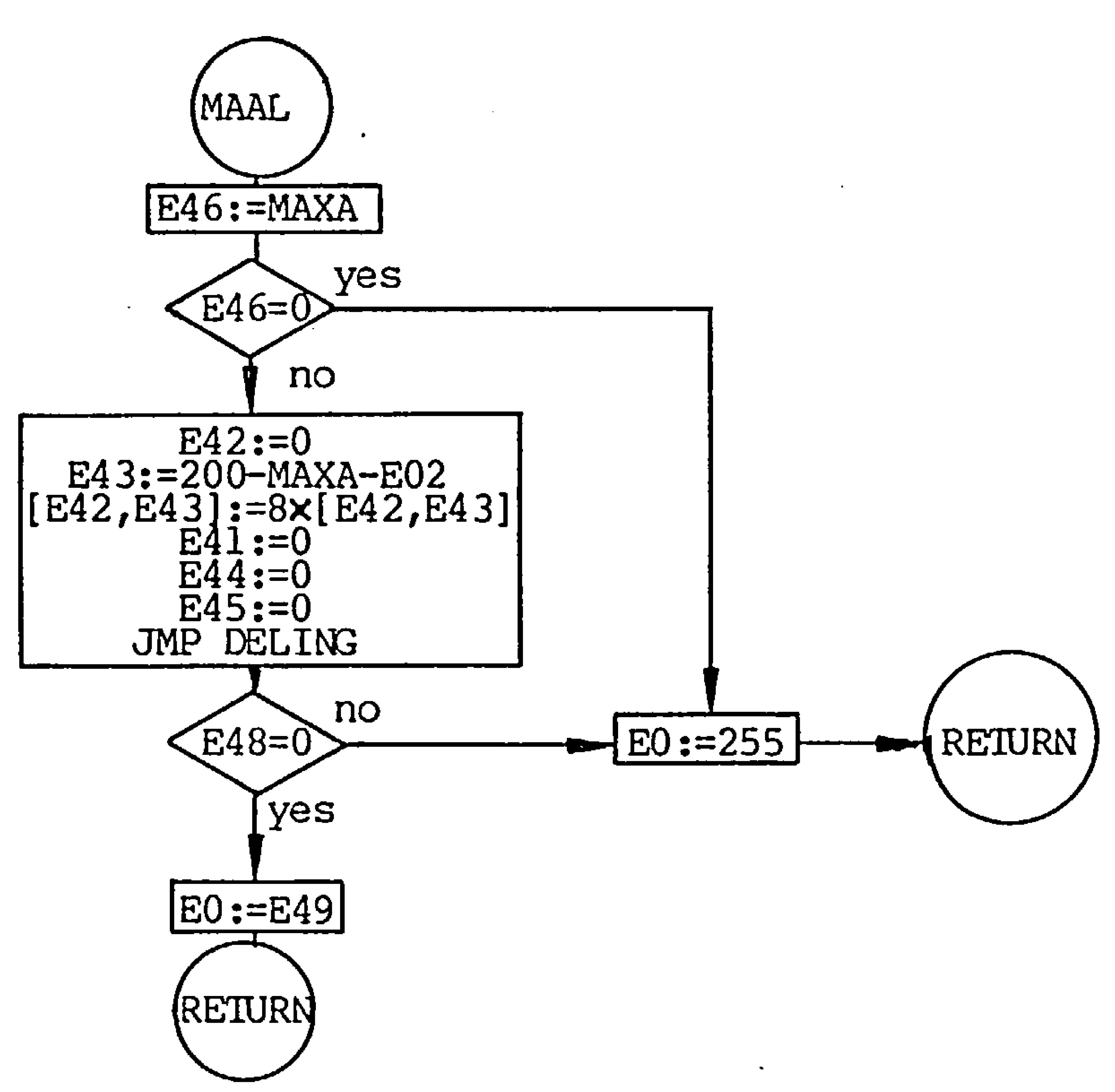
LABEL	OPERATION	BYTES	CYCLES
.MAAL	CLD	1	2
	LDA MAXA	3	4
	STA E46	3	4
	BEQ MAAL1	2	2,3
	LDA #&0	2	2
	STA E42	3	4
	SEC	1	2
	LDA #&C8	2	2
	SBC MAXA	3	4
	SBC E02	3	4
	STA E43	3	4
	CLC	1	2
	ROL E43	3	6
	ROL E42	3	6
	ROL E43	3	6
	ROL E42	3	6
	ROL E43	3	6
	ROL E42	3	6
	LDA #&0	2	2
	STA E41	3	4
	STA E44	3	4
	STA E45	3	4
	JSR DELING	3	6
	LDA E48	3	4
	BEQ MAAL2	2	2,3
.MAAL1	LDA #&FF	2	2
	STA E0	3	4
	RTS	1	6
.MAAL2	LDA E49	3	4
	STA E0	3	4
	RTS	1	6

subroutine .SCALE

SCALE introduces a delay time in the scanning rate of the photodiode array. The routine is called by SCALI and START. The extra delay time depends on the value E0.

LABEL	OPERATION	BYTES	CYCLES
.SCALE	STX TEMPX	3	4
	LDX E0	3	4
.CYCLE	CPX #&0	2	2
	BEQ SCAF	2	2,3
	DEX	1	2
	JMP CYCLE	3	3
.SCAF	LDX TEMPX	3	4
	JMP WEER	3	3

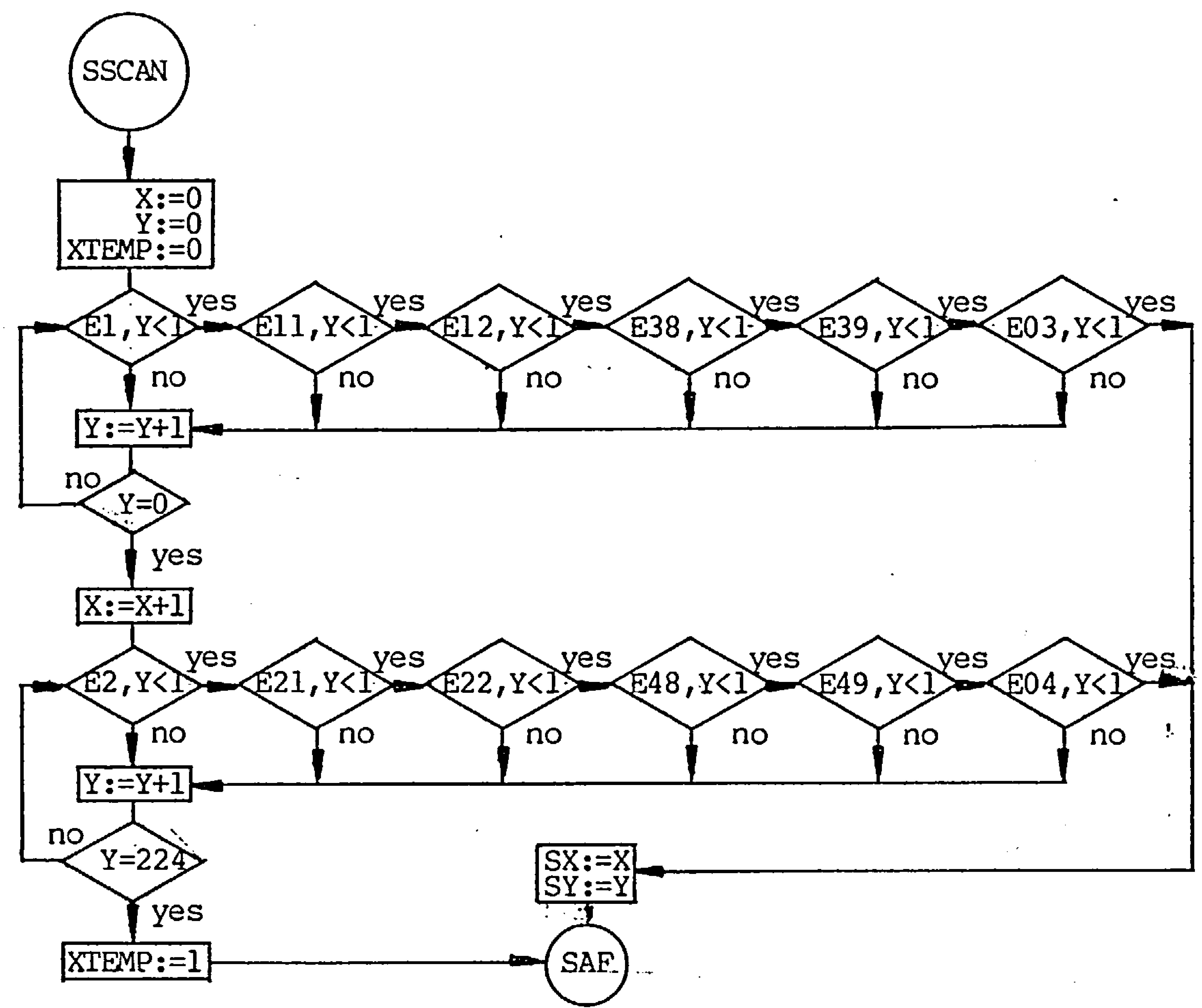

SCALE
TEMPX:=X
X:=E0
CYCLE
X=0
yes
X:=TEMPX
WEER
no
WASTE
TIME
X:=X-1


MAAL
E46:=MAXA
E46=0
yes
no
E42:=0
E43:=200-MAXA-E02
[E42,E43]:=8×[E42,E43]
E41:=0
E44:=0
E45:=0
JMP DELING
E48=0
no
E0:=255
RETURN
yes
E0:=E49
RETURN

subroutine .SSCAN

SSCAN locates the start position of one complete scan, and stores these values in SX and SY. These values are used by GROEP for further processing of the data. If the routine fails to locate the starting position, it returns XTEMP=1.

LABEL	OPERATION	BYTES	CYCLES
.SSCAN	CLD	1	2
	LDY #&0	2	2
	STY XTEMP	3	4
	LDX #&0	2	2
.SCAN	LDA E1,Y	3	4
	CMP #&1	2	2
	BCC UIT1	2	2,3
.SCAN1	INY	1	2
	BEQ PL	2	2,3
	JMP SCAN	3	3
.PL	INX	1	2
.PLUS	LDA E2,Y	3	4
	CMP #&1	2	2
	BCC UIT2	2	2,3
.PLUS1	INY	1	2
	CPY #&E0	2	2
	BEQ TERUG	2	2,3
	JMP PLUS	3	3
.UIT1	LDA E11,Y	3	4
	CMP #&1	2	2
	BCS SCAN1	2	2,3
	LDA E12,Y	3	4
	CMP #&1	2	2
	BCS SCAN1	2	2,3
	LDA E38,Y	3	4
	CMP #&1	2	2
	BCS SCAN1	2	2,3
	LDA E39,Y	3	4
	CMP #&1	2	2
	BCS SCAN1	2	2,3
	LDA E03,Y	3	4
	CMP #&1	2	2
	BCC UIT	2	2,3
	JMP SCAN1	3	3
.UIT2	LDA E21,Y	3	4
	CMP #&1	2	2
	BCS PLUS1	2	2,3
	LDA E22,Y	3	4
	CMP #&1	2	2
	BCS PLUS1	2	2,3
	LDA E48,Y	3	4
	CMP #&1	2	2
	BCS PLUS1	2	2,3
	LDA E49,Y	3	4
	CMP #&1	2	2
	BCS PLUS1	2	2,3
	LDA E04,Y	3	4
	CMP #&1	2	2
	BCC UIT	2	2,3
	JMP PLUS1	3	3
.TERUG	LDA #&1	2	2
	STA XTEMP	3	4
	JMP SAF	3	3
.UIT	STX SX	3	4
	STY SY	3	4
.SAF	RTS	1	6


SSCAN
X:=0
Y:=0
XTEMP:=0
E1,Y<1
E11,Y<1
E12,Y<1
E38,Y<1
E39,Y<1
E03,Y<1
yes
no
Y:=Y+1
Y=0
X:=X+1
E2,Y<1
E21,Y<1
E22,Y<1
E48,Y<1
E49,Y<1
E04,Y<1
Y=224
XTEMP:=1
SX:=X
SY:=Y
SAF

subroutine .GROEP

GROEP replaces one complete scan in the memory locations &5400-&55FF. The input information is created by routine SSCAN as the values SY and SX. The second part of the routine removes spurious spikes present in the spectrum. These spikes are caused by electronic noise pick-up.

LABEL	OPERATION	BYTES	CYCLES
.GROEP	CLD	1	2
	LDX SX	3	4
	LDY SY	3	4
	CPX #&0	2	2
	BEQ P5	2	2,3
	LDX #&0	2	2
.P1	LDA E24,Y	3	4
	STA E1,X	3	5
	INX	1	2
	INY	1	2
	BNE P1	2	2,3
.P2	LDA E34,Y	3	4
	STA E1,X	3	5
	INY	1	2
	INX	1	2
	BNE P2	2	2,3
.P3	LDA E34,Y	3	4
	STA E2,X	3	5
	INX	1	2
	INY	1	2
	BNE P3	2	2,3
.P4	LDA E44,Y	3	4
	STA E2,X	3	5
	INY	1	2
	INX	1	2
	BNE P4	2	2,3
	JMP KLAAR	3	3
.P5	LDA E14,Y	3	4
	STA E1,X	3	5
	INX	1	2
	INY	1	2
	BNE P5	2	2,3
.P6	LDA E24,Y	3	4
	STA E1,X	3	5
	INY	1	2
	INX	1	2
	BNE P6	2	2,3
.P7	LDA E24,Y	3	4
	STA E2,X	3	5
	INX	1	2
	INY	1	2
	BNE P7	2	2,3
.P8	LDA E34,Y	3	4
	STA E2,X	3	5
	INY	1	2
	INX	1	2
	BNE P8	2	2,3

subroutine .GROEP

LABEL	OPERATION	BYTES	CYCLES
.KLAAR	LDY #&0	2	2
.VP1	LDA E11,Y	3	4
	SEC	1	2
	SBC E1,Y	3	4
	BCC NEG1	2	2,3
	CMP #&6	2	2
	BCC VP2	2	2,3
	LDA E11,Y	3	4
	SEC	1	2
	SBC E12,Y	3	4
	BCC VP2	2	2,3
	CMP #&6	2	2
	BCC VP2	2	2,3
.NP11	CLC	1	2
	LDA E1,Y	3	4
	ADC E12,Y	3	4
	ROR A	1	2
	STA E11,Y	3	5
	JMP VP2	2	2,3
.NEG1	LDA E1,Y	3	4
	SEC	1	2
	SBC E11,Y	3	4
	CMP #&6	2	2
	BCC VP2	2	2,3
	LDA E12,Y	3	4
	SEC	1	2
	SBC E11,Y	3	4
	CMP #&6	2	2
	BCC VP2	2	2,3
	JMP NP11	3	3
.VP2	LDA E21,Y	3	4
	SEC	1	2
	SBC E2,Y	3	4
	BCC NEG2	2	2,3
	CMP #&6	2	2
	BCC VP0	2	2,3
	LDA E21,Y	3	4
	SEC	1	2
	SBC E22,Y	3	4
	CMP #&6	2	2
	BCC VP0	2	2,3
.NP21	CLC	1	2
	LDA E2,Y	3	4
	ADC E22,Y	3	4
	ROR A	1	2
	STA E21,Y	3	5
	JMP VP0	3	3
.NEG2	LDA E2,Y	3	4
	SEC	1	2
	SBC E21,Y	3	4
	CMP #&6	2	2
	BCC VP0	2	2,3
	LDA E22,Y	3	4
	SEC	1	2
	SBC E21,Y	3	4
	BCC VP0	2	2,3
	CMP #&6	2	2
	BCC VP0	2	2,3
	JMP NP21	3	3
.VP0	INY	1	2
	BEQ NSPIKE	2	2,3
	JMP VP1	3	3
.NSPIKE	LDA E12	3	4
	STA E1	3	4
	STA E11	3	4
	LDY #&FD	2	2
	LDA E2,Y	3	4
	STA E21,Y	3	5
	JMP FILT	3	3

GROEP
X:=SX
Y:=SY
X=0
yes
no
X:=0
E1,X:=E14,Y
X:=X+1
Y:=Y+1
Y=0
E1,X:=E24,Y
Y:=Y+1
X:=X+1
X=0
E2,X:=E24,Y
X:=X+1
Y:=Y+1
E2,X:=E34,Y
Y:=Y+1
X:=X+1
E1,X:=E34,Y
E2,X:=E34,Y
E2,X:=E44,Y
Y:=0
E11,Y<E1,Y
E11,Y<E1,Y+6
E11,Y<E12,Y
E11,Y<E12+6
E1,Y<E11,Y+6
E21,Y<E2,Y
E21,Y<E2,Y+6
E21,Y<E22,Y+6
E12,Y<E11,Y+6
E2,Y<E21,Y+6
Y:=Y+1
E11,Y:=(E1,Y+E12,Y)/2
E22,Y<E21,Y
Y=0
E1:=E12
E11:=E12
Y:=#&FD
E21,Y:=E2,Y
JMP FILT
E22,Y<E21,Y+6
E21,Y:=(E2,Y+E22,Y)/2

subroutine .MAX

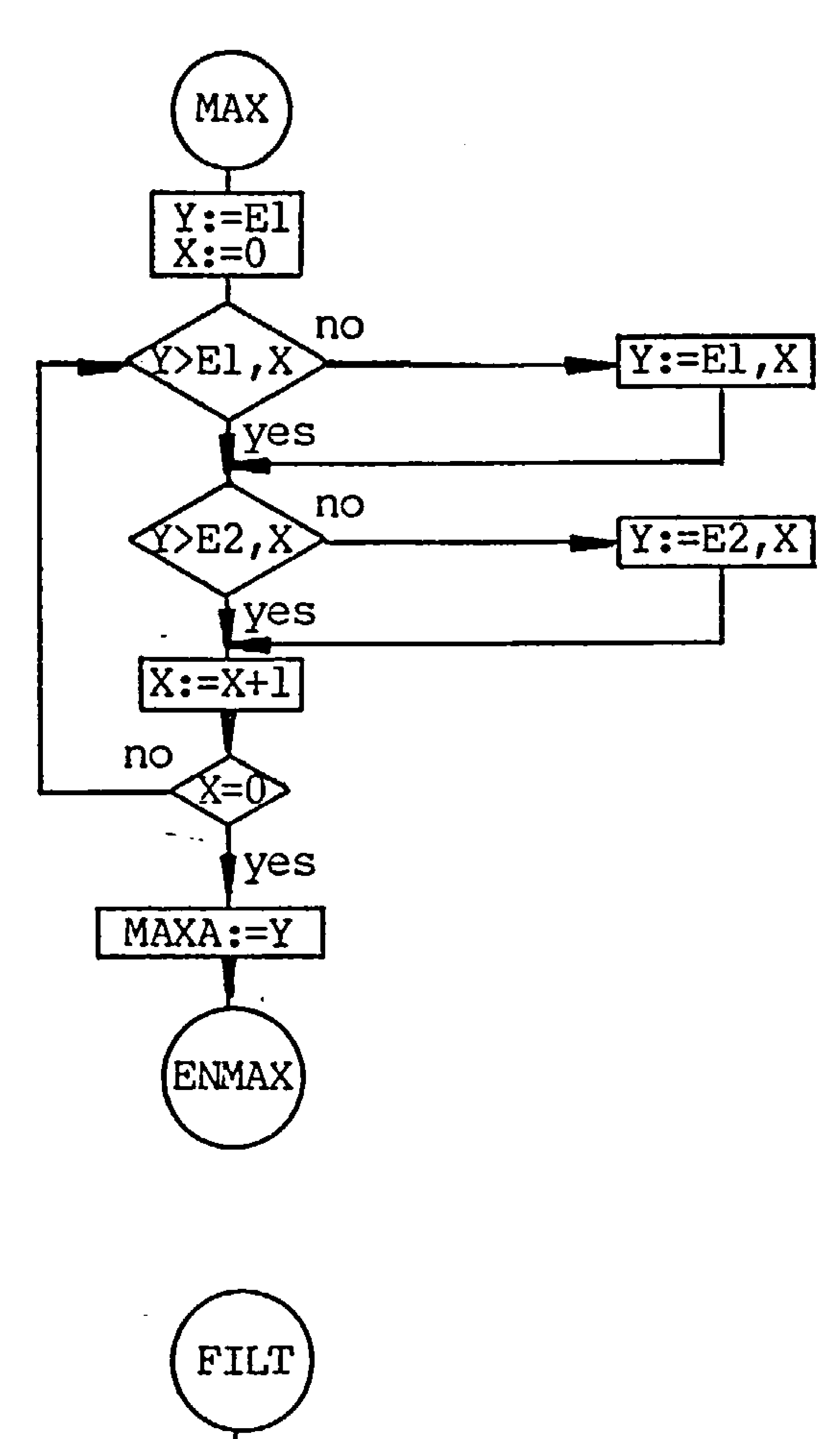
MAX calculates the maximum value present in the spectrum and stores its value in MAXA.

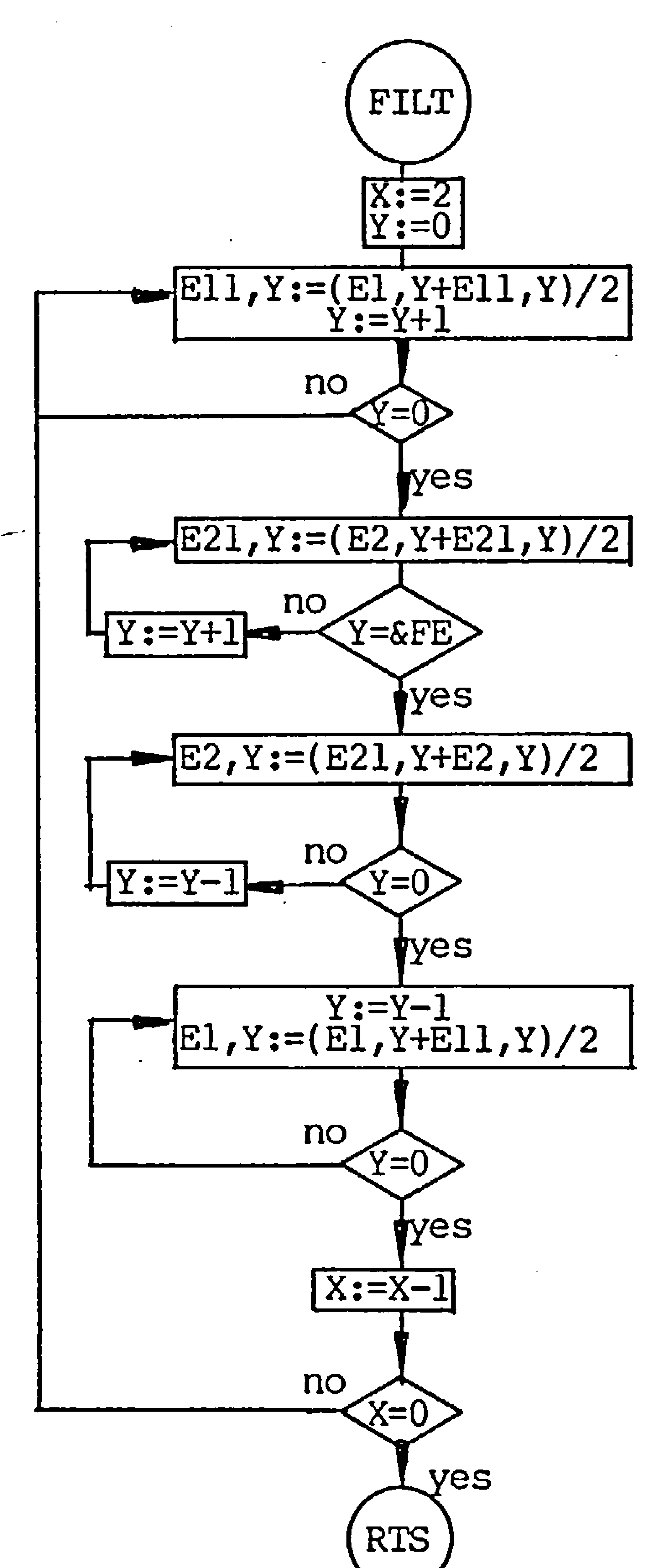
LABEL	OPERATION	BYTES	CYCLES
.MAX	CLD	1	2
	LDY E1	3	4
	LDX #&0	2	2
.MP1	TYA	1	2
	CMP E1,X	3	4
	BCS MP2	2	2,3
	LDY E1,X	3	4
.MP2	TYA	1	2
	CMP E2,X	3	4
	BCS MP0	2	2,3
	LDY E2,X	3	4
.MP0	INX	1	2
	BNE MP1	2	2,3
	STY MAXA	3	4
	RTS	1	6

subroutine .FILT

FILT is a digital low pass filter, which eliminates the high frequencies in the sensor signal. These high frequencies are caused by the differences in the integration times of the individual photo diodes in the array.

LABEL	OPERATION	BYTES	CYCLES
.FILT	CLD	1	2
	LDX #&1	2	2
	LDY #&0	2	2
.FILT1	LDA E1,Y	3	4
	CLC	1	2
	ADC E11,Y	3	4
	ROR A	1	2
	STA E11,Y	3	5
	INY	1	2
	BEQ FILT2	2	2,3
	JMP FILT1	3	3
.FILT2	LDA E2,Y	3	4
	CLC	1	2
	ADC E21,Y	3	4
	ROR A	1	2
	STA E21,Y	3	5
	CPY #&FE	2	2
	BEQ FILT3	2	2,3
	INY	1	2
	JMP FILT2	3	3
.FILT3	LDA E21,Y	3	4
	CLC	1	2
	ADC E2,Y	3	4
	ROR A	1	2
	STA E2,Y	3	5
	CPY #&0	2	2
	BEQ FILT4	2	2,3
	DEY	1	2
	JMP FILT3	3	3
.FILT4	DEY	1	2
	LDA E11,Y	3	4
	CLC	1	2
	ADC E1,Y	3	4
	ROR A	1	2
	STA E1,Y	3	5
	CPY #&0	2	2
	BEQ FILT5	2	2,3
	JMP FILT4	3	3
.FILT5	DEX	1	2
	BNE FILT1	2	2,3
	JMP REOV	3	3

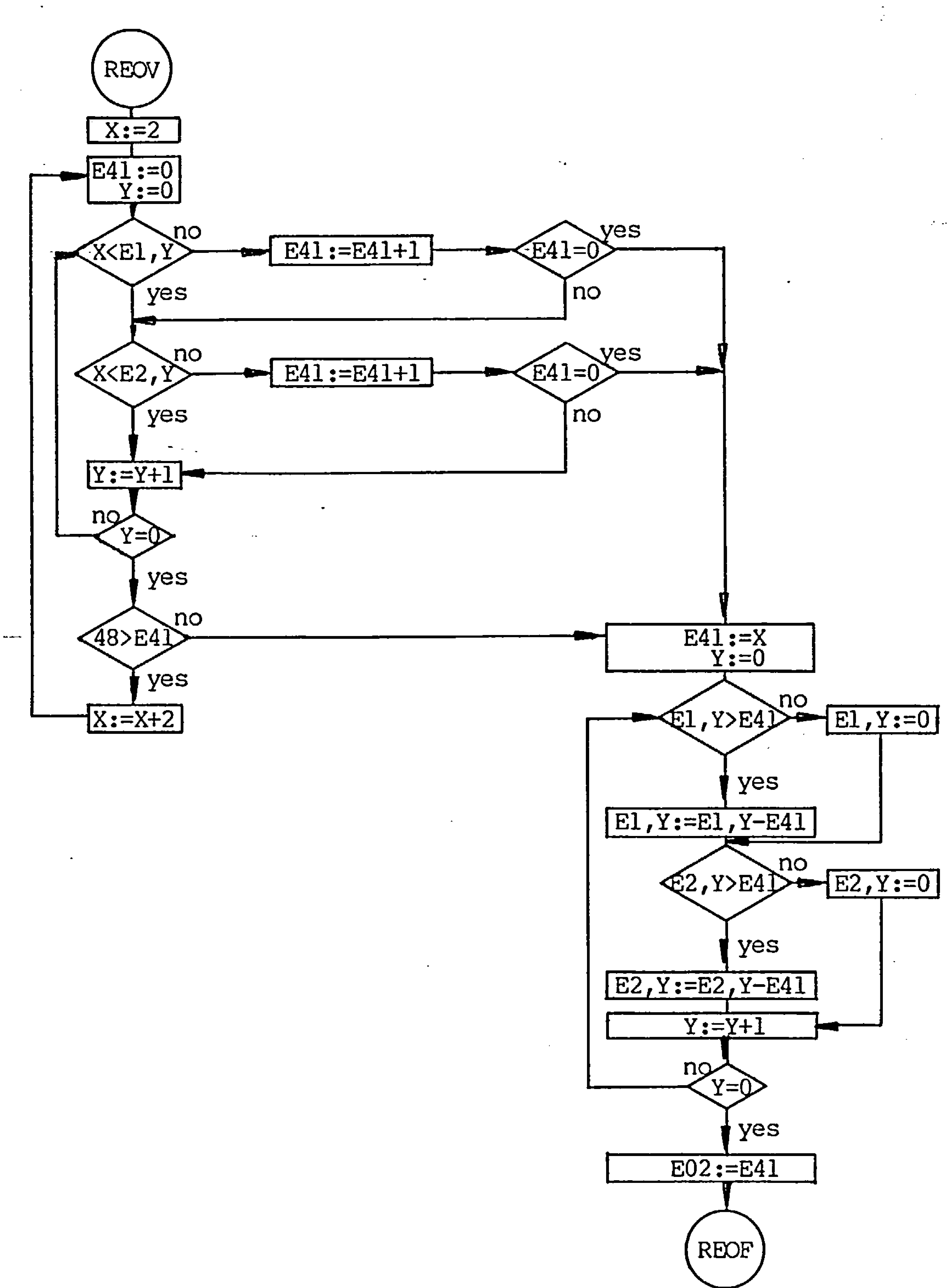

MAX
Y:=E1
X:=0
Y>E1,X
no
Y:=E1,X
yes
Y>E2,X
no
Y:=E2,X
yes
X:=X+1
no
X=0
yes
MAXA:=Y
ENMAX


FILT
X:=2
Y:=0
E11,Y:=(E1,Y+E11,Y)/2
Y:=Y+1
no
Y=0
yes
E21,Y:=(E2,Y+E21,Y)/2
no
Y:=Y+1
Y=&FE
yes
E2,Y:=(E21,Y+E2,Y)/2
no
Y:=Y-1
Y=0
yes
Y:=Y-1
E1,Y:=(E1,Y+E11,Y)/2
no
Y=0
yes
X:=X-1
no
X=0
yes
RTS

subroutine .REOV

REOV removes the dc offset present in the signal. This dc offset is caused by the thermal instability of the electronics boards in the sensor head. The value of the offset is stored in memory E02.

LABEL	OPERATION	BYTES	CYCLES
.REOV	CLD	1	2
	LDX #&2	2	2
.OP3	LDA #&0	2	2
	STA E41	3	4
	TAY	1	2
.OP0	TXA	1	2
	CMP E1,Y	3	4
	BCC OP1	2	2,3
	INC E41	3	6
	BEQ TREKAF1	2	2,3
.OP1	TXA	1	2
	CMP E2,Y	3	4
	BCC OP2	2	2,3
	INC E41	3	6
	BEQ TREKAF1	2	2,3
.OP2	INY	1	2
	BEQ TESTX	2	2,3
	JMP OP0	3	3
.TESTX	LDA #&30	2	2
	CMP E41	3	4
	BCS NEWX	2	2,3
.TREKAF1	STX E41	3	4
	LDY #&0	2	2
.TREKAF	LDA E1,Y	3	4
	SEC	1	2
	SBC E41	3	4
	BCS NEWY1	2	2,3
	LDA #&0	2	2
.NEWY1	STA E1,Y	3	5
	LDA E2,Y	3	4
	SEC	1	2
	SBC E41	3	4
	BCS NEWY2	2	2,3
	LDA #&0	2	2
.NEWY2	STA E2,Y	3	5
	INY	1	2
	BNE TREKAF	2	2,3
	JMP REOF	3	3
.NEWX	INX	1	2
	INX	1	2
	JMP OP3	3	3
.REOF	LDA E41	3	4
	STA E02	3	4
	RTS	1	6

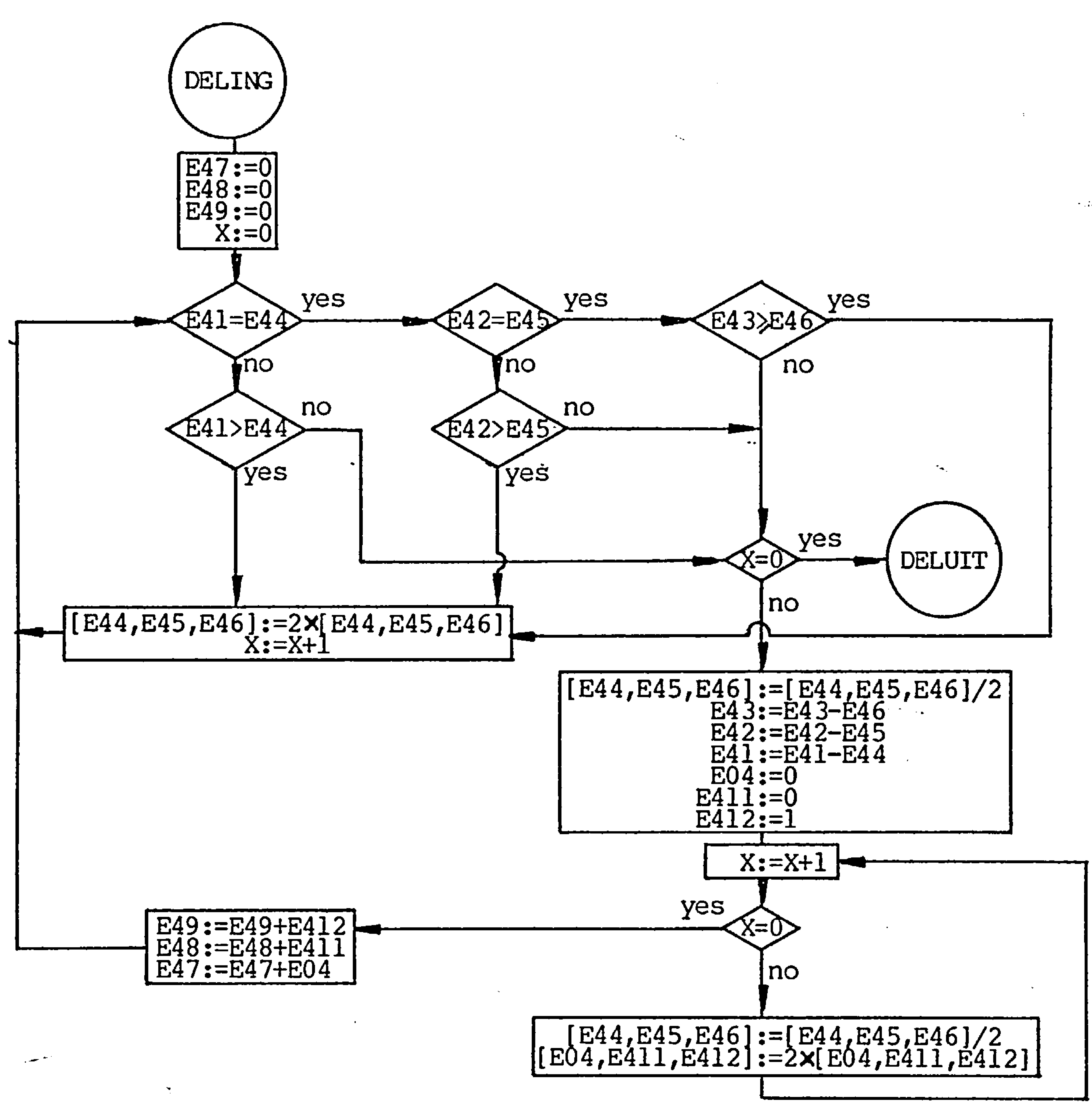


REOV
X:=2
E41:=0
Y:=0
X<E1,Y
no
yes
E41:=E41+1
E41=0
yes
no
X<E2,Y
no
yes
E41:=E41+1
E41=0
yes
no
Y:=Y+1
no
Y=0
yes
48>E41
no
yes
X:=X+2
E41:=X
Y:=0
E1,Y>E41
no
E1,Y:=0
yes
E1,Y:=E1,Y-E41
E2,Y>E41
no
E2,Y:=0
yes
E2,Y:=E2,Y-E41
Y:=Y+1
no
Y=0
yes
E02:=E41
REOF

subroutine .DELING

DELING performs a division. The numerator, divisor and quotient are resp. stored in (E41,E42,E43), (E44,E45,E46) and (E47,E48,E49).

LABEL	OPERATION	BYTES	CYCLES
.DELING	CLD	1	2
	LDA #&0	2	2
	STA E47	3	4
	STA E48	3	4
	STA E49	3	4
	TAX	1	2
.DEL1	LDA E41	3	4
	CMP E44	3	4
	BEQ DEL2	2	2,3
	BCS DEL5	2	2,3
	JMP DEL4	3	3
.DEL2	LDA E42	3	4
	CMP E45	3	4
	BEQ DEL3	2	2,3
	BCS DEL5	2	2,3
	JMP DEL4	3	3
.DEL3	LDA E43	3	4
	CMP E46	3	4
	BCS DEL5	2	2,3
	JMP DEL4	3	3
.DEL5	CLC	1	2
	ROL E46	3	6
	ROL E45	3	6
	ROL E44	3	6
	INX	1	2
	JMP DEL1	3	3
.DEL4	CPX #&0	2	2
	BEQ DELUIT	2	2,3
	CLC	1	2
	ROR E44	3	6
	ROR E45	3	6
	ROR E46	3	6
	SEC	1	2
	LDA E43	3	4
	SBC E46	3	4
	STA E43	3	4
	LDA E42	3	4
	SBC E45	3	4
	STA E42	3	4
	LDA E41	3	4
	SBC E44	3	4
	STA E41	3	4
	LDA #&0	2	2
	STA E04	3	4
	STA E411	3	4
	LDA #&1	2	2
	STA E412	3	4
.DEL7	DEX	1	2
	BEQ DEL6	2	2,3
	CLC	1	2
	ROR E44	3	6
	ROR E45	3	6
	ROR E46	3	6
	CLC	1	2
	ROL E412	3	6
	ROL E411	3	6
	ROL E04	3	6
	JMP DEL7	3	3
.DEL6	CLC	1	2
	LDA E49	3	4
	ADC E412	3	4
	STA E49	3	4
	LDA E48	3	4
	ADC E411	3	4
	STA E48	3	4
	LDA E47	3	4
	ADC E04	3	4
	STA E47	3	4
	JMP DEL1	3	3
.DELUIT	RTS	1	6


DELING
E47:=0
E48:=0
E49:=0
X:=0
E41=E44
yes
no
E42=E45
yes
no
E43>E46
yes
no
E41>E44
no
yes
E42>E45
no
yes
X=0
yes
DELUIT
no
[E44,E45,E46]:=2×[E44,E45,E46]
X:=X+1
[E44,E45,E46]:=[E44,E45,E46]/2
E43:=E43-E46
E42:=E42-E45
E41:=E41-E44
E04:=0
E411:=0
E412:=1
X:=X+1
yes
X=0
no
E49:=E49+E412
E48:=E48+E411
E47:=E47+E04
[E44,E45,E46]:=[E44,E45,E46]/2
[E04,E411,E412]:=2×[E04,E411,E412]

subroutine AlM

AlM calculates the centre of gravity of the spectrum. Its value is stored in the memory locations E03 and E37.

LABEL	OPERATION	BYTES	CYCLES
.AlM	CLD	1	2
	LDA #&0	2	2
	STA E41	3	4
	STA E42	3	4
	STA E43	3	4
	STA E31	3	4
	STA E32	3	4
	STA E33	3	4
	STA E34	3	4
	STA E35	3	4
	STA E36	3	4
	STA E37	3	4
	STA E38	3	4
	LDY #&0	2	2
.AlM1	CLC	1	2
	LDA E41	3	4
	ADC El,Y	3	4
	STA E42	3	4
	LDA E41	3	4
	ADC #&0	3	4
	STA E41	3	4
	INY	1	2
	BEQ AlM9	2	2,3
	JMP AlM1	3	3
.AlM9	LDX #&1	2	2
.AlM2	TXA	1	2
	TAY	1	2
	CLC	1	2
	LDA El,X	3	4
	ADC E2,X	3	4
	STA E36	3	4
	LDA #&0	2	2
	STA E34	3	4
	ADC #&0	2	2
	STA E35	3	4
.AlM3	TYA	1	2
	BEQ AlM5	2	2,3
	ROR A	1	2
	TAY	1	2
	BCC AlM4	2	2,3
	CLC	1	2
	LDA E33	3	4
	ADC E36	3	4
	STA E33	3	4
	LDA E32	3	4
	ADC E35	3	4
	STA E32	3	4
	LDA E31	3	4
	ADC E34	3	4
	STA E31	3	4
.AlM4	CLC	1	2
	ROL E36	3	6
	ROL E35	3	6
	ROL E34	3	6
	JMP AlM3	3	3
.AlM5	INX	1	2
	BEQ AlM55	2	2,3
	JMP AlM2	3	3
.AlM55	LDA E41	3	4
	CMP E31	3	4
	BEQ AlM6	2	2,3
	BCS CNEG	2	2,3
	JMP AlM8	3	3
.AlM6	LDA E42	3	4
	CMP E32	3	4
	BEQ AlM7	2	2,3
	BCS CNEG	2	2,3
	JMP AlM8	3	3
.AlM7	LDA E43	3	4
	CMP E33	3	4

subroutine A1M

LABEL	OPERATION	BYTES	CYCLES
	BEQ A1M8	2	2,3
	BCS CNEG	2	2,3
.A1M8	SEC	1	2
	LDA E33	3	4
	STA E43	3	4
	LDA E32	3	4
	SBC E42	3	4
	STA E42	3	4
	LDA E31	3	4
	SBC E41	3	4
	STA E41	3	4
	LDA #&1	2	2
	STA E38	3	4
	JMP A2M	3	3
.CNEG	SEC	1	2
	LDA E43	3	4
	SBC E33	3	4
	STA E43	3	4
	LDA E42	3	4
	SBC E32	3	4
	STA E42	3	4
	LDA E41	3	4
	SBC E31	3	4
	STA E41	3	4
	LDA #&1	2	2
	STA E37	3	4
.A2M	LDA E37	3	4
	STA SIGN	3	4
	LDA #&0	2	2
	STA E341	3	4
	STA E351	3	4
	STA E361		
	TAX	1	2
.OPA21	CLC	1	2
	LDA E1,X	3	4
	ADC E341	3	4
	STA E341	3	4
	BCC OPA22	2	2,3
	INC E351	3	6
	BNE OPA22	2	2,3
	INC E361	3	6
.OPA22	CLC	1	2
	LDA E2,X	3	4
	ADC E341	3	4
	STA E341	3	4
	BCC OPX	2	2,3
	INC E351	3	6
	BNE OPX	2	2,3
	INC E361	3	6
.OPX	INX	1	2
	BNE OPA21	2	2,3
	LDA E361	3	4
	STA E44	3	4
	LDA E351	3	4
	STA E45	3	4
	LDA E341	3	4
	STA E46	3	4
	JSR DELING	3	6
	CLC	1	2
	LDA E49	3	4
	ADC E37	3	4
	STA E03	3	4
.A2MUIT	RTS	1	6

A1M
E41:=0 E34:=0
E42:=0 E35:=0
E43:=0 E36:=0
E31:=0 E37:=0
E32:=0 E38:=0
E33:=0 Y:=0
[E41,E47]:=E42+E1,Y
Y:=Y+1
no
Y=0
yes
X:=1
Y:=X
[E35,E36]:=E1,X+E2,X
E34:=0
no
Y=0
Y:=Y/2
C=0
no
E33:=E33+E36
E32:=E32+E35
E31:=E31+E34
yes
yes
X:=X+1
[E34,E35,E36]:=2 [E34,E35,E36]
no
X=0
yes
E41=E31
yes
E42=E32
yes
E43<E33
yes
no
no
no
E41>E31
no
E43:=E33
E42:=E32-E42
E41:=E31-E41
E38:=1
yes
E42>E32
no
yes
E43:=E43-E33
E42:=E42-E32
E41:=E41-E31
E37:=1
SIGN:=E37
E341:=0
E351:=0
E361:=0
X:=0
E341:=E341+E1,X
C=0
yes
no
E341:=E342+E2,X
C=0
yes
no
X:=X+1
E351:=E351+1
E351=0
no
yes
E361:=E361+1
E351:=E351+1
E351=0
no
yes
E361:=E361+1
X=0
no
yes
E44:=E361
E45:=E351
E46:=E341
JSR DELING
E03:=E49+E32
A2MUIT

subroutine .R11

R11 calculates the rms symmetrie of the spectrum. Its value is stored in E49.

LABEL	OPERATION	BYTES	CYCLES
.R11	LDA #&0	2	2
	STA E32	3	4
	STA E31	3	4
	STA E38	3	4
	LDX #&FF	2	2
	LDY #&1	2	2
.R1Z1	JSR PG12	3	6
	CPX #&0	2	2
	BEQ ROOT	2	2,3
	JMP R1Z1	3	3
.PG12	LDA E1,X	3	4
	CMP E2,Y	3	4
	BCS PG12P	2	2,3
	SEC	1	2
	LDA E2,Y	3	4
	SBC E1,X	3	4
	STA E33	3	4
.PG12K	JSR KWAD	3	6
	DEX	1	2
	INY	1	2
	RTS	1	6
.PG12P	SBC E2,Y	3	4
	STA E33	3	4
	JMP PG12K	3	3
.KWAD	LDA #&0	2	2
	STA E39	3	4
	LDA E33	3	4
	STA E37	3	4
.KWAD1	LDA E33	3	4
	CMP #&0	2	2
	BEQ KWADK	2	2,3
	CLC	1	2
	ROR E33	3	6
	BCS KWAD2	2	2,3
	JMP KWAR	3	3
.KWAD2	CLC	1	2
	LDA E31	3	4
	ADC E37	3	4
	STA E31	3	4
	LDA E32	3	4
	ADC E39	3	4
	STA E32	3	4
	LDA #&0	2	2
	ADC E38	3	4
	STA E38	3	4
.KWAR	CLC	1	2
	ROL E37	3	6
	ROL E39	3	6
	JMP KWAD1	3	3
.KWADK	RTS	1	6
.ROOT	CLD	1	2
	CLC	1	2
	LDA E38	3	4
	STA E33	3	4
	STA E44	3	4
	LDA E32	3	4
	STA E34	3	4
	STA E45	3	4
	LDA E31	3	4
	STA E46	3	4
	ROR E44	3	6
	ROR E45	3	6
	ROR E46	3	6
.ROT1	LDA E38	3	4
	STA E41	3	4
	LDA E32	3	4
	STA E42	3	4
	LDA E31	3	4
	STA E43	3	4
	JSR DELING	3	6

subroutine .R11

LABEL	OPERATION	BYTES	CYCLES
	SEC	1	2
	LDA E46	3	4
	SBC E49	3	4
	STA E35	3	4
	LDA E45	3	4
	SBC E48	3	4
	STA E34	3	4
	LDA E44	3	4
	SBC E47	3	4
	STA E33	3	4
	CLC	1	2
	ROR E33	3	6
	ROR E34	3	6
	ROR E35	3	6
	SEC	1	2
	LDA E46	3	4
	SBC E35	3	4
	STA E46	3	4
	LDA E45	3	4
	SBC E34	3	4
	STA E45	3	4
	LDA E44	3	4
	SBC E33	3	4
	STA E44	3	4
	LDA #&0	2	2
	CMP E33	3	4
	BEQ ROT2	2	2,3
	JMP ROT1	3	3
.ROT2	CMP E34	3	4
	BEQ ROT3	2	2,3
	JMP ROT1	3	3
.ROT3	CMP E35	3	4
	BEQ W100	2	2,3
	JMP ROT1	3	3
.W100	LDA #&0	2	2
	STA E41	3	4
	STA E42	3	4
	STA E43	3	4
	LDA #&64	2	2
	STA E49	3	4
.C100	LDA E49	3	4
	CMP #&0	2	2
	BEQ RINT	2	2,3
	CLC	1	2
	ROR E49	3	6
	BCC C101	2	2,3
	CLC	1	2
	LDA E43	3	4
	ADC E46	3	4
	STA E43	3	4
	LDA E42	3	4
	ADC E45	3	4
	STA E42	3	4
	LDA E41	3	4
	ADC E44	3	4
	STA E41	3	4
.C101	CLC	1	2
	ROL E46	3	6
	ROL E45	3	6
	ROL E44	3	6
	JMP C100	3	3
.RINT	LDA E361	3	4
	STA E44	3	4
	LDA E351	3	4
	STA E45	3	4
	LDA E341	3	4
	STA E46	3	4
	JSR DELING	3	6
.RIUIT	RTS	1	6

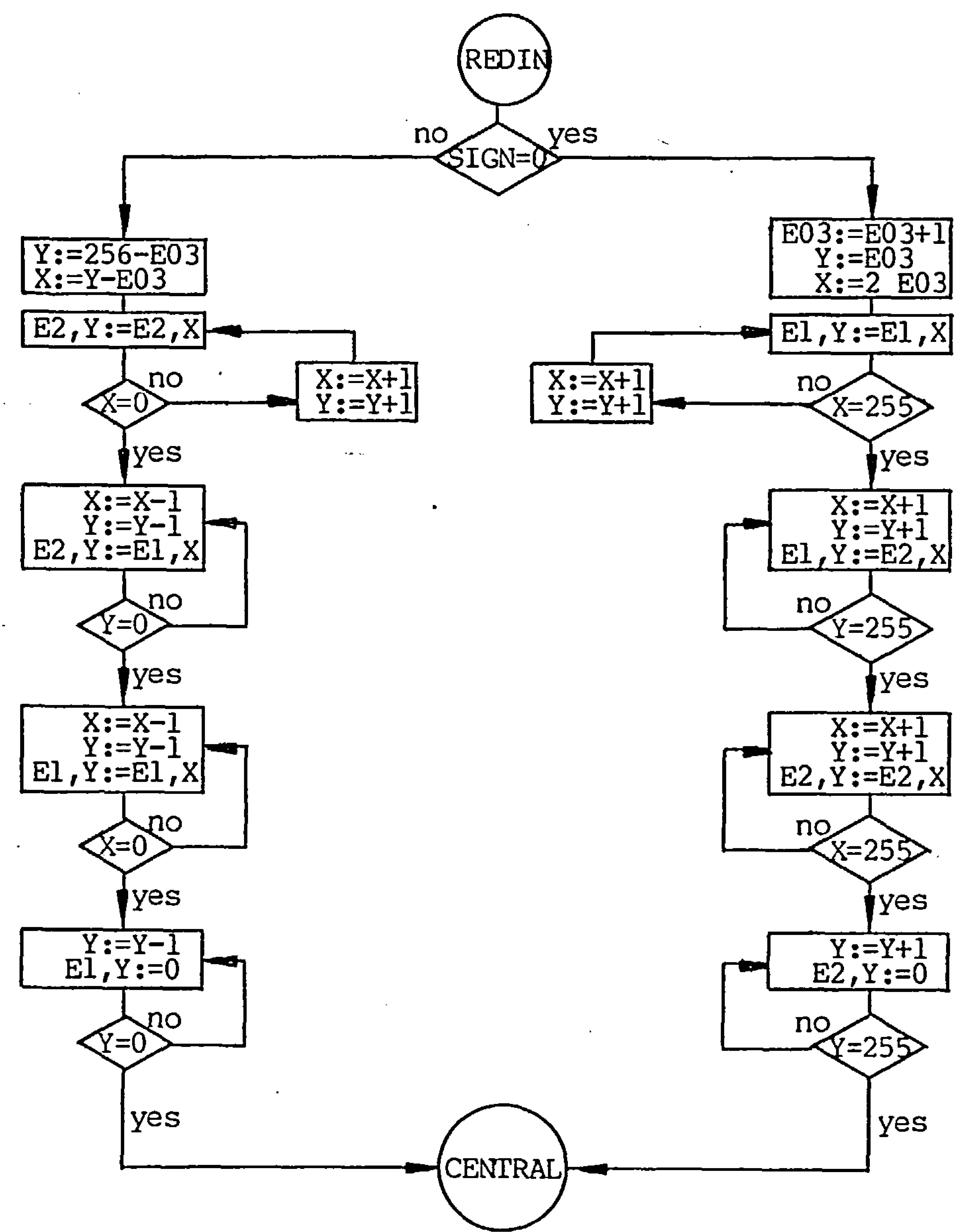
R11
E33:=0
E32:=0
E38:=0
X:=255
Y:=1
E1,X>E2,X
no
yes
E33:=E2,Y-E1,Y
E33:=E1,X-E2,Y
JSR KWAD
X:=X-1
Y:=Y+1
no
X=0
yes
E33:=E38 E34:=E32
E44:=E38 E45:=E32
E35:=E31 E46:=E31
[E44,E45,E46]:=[E44,E45,E46]/2
E41:=E38
E42:=E32
E43:=E31
JSR DELING
E35:=E46-E49
E34:=E45-E48
E33:=E44-E42
[E33,E34,E35]:=[E3,E34,E35]/2
E46:=E46-E35
E45:=E45-E34
E44:=E44-E32
no
E33=0
yes
no
E34=0
yes
no
E35=0
yes
E41:=0
E42:=0
E43:=0
E49:=0
yes
E49=0
no
E49:=E49/2
yes
C=0
[E44,E45,E46]:=2 [E44,E45,E46]
no
E43:=E43+E46
E42:=E42+E45
E41:=E41+E44
E44:=E361
E45:=E351
E46:=E341
JSR DELING
RIUIT
KWAD
E39:=0
E37:=E33
yes
E33=0
RTS
no
E33:=E33/2
yes
C=0
E31:=E31+E37
E32:=E32+E39
E38:=E38+C
no
[E39,E37]:=2 [E39,E32]

subroutine .REDIN

REDIN shifts the spectrum to the centre of the memory locations &5400-&55Ff for ease of further processing.

LABEL OPERATION BYTES CYCLES

LABEL	OPERATION	BYTES	CYCLES
.REDIN	CLD	1	2
	LDA SIGN	3	4
	BEQ POSC	2	2,3
	SEC	1	2
	LDA #&0	2	2
	SBC E03	3	4
	TAY	1	2
	SBC E03	3	4
	TAX	1	2
.NEGC2	JSR NEGC22	3	6
	CPX #&0	2	2
	BEQ NEGC1	2	2,3
	DEX	1	2
	DEY	1	2
	JMP NEGC2	3	3
.NEGC1	DEX	1	2
	DEY	1	2
	JSR NEGC11	3	6
	CPY #&0	2	2
	BEQ NEGC3	2	2,3
	JMP NEGC1	3	3
.NEGC3	DEX	1	2
	DEY	1	2
	JSR NEGC33	3	6
	CPX #&0	2	2
	BEQ NEGC4	2	2,3
	JMP NEGC3	3	3
.NEGC4	DEY	1	2
	LDA #&0	2	2
	STA E1,Y	3	5
	CPY #&0	2	2
	BEQ CENTRAL	2	2,3
	JMP NEGC4	3	3
.POSC	CLC	1	2
	LDA E03	3	4
	TAY	1	2
	ROL A	1	2
	TAX	1	2
.POSC2	JSR NEGC33	3	6
	CPX #&FF	2	2
	BEQ POSC1	2	2,3
	INX	1	2
	INY	1	2
	JMP POSC2	3	3
.POSC1	INY	1	2
	INX	1	2
	JSR NEGC44	3	6
	CPY #&FF	2	2
	BEQ POSC3	2	2,3
	JMP POSC1	3	3
.POSC3	INY	1	2
	INX	1	2
	JSR NEGC22	3	6
	CPX #&FF	2	2
	BEQ POSC4	2	2,3
	JMP POSC3	3	3
.POSC4	INY	1	2
	LDA #&0	2	2
	STA E2,Y	3	5
	CPY #&FF	2	2
	BEQ CENTRAL	2	2,3
	JMP POSC4	3	3
.CENTRAL	JMP R11	3	3
.NEGC22	LDA E2,X	3	4
	STA E2,Y	3	5
	RTS	1	6
.NEGC33	LDA E1,X	3	4
	STA E1,Y	3	5
	RTS	1	6
.NEGC11	LDA E1,X	3	4
	STA E2,Y	3	5
	RTS	1	6
.NEGC44	LDA E2,X	3	4
	STA E1,Y	3	5
	RTS	1	6


REDIN
no
SIGN=0
yes
Y:=256-E03
X:=Y-E03
E03:=E03+1
Y:=E03
X:=2 E03
E2,Y:=E2,X
E1,Y:=E1,X
X=0
no
X:=X+1
Y:=Y+1
X:=X+1
Y:=Y+1
no
X=255
yes
yes
X:=X-1
Y:=Y-1
E2,Y:=E1,X
X:=X+1
Y:=Y+1
E1,Y:=E2,X
Y=0
no
no
Y=255
yes
yes
X:=X-1
Y:=Y-1
E1,Y:=E1,X
X:=X+1
Y:=Y+1
E2,Y:=E2,X
X=0
no
no
X=255
yes
yes
Y:=Y-1
E1,Y:=0
Y:=Y+1
E2,Y:=0
Y=0
no
no
Y=255
yes
yes
CENTRAL

subroutine .LOC

LOC defines the blocks in which the zero, first and second orders are located. The results are stored in E41-E46. E49=0 if the routine fails.

LABEL	OPERATION	BYTES	CYCLES
.LOC	CLD	1	2
	LDA #&1	2	2
	STA E47	3	4
.LOC1	LDY E47	3	4
	LDA E2,Y	3	4
	STA E48	3	4
	CLC	1	2
	LDA E47	3	4
	ADC #&20	2	2
	STA E49	3	4
	LDX #&0	2	2
.LOC4	LDA E48	3	4
	CMP E2,X	3	4
	BCS LOC2	2	2,3
.LOC3	INX	1	2
	CPX E47	3	4
	BNE LOC4	2	2,3
.LOC7	LDA E2,X	3	4
	CMP E48	3	4
	BCS LOC5	2	2,3
	JMP LOC2	3	3
.LOC5	INX	1	2
	CPX E49	3	4
	BNE LOC7	2	2,3
	JMP LOC6	3	3
.LOC2	INC E47	1	2
	LDA E47	3	4
	CMP #&50	2	2
	BCS NLOC	2	2,3
	JMP LOC1	3	3
.LOC6	LDA E47	3	4
	STA E44	3	4
	LDA #&FC	2	2
	STA E47	3	4
.LOC8	LDY E47	3	4
	LDA E1,Y	3	4
	STA E48	3	4
	SEC	1	2
	LDA E47	3	4
	SBC #&20	2	2
	STA E49	3	4
	LDX E49	3	4
.LOC11	LDA E48	3	4
	CMP E1,X	3	4
	BCS LOC9	2	2,3
.LOC10	INX	1	2
	CPX E47	3	4
	BNE LOC11	2	2,3
.LOC14	LDA E1,X	3	4
	CMP E48	3	4
	BCS LOC12	2	2,3
	JMP LOC9	3	3
.LOC12	INX	3	3
	CPX #&FF	2	2
	BNE LOC14	2	2,3
	JMP LOC13	3	3
.LOC9	DEC E47F	3	6
	LDA E47	3	4
	CMP #&C8	2	2
	BEQ NLOC	2	2,3
	JMP LOC8	3	3
.NLOC	LDA #&0	2	2
	STA E49	3	4
	RTS	1	6
.LOC13	LDA E47	3	4
	STA E43	3	4
	SEC	1	2
	LDA E44	3	4
	SBC E43	3	4
	STA E49	3	4
	ROR A	1	2
	CLC	1	2
	ROR A	1	2
	ADC E49	3	4
	STA E49	3	4
	SEC	1	2
	LDA E43	3	4
	SBC E49	3	4
	STA E42	3	4

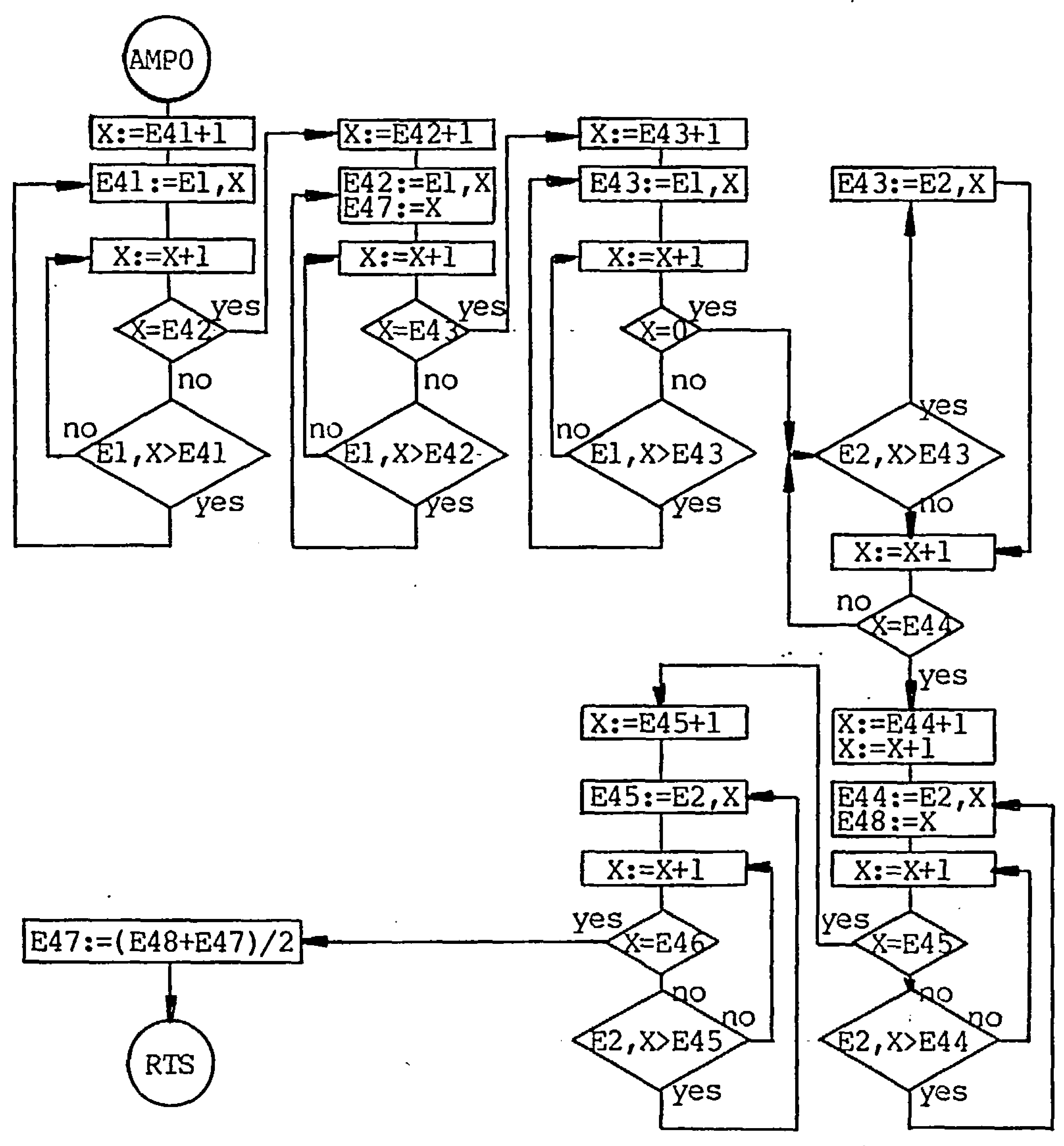
LABEL	OPERATION	BYTES	CYCLES
	SBC E49	3	4
	STA E41	3	4
	CLC	1	2
	LDA E44	3	4
	ADC E49	3	4
	STA E45	3	4
	ADC E49	3	4
	STA E46	3	4
.LO45	LDX E45	3	4
	LDA E2,X	3	4
	STA E48	3	4
	INX	1	2
	LDA E2,X	3	4
	STA E49	3	4
	DEX	1	2
	DEX	1	2
	LDA E2,X	3	4
	STA E47	3	4
	INX	1	2
	LDA E47	3	4
	CMP E48	3	4
	BCS LO2	2	2,3
	LDA E49	3	4
	CMP E48	3	4
	BCS LO3	2	2,3
	LDA E47	3	4
	CMP E49	3	4
	BCS LO4	2	2,3
.LO3	DEC E45	3	6
	JMP LO45	3	3
.LO2	LDA E49	3	4
	CMP E48	3	4
	BCS LO5	2	2,3
.LO4	INC E45	3	6
	JMP LO45	3	3
.LO5	SEC	1	2
	LDA E45	3	4
	SBC E44	3	4
	STA E49	3	4
	CLC	1	2
	LDA E45	3	4
	ADC E49	3	4
	STA E46	3	4
.LO42	LDX E42	3	4
	LDA E1,X	3	4
	STA E48	3	4
	INX	1	2
	LDA E1,X	3	4
	STA E49	3	4
	DEX	1	2
	DEX	1	2
	LDA E1,X	3	4
	STA E47	3	4
	INX	1	2
	LDA E47	3	4
	CMP E48	3	4
	BCS L42	2	2,3
	LDA E49	3	4
	CMP E48	3	4
	BCS L43	2	2,3
	LDA E47	3	4
	CMP E49	3	4
	BCS L44	2	2,3
.L43	DEC E42	3	6
	JMP LO42	3	3
.L42	LDA E49	3	4
	CMP E48	3	4
	BCS L45	2	2,3
.L44	INC E42	3	6
	JMP LO42	3	3
.L45	SEC	1	2
	LDA E43	3	4
	SBC E42	3	4
	STA E49	3	4
	SEC	1	2
	LDA E42	3	4
	SBC E49	3	4
	STA E41	3	4
	RTS	1	6

LOC
E47:=1
Y:=E47
E48:=E2,Y
E49:=E47+32
X:=0
E48>E2,X
yes
no
E47:=E47+1
E47>80
no
yes
X:=X+1
X=E47
no
yes
E2,X>E48
no
yes
X:=X+1
X=E49
no
yes
E49:=0
RTS
E44:=E47
E47:=252
Y:=E47
E48:=E1,Y
E49:=E47-32
X:=E49
E48>E1,X
yes
no
E47:=E47-1
E47=200
no
yes
X:=X+1
X=E47
no
yes
E1,X>E48
no
yes
X:=X+1
X=255
no
yes
E43:=E47
E49:=E44-E43
E42:=E44-E43
E41:=E42-E49
E45:=E44+E49
E45:=E45+E49
X:=E45
E48:=E2,X
X:=X+1
E49:=E2,X
X:=X-2
E47:=E2,X
X:=X+1
E47>E48
yes
no
E49>E48
yes
no
E49:=E45-E44
E46:=E45+E49
E49>E48
yes
no
E45:=E45-1
E47>E49
no
yes
E45:=E45+1
X:=E42
E48:=E1,X
X:=X+1
E49:=E1,X
X:=X-2
E47:=E1,X
X:=X+1
E47>E48
yes
no
E49>E48
yes
no
E49:=E43-E42
E41:=E42-E49
RTS
E42:=E42-1
E49>E48
yes
no
E47>E49
no
yes
E42:=E42+1

subroutine .AMP0

AMP0 calculates the local maxima in the blocks defined by .LOC and calculates the spacing between the two first order maxima. The amplitudes are stored in E41-E45. The spacing is stored in E47.

```
LABEL    OPERATION  BYTES CYCLES

.AMP0    LDX E41      3     4
         INX          1     2
         LDA E1,X     3     4
.AMP1    STA E41      3     4
.AMP3    INX          1     2
         CPX E42      3     4
         BEQ AMP2     2     2,3
         LDA E1,X     3     4
         CMP E41      3     4
         BCS AMP1     2     2,3
         JMP AMP3     3     3
.AMP2    LDX E42      3     4
         INX          1     2
         LDA E1,X     3     4
.AMP5    STA E42      3     4
         STX E47      3     4
.AMP6    INX          1     2
         CPX E43      3     4
         BEQ AMP4     2     2,3
         LDA E1,X     3     4
         CMP E42      3     4
         BCS AMP5     2     2,3
         JMP AMP6     3     3
.AMP4    LDX E43      3     4
         INX          1     2
         LDA E1,X     3     4
.AMP8    STA E43      3     4
.AMP9    INX          1     2
         CPX #&0      2     2
         BEQ AMP7     2     2,3
         LDA E1,X     3     4
         CMP E43      3     4
         BCS AMP8     2     2,3
         JMP AMP9     3     3
.AMP7    LDA E2,X     3     4
         CMP E43      3     4
         BCS AMP10    2     2,3
         JMP AMP11    3     3
.AMP10   STA E43      3     4
.AMP11   INX          1     2
         CPX E44      3     4
         BEQ AMP12    2     2,3
         JMP AMP7     3     3
.AMP12   LDX E44      3     4
         INX          1     2
         LDA E2,X     3     4
.AMP14   STA E44      3     4
         STX E48      3     4
.AMP15   INX          1     2
         CPX E45      3     4
         BEQ AMP13    2     2,3
         LDA E2,X     3     4
         CMP E44      3     4
         BCS AMP14    2     2,3
         JMP AMP15    3     3
.AMP13   LDX E45      3     4
         INX          1     2
         LDA E2,X     3     4
.AMP17   STA E45      3     4
.AMP18   INX          1     2
         CPX E46      3     4
         BEQ AMP16    2     2,3
         LDA E2,X     3     4
         CMP E45      3     4
         BCS AMP17    2     2,3
         JMP AMP18    3     3
.AMP16   SEC          1     2
         LDA E48      3     4
         SBC E47      3     4
         CLC          1     2
         ROR A        1     2
         STA E47      3     4
         RTS          1     6
```


AMPO
X:=E41+1
E41:=E1,X
X:=X+1
X=E42
yes
no
E1,X>E41
no
yes
X:=E42+1
E42:=E1,X
E47:=X
X:=X+1
X=E43
yes
no
E1,X>E42
no
yes
X:=E43+1
E43:=E1,X
X:=X+1
X=0
yes
no
E1,X>E43
no
yes
E43:=E2,X
E2,X>E43
yes
no
X:=X+1
X=E44
no
yes
X:=E44+1
X:=X+1
E44:=E2,X
E48:=X
X:=X+1
X=E45
yes
no
E2,X>E44
no
yes
X:=E45+1
E45:=E2,X
X:=X+1
X=E46
yes
no
E2,X>E45
no
yes
E47:=(E48+E47)/2
RTS

ASSEMBLER MNEMONICS FOR 6502

ADC	ADd with Carry	A,C=A+M+C
AND	logical AND	A=A AND M
ASL	Arithmetic Shift Left	M=M*2, C=M7 (or accumulator)
BCC	Branch on Carry Clear	Branch if C=0
BCS	Branch on Carry Set	Branch if C=1
BEQ	Branch on result EQuals zero	Branch if Z=1
BIT	test memory BITs with accumulator	A AND M, N=M7, V=M6
BMI	Branch if negative flag set	Branch if N=1
BNE	Branch if result Not Equal zero	Branch if Z=0
BPL	Branch on Positive resuLt	Branch if N=0
BRK	forced interrupt	PC and P pushed on stack
BVC	Branch if oVerflow clear	Branch if V=0
BVS	Branch if oVerflow set	Branch if V=1
CLC	CLear Carry flag	C=0
CLD	CLear Decimal flag	D=0
CLI	CLear Interrupt disable flag	I=0
CLV	CLear the oVerflow flag	V=0
CMP	CoMPare memory and accumulator	A-M
CPX	ComPare memory with X register	X-M
CPY	ComPare memory with Y register	Y-M
DEC	DECrement memory by one	M=M-1
DEX	DEcrement X register by one	X=X-1
DEY	DEcrement Y register by one	Y=Y-1
EOR	Exclusive OR memory with accumulator	A=A EOR M
INC	INCrement memory by one	M=M+1
INX	INcrement X register by one	X=X+1
INY	INcrement Y register by one	Y=Y+1
JMP	JuMP to new location	PC=new address
JSR	Jump SubRoutine	Save current PC;PC=new address
LDA	LoaD Accumulator from memory	A=M
LDX	LoaD X register from memory	X=M
LDY	LoaD Y register from memory	Y=M
LSR	Logical Shift Right by one bit	M=M/2 (or A)
NOP	No OPeration	
ORA	OR memory with Accumulator	A=A OR M
PHA	PusH Accumulator onto stack	Push A

PHP	PusH status register onto stack	Push P
PLA	PulL Accumulator off stack	Pull A
PLP	PulL status register off stack	Pull P
ROL	ROtate one bit Left	M=M*2, M0=C, C=M7 (A or M)
ROR	ROtate one bit Right	M=M/2, M7=C, C=M0 (A or M)
RTI	ReTurn from Interrupt	Status register and PC returned
RTS	ReTurn from Subroutine	Pull PC from stack
SBC	SuBstract mem. from accu. with Carry	A,C=A-M-(1-C)
SEC	SEt Carry flag	C=1
SED	SEt Decimal mode	D=1
SEI	SEt Interrupt disable flag	I=1
STA	STore Accumulator contents in memory	M=A
STX	STore X contents in memory	M=X
STY	STore Y contents in memory	M=Y
TAX	Transfer A to X	X=A
TAY	Transfer A to Y	Y=A
TSX	Transfer S to X	X=S
TXA	Transfer X to A	A=X
TXS	Transfer X to S	S=X
TYA	Transfer Y to A	A=Y

ASSEMBLER MNEMONICS FOR 6502

Processor status after use. Only the effected ones are listed.

C=carry flag
Z=zero flag
I=interrupt flag
D=decimal flag
B=break command
V=overflow flag
N=negative flag

ASL	C set to old contents of bit 7
	Z set if result=0
	N set if bit 7 of the result is set
BIT	Z set if the result=0
	V set to bit 6 of memory
	N set to bit 7 of memory
BRK	B set
CLC	C cleared
CLD	D cleared
CLI	I cleared
CLV	V cleared
CMP	C set if A greater than or equal to M
	Z set if A=M
	N set if bit 7 of the result is set
CPX	C set if X greater than or equal to M
	Z set if X=M
	N set if bit 7 of the result is set
CPY	C set if Y greater than or equal to M
	Z set if Y=M
	N set if bit 7 of the result is set
DEC	Z set if memory contents becomes 0
	N set if bit 7 of the result is set
DEX	Z set if X becomes 0
	N set if bit 7 of X becomes set
DEY	Z set if Y becomes 0
	N set if bit 7 of Y becomes set
EOR	Z set if A becomes 0
	N set if bit 7 of A becomes set

INC	Z set if memory contents becomes 0
	N set if bit 7 of memory becomes set
INX	Z set if X becomes 0
	N set if bit 7 of X becomes set
INY	Z set if Y becomes 0
	N set if bit 7 of Y becomes set
LDA	Z set if A=0
	N set if bit 7 of A set
LDX	Z set if X=0
	N set if bit 7 of X set
LDY	Z set if Y=0
	N set if bit 7 of Y set
LSR	C set to bit 0 of operand
	Z set if result=0
	N cleared
ORA	Z set if A=0
	N set if bit 7 of A set
PLA	Z set if A=0
	N set if bit 7 of A set
PLP	C bit 0 from stack
	Z bit 1 from stack
	I bit 2 from stack
	D bit 3 from stack
	B bit 4 from stack
	V bit 6 from stack
	N bit 7 from stack
ROL	C set to old value of bit 7
	Z set if result=0
	N set if bit 7 of the result is set
ROR	C set to old value of bit 0
	Z set if result=0
	N set if bit 7 of the result is set
RTI	C bit 0 from stack
	Z bit 1 from stack
	I bit 2 from stack
	D bit 3 from stack
	B bit 4 from stack
	V bit 6 from stack
	N bit 7 from stack
SBC	C cleared if borrow occurs

Z set if result=0

V set if the sign of the result is wrong

N set if bit 7 of the result is set

SEC C set

SED D set

SEI I set

TAX Z set if X becomes 0

N set if bit 7 of X is set

TAY Z set if Y becomes 0

N set if bit 7 of Y is set

TSX Z set if X becomes 0

N set if bit 7 of X is set

TXA Z set if A becomes 0

N set if bit 7 of A is set

TYA Z set if A becomes 0

N set if bit 7 of A is set

ASSEMBLER MNEMONICS FOR 6502

Instruction	Addressing mode	Bytes	Cycles
BCC	relative	2	2(+1 if branch succeeds +2 if to new page)
BCS	relative	2	2(+1 if branch succeeds +2 if to new page)
BEQ	relative	2	2(+1 if branch succeeds +2 if to new page)
BIT	zero page	2	3
	absolute	3	4
BMI	relative	2	2(+1 if branch succeeds +2 if to new page)
BNE	relative	2	2(+1 if branch succeeds +2 if to new page)
BPL	relative	2	2(+1 if branch succeeds +2 if to new page)
BRK	implied	1	7
BVC	relative	2	2(+1 if branch succeeds +2 if to new page)
BVS	relative	2	2(+1 if branch succeeds +2 if to new page)
CLC	implied	1	2
CLD	implied	1	2
CLI	implied	1	2
CLV	implied	1	2
CPX	immediate	2	2
	zero page	2	3
	absolute	3	4
CPY	immediate	2	2
	zero page	2	3
	absolute	3	4
DEX	implied	1	2
DEY	implied	1	2
INX	implied	1	2
INY	implied	1	2
JMP	absolute	3	3
	indirect	3	5
JSR	absolute	3	6
NOP	implied	1	2
PHA	implied	1	3
PHP	implied	1	3
PLA	implied	1	4
PLP	implied	1	4
RTI	implied	1	6
RTS	implied	1	6
SEC	implied	1	2
SED	implied	1	2

SEI	implied	1	2
STX	zero page	2	3
	zero page,Y	2	4
	absolute	3	4
STY	zero page	2	3
	zero page,X	2	4
	absolute	3	4
TAX	implied	1	2
TAY	implied	1	2
TSX	implied	1	2
TXA	implied	1	2
TXS	implied	1	2
TYA	implied	1	2

ASSEMBLER MNEMONICS FOR 6502

Instruction	ADC	AND	ASL	CMP	DEC	EOR	INC	LDA	LDX	LDY	LSR	ORA	ROL	ROR	STA
immediate	2,2	2,2		2,2		2,2		2,2	2,2	2,2		2,2			
zero page	2,3	2,3	2,5	2,3	2,5	2,3	2,5	2,3	2,3	2,3	2,5	2,3	2,5	2,5	2,3
zero page,X	2,4	2,4	2,6	2,4	2,6	2,4	2,6	2,4		2,4	2,6	2,4	2,6	2,6	2,4
zero page,Y									2,4						
absolute	3,4	3,4	3,6	3,4	3,6	3,4	3,6	3,4	3,4	3,4	3,6	3,4	3,6	3,6	3,4
absolute,X	3,4*	3,4*	3,7	3,4*	3,7	3,4*	3,7	3,4*		3,4*	3,7*	3,4*	3,7	3,7	3,5
absolute,Y	3,4*	3,4*		3,4*		3,4*		3,4*	3,4*			3,4*			3,5
(indirect,X)	2,6	2,6		2,6		2,6		2,6				2,6			2,6
(indirect),Y	2,5*	2,5*		2,5*		2,5*		2,5*				2,5*			2,6
accumulator			1,2								1,2		1,2	1,2	

*=(+1 if page crossed)

a,b=bytes,cycles

Operand fields for various addressing modes:

addressing mode	operand	examples
Implied	n/a	RTS
Accumulator	A	ROL A
Immediate	operand preceeded by #	LDA #&FF LDX #count
Absolute	absolute address	CMP &1900 JMP label
Zero page	the absolute address is in the zero page	CPY &80
Indirect	the address is stored in the (memory)	JMP (&1900)
Absolute,X or Y	the address is an absolute address +XorY	LDA &2800,X ADC table,Y
Zero page,X	using zero page address +X	LDX &72,X
Pre indexed	indirect addressing using a table of addresses in zero page	LDA (&80,X)
Post indexed	indirect addressing using an indirect address in zero page plus offset in Y	STA (&80),Y
Relative	has to be used with branch instructions. The maximum jump forward is 127 bytes. The maximum jump backwards is 128 bytes.	BNE label