CoA/N/M&P-2

CoA Note M & P No. 2.

ST. NO. R27664
U.D.C. Copy A.
AUTH.

THE COLLEGE OF AERONAUTICS

CRANFIELD

COLLEGE OF AERONAUTICS LIBRARY

DIFFRACTION GRATINGS - THEIR PRINCIPLES AND APPLICATIONS TO MACHINE TOOLS

by

M. Healey

R27664

CoA Note M and P No. 2

December, 1963

3 8006 10057 9187

THE COLLEGE OF AERONAUTICS

DEPARTMENT OF PRODUCTION AND INDUSTRIAL ADMINISTRATION

Instrumentation and automation section

Diffraction gratings

their principles and applications to machine tools

- by -

M. Healey, B.Sc. (Eng.)

SUMMARY

The use of diffraction gratings as measuring elements on machine tools is nowadays an accepted procedure. Part I of this paper therefore sets out to give a brief understanding of the principles of diffraction gratings, and the ways in which they have been utilised commercially. In addition, comparisons are made with other reference elements and some consideration is given to manufacturing techniques.

Part II of the paper deals in more detail with the pure theory of crossed diffraction gratings.

Contents

PART 1

Introduction

A rough screen will distribute light by refraction or reflection, in a random manner. If, however, the roughness has a regularly repeated pattern, so that the distributed light is regimented, then this is called a diffraction grating. The simplest form of diffraction grating is a piece of glass printed with regularly spaced opaque lines leaving regular transparent gaps, so that collimated light shone through is split into regular beams. This is known as a line and space grating.

Straight away, it must be made clear that light is not a number of 'beams', but is a high frequency energy wave which has amplitude and phase. It is the interaction of light waves of differing phase which, in fact, causes diffraction patterns in crossed gratings.

However, if, initially, we consider coarse pitch gratings, the approximation of beams of light, i.e. ray optics, will suffice.

As a means of detecting movement, gratings must be used in pairs, the actual movement then detected being the relative movement of the two gratings. The changes in light intensity are measured by a photo-cell and from a practical point the index grating need only be as large as the photo-cell.

Parallel line and space gratings

If the lines are ruled with even mark-space ratio, and index and reference gratings have their lines parallel, (figure 1), then, when the lines are coincident, the maximum amount of light will pass through the gratings; when the lines on one grating and spaces on the other are coincident, the minimum amount of light will pass through. The photo-cell will then give an output as in figure 3. Measurement of the d.c. level of this signal will then indicate the relative position within one ruling pitch of movement. It is an obvious limitation of diffraction grating systems that the system is 'incremental'. In other words, by counting the number of cycles of the photo-cell output, the displacement can be measured, but all measurements must be relative to some datum. One point on the grating cannot be distinguished from another an exact number of pitches away.

Further, since the photo-cell output is in practice not of a very precise waveform, measurement of d.c. values makes it impractical to divide the fringe by more than 4. However, using a grating of say 1000 lines/inch, a movement of 1 inch will generate 1000 cycles from which a pulse counter can indicate a readout in thousandths of an inch. The problem now arises in that moving the index back to the datum point, another 1000 pulses will be counted, leaving an indicated 2000 instead of 0. Hence, some means of sensing direction of movement is required.

Moiré Fringes

If the rulings of the grating and index are now inclined slightly, an interference pattern is formed which is known as a Moiré Fringe, figure 2. These fringes are of a serrasoidal distribution over a pitch, at right angles to the rulings, which is proportional to the ruling pitch and inversely proportional to the angle of inclination.

If the gratings are now moved as before in a direction perpendicular to the ruled lines, the fringes move in a direction at right angles to the direction of movement of the gratings. More importantly, the fringe pattern will move up for one direction of grating movement, and down for an opposite grating movement. The fringe will move one pitch for a movement of the grating of one ruled pitch.

The amplification can be made very high by decreasing the angle of inclination, although a practical limit is set by the clarity of the fringe pattern and the degree of amplification which is useful. Also, direction sensing is provided by detecting whether the fringe is moving up or down. To do this, at least two photo-cells are required, spaced $^1/_4$ of a fringe apart, (figure 4); both photo-cells will generate serrasoidal voltages, as the gratings are moved, with a 90° phase shift between them.

This means that photo-cell A is either leading or lagging photo-cell B, dependent upon the direction of motion, which can be utilised to trigger add or subtract circuits in the pulse counter. In practice, since errors may now be introduced by the spacing of the fringe and positioning of the photo-cells, 4 photo-cells are used, displaced $^1/_4$ of a fringe apart, with alternate photo-cell voltages subtracted, giving an average effect and also eliminating the d.c. component of the signal.

Four phase index pieces

Instead of using Moiré fringes, two or four phase signals can be generated using parallel rulings with an index piece comprising 4 sections, the ruled lines of which are $^1/_4$ of a pitch displaced from each other.

Vernier patterns

It is obvious that Vernier effects can be produced by using two gratings of different pitch, but these have little practical application to machine tools.

Basic methods of using gratings on machine tools

Diffraction gratings have been employed in two types of machine.

1. A co-ordinate measuring machine,[8] in which the table of the machine is set to a datum point, all counters are set to zero and movement relative to the datum is indicated directly. Moiré fringes are utilised and division

of one fringe by 2 is achieved, using 4 photo-cells, so that machines are marketed using 1000 lines/inch gratings, to give a readout with a least significant figure of .0005". A single axis arrangement for attaching to m/c tools using 2500 lines/inch gratings and dividing by 4 giving a readout in tenths of a thou. is also available.

2. A numerically controlled machine tool in which the machine is, as always, set to datum and then fed with information from punched or magnetic tape. Pulses from the tape are fed into a store and pulses from the gratings are subtracted from this store. The servo drives the table to reduce the total number of pulses in the store to zero. The maximum number of pulses in the store at any instant indicates the maximum error in the system. This system is difficult to engineer since reliable bi-input, or synchronous, reversible counters are far more involved than single input counters as used in measuring machines. It was used, however, on the Ferranti Mk III system now replaced by the phase modulated Mk IV system.

Phase modulated systems[5]

To increase the accuracy of a system of this type, it is necessary to either increase the number of lines/inch or to subdivide each fringe more accurately. The finer the pitch of the line, the more prone the system is to miss a pulse, or for reversible counters not to reverse in time.

To subdivide the fringe more accurately has many advantages. For example, if the fringe is divided by 100, then for 1 tenth of a thou. resolution, we require gratings of 100 lines/inch. This has an advantage which will be mentioned later in that steel gratings can be used, with an improvement in ruggedness and co-efficient of expansion. Further, the system can move half a fringe, i.e., 5 thou. without being ambiguous, which would allow a much higher rate of fast traverse.

So the problem to be faced is to divide one fringe accurately by up to 100. It has already been stated that, by measuring the d.c. level of the photo-cell signals, inaccuracies, due to the poor waveform and d.c. amplification, limit the direct subdivision of the fringe to 4 at the most.

The problem is overcome by converting the system from a static spatial problem to a temporal one. In other words, continuous a.c. signals are generated from a fixed index and a movable grating attached to the slide and these two signals compared in phase.

Consider figure 5. The drum is ruled around its periphery, with a pitch equal to that of the gratings, and rotated at a frequency w. If the movable grating is stationary, both gratings will generate serrasoidal signals of similar frequency Nw, where N is the number of lines around the drum. The waveforms will be similar, but the phase will depend upon the position of the movable grating. Measurement of the phase of these two signals to 1/100th of a cycle is easily possible, thus achieving the desired aim.

This technique is utilised in two ways:-

a. The Ferranti Mark IV system[9]. Figure 6

Here the tape carries a reference signal to which is synchronised the drum motor. Feedback of the signal from the fixed grating, w_f, will eliminate errors due to stretching of the tape, etc. Another track on the same tape carries information about the required table position and the error signal, derived from comparing the phase of this signal with that from the movable grating drives the table. To keep the table still, the reference frequency w_r and the information frequency w_m will be equal; to move the table one way, w_m will be higher in frequency than w_r and lower in frequency to move the table in the opposite direction. The velocity of the table is proportional to $w_m \pm w_r$. Since the tape information must be continuous, magnetic tape is used, and the signal is printed in pulse rather than serrasoidal form. Also, since a pulse cannot be intentionally missing, flywheel synchronisation circuits can be employed, eliminating errors due to occasional misreading of the tape. In practice a spiral ruled disc is placed against the grating with its axis of rotation perpendicular to the grating instead of using a drum, (figures 7a, 7b and 7c).

b. The Staveley system[6]

In this system, the continuous a/c signal is achieved by cyclicly sampling the 4 photo-cells deployed across the fringe, the phase of the a.c. signal being dependent upon the relative position of the fringe. The reference channel on the tape would provide the switching frequency, thus eliminating the need to synchronise a rotating drum.

Steel gratings

Recent developments have led to gratings being produced on thin strips of stainless steel, the lines being acid etched, the grating then comprising alternate absorbing and reflecting lines. The reading heads are necessarily more complicated optically, due to the photo-cell and light source being situated on the same side of the grating. The advantages are mechanical durability and a temperature coefficient of expansion similar to that of the machine tool. These gratings are limited in pitch, as the granular structure of the steel is exposed by the etching causing the edge of the line to blurr. 1000 lines/inch is the maximum, but it has already been pointed out that phase modulated systems using 100 and 500 lines/inch are in practical use.

Radial gratings

Radial gratings are at present made on glass as direct exposure copies of precision ruled masters. The masters are available in a limited number of total lines ranging from 100 to 64,800 lines per 360°. Further, the accuracy of some of these configurations is better than others. Typically, a 1000 line disc has a maximum error of 2 seconds of arc. Errors further to this are introduced by eccentricity, a source of error which, however, can be reduced by using two reading heads placed diametrically opposite. A photographic 'integrating' copying technique, which, it is hoped, will increase the accuracy, is being developed by the N.P.L. and the N.E.L.

A summary of some applications of diffraction gratings

i) Spectrometry - the use of phase gratings, (figure 9a), of a very fine pitch to isolate the different frequency components of a light source.

ii) Linear measurement - the use of photographic line and space gratings to indicate the distance from some datum, either by direct readout of one electrical cycle per ruled pitch or by subdivision, usually by phase modulation.

iii) Reference elements for control systems - as (ii), except that error signals between the output from the grating and an input reference, drive a servo to a null condition.

iv) Gear error measurement[13] - the use of two radial gratings and a master gear to measure the accuracy of transmission of a gear. The master and test gear are correctly meshed and one gear is driven. Similar radial gratings are attached to each gear shaft. If the master gear has m teeth and the test gear n teeth, then the a.c. signal from the test gear grating is frequency multiplied by n and frequency divided by m. The master gear signal and the modified test gear signal will now be of equal frequency and any difference in phase will be a measure of errors in motion due to errors in the test gear. Known errors in the master gear can be calibrated out.

v) Strain measurement - the use of the grating to measure stretching in strain testing. Radial gratings can also be used to measure torque, particularly under dynamic conditions. A grating is placed at each end of a twistable shaft, which is transmitting the torque to be measured. Twisting of the shaft, will cause phase differences between the signals from the two gratings which can be calibrated in terms of torque.

Manufacture of diffraction gratings

a) Phase gratings:[11]

Phase gratings, having a prismatic form as shown in figure 8, have been used for spectrometry for a long time. Such gratings are very fine in pitch ranging from 1000 to 50,000 lines per inch. They are used singly, and light passing through them will split into various order beams dependent upon the pitch of the grating and more important the wave length of the light. The grating can therefore be used to measure length relative to the wavelength of light. They are made by impressing polished brass cylinders of 1 or 2 inches diameter with a diamond tool on a lathe fitted with a control system, in which the tool movement is dependent upon the mean pitch of a previously directly machined cylinder. The cylinder is then coated with a Polystyreme skin about 1 thou. thick, which is split axially, peeled off and laid out flat. This is called a 'pelicle'.

A skin of gelatine about $1\frac{1}{2}$ thou. thick is moulded onto an optical flat using a 2nd optical flat to form the top surface. The pelicle is then 'floated' onto the gelatine by surface tension and immersed in a 2% solution of Ammonium Dichromate which renders the gelatine liable to hardening by exposure to U.V.L. After soaking, the flat, gelatine and pelicle are dried very slowly in a dark box. During this period, with no applied pressure, 'plastic flow' of the gelatine occurs moulding it to the shape of the pelicle.

The pelicle is peeled off and the gelatine exposed, thereby being hardened.

Acrylic resin is now poured on the gelatine and a sheet of glass placed on top. When set, the gelatine mould can be broken from the resin copy.

1 cylinder makes approximately 120 pelicles
1 pelicle makes 2 or 3 gelatine moulds
1 gelatine mould makes about 10 gratings

b) Amplitude gratings:[12]

Line and space gratings are made as photographic copies of a precision ruled master. However, an integrating of the errors during this process which will be described hereafter, results in copies of greater accuracy than the originals. Direct contact printing is possible for gratings of lower accuracy.

Flat plates of glass usually about 10″ long, although up to 36″ lengths are available without joining, are coated with a photosensitive emulsion. This is placed end to end with the master copy on a precision carriage. One index piece is employed in conjunction with the master to produce a reference signal. This signal triggers a flash tube placed above a second index piece which in turn is placed above the emulsioned glass plate. Let us say there are 40 lines exposed under the flash, then each individual flash should perform $\frac{1}{40}$th of the total required exposure. This integrating effect will reduce any short term errors present in the master. Long term errors are calibrated in the master and corrected by hand, using a phase shifter in the flash trigger circuit. Using repeated copying, it is now possible to produce gratings which are accurate at all points over a 10″ length to ± 20 micro inches.

Refinements incorporated in some commercial gratings are chrome plated lines in which the exposed grating is soaked in an agent which will dissolve only the unexposed emulsion. The grating is then plated and a second agent used which dissolves away the remaining emulsion leaving chrome lines only, which are more durable.

Comparison of diffraction gratings with other forms of reference systems

The following is a brief description of alternative systems in use.

i) End position indicators - usually used on a lead screw with a geared coarse positioning element. These devices are in the form of comutators, one segment of which is excited and the system driven until this segment is in contact with a datum brush, thus producing a stop signal. They are not as complex or as expensive as diffraction gratings nor are they as accurate. They are not suitable for continuous control.

ii) Resistance potentiometers - the simplest absolute system, it is used only for coarse positioning, due to low accuracy and contact problems.

iii) Inductive potentiometers - effectively a tapped toroidally wound transformer the tap connection being made via a rotary switch geared to the machine lead screw. This is limited by the size of the tap increment. Fine positioning is therefore achieved by using a continuous rotary inductive potentiometer such as a Linvar which is constructed similar to a synchro and gives a continuous output voltage proportional to angular position over 80°. The output voltage from the Linvar is added to that from the tapped transformer. This system is absolute and was used by E.M.I. but has since been superseded by other systems.

iv) Capacitive potentiometers - comprise two concentric cylinders one much shorter than the other. An a.c. voltage is applied across the ends of the long cylinder so that a voltage proportional to displacement is developed between the short cylinder and one end of the long cylinder. As with the other forms of potentiometer the voltage from the potentiometer is compared with a voltage from a precision tapped transformer in a simple bridge. This system is completely absolute. It is made by Sogenique in a form where the reference element comprises 1 inch long sections fed at specific voltages with the shorter element covering two sections. Fine adjustment of the bar voltages results in accuracies better than .0001". This system is as yet in limited use, its initial application being as a displacement indicator with digital readout. It should however lead to excellent results in continuous control mechanisms. Circular forms are also being developed.

v) Synchros and resolvers - available with many variations and other names such as selsyns and magslips. A single phase, a.c. excited rotor develops voltages in a 3 phase stator winding, of the same frequency and phase as the input voltage, but of amplitudes proportional to Sin θ, Sin $(\theta + 120)$ and Sin $(\theta - 120)$ where θ is the rotor angle. Conversely voltages of V Sin θ Sin wt; V Sin $(\theta + 120)$ Sin wt and V Sin $(\theta - 120)$ Sin wt fed to the stator windings will set up an alternating field such that no voltage will be induced in the rotor when it is 90° from θ. Thus the rotor voltage can be used as an error signal for a servo. A resolver is basically similar but with a 2 phase rotor and stator. If the two stator windings are fed with signals of the same amplitude but with 90° phase difference, a rotating field will be set up, so that the rotor voltages will be of constant amplitude but of phase dependant upon the rotor position θ. Thus the two phase rotor voltages will be V Sin $(wt + \theta)$ and V Cos $(wt + \theta)$ where w is the excitation frequency. This is used in a phase comparison system. Used to measure leadscrew displacement, accuracies, is closed loops, of 10 minutes of arc can be achieved. Thus position accuracies will be about .0005". Used with coarse and fine ranges, these systems are absolute. They are cheaper than diffraction gratings but not so accurate in terms of linear displacement. Used on some of the I.G.E. continuous control systems but mainly for position controllers such as the I.G.E. and E.M.I. They are however being superseded for the simple position controller by E.P.I.'s as in i).

vi) Linear resolvers - these are similar to synchro-resolvers in principle except that they are printed on linear scales, so that one pitch, usually about 0.1", will give one cycle of electrical output. These devices are therefore analogue over one pitch and must be used in connection with coarse position indicators. The accuracy can be .0001". Marketed with basic differences in construction by E.M.I. as the 'Farrand Industosyn' (inductively coupled) and A.E.I. as the 'Helixyn' (capacitively coupled). These elements are directly competative with diffraction gratings in cost and accuracy and relative merits lie more in the cost and versatility of associated computing equipment.

vii) Variable reluctance elements - precision ground steel pins .1" diameter are set in a non-magnetic base in a comb like structure. A C-core head is fixed so that the pins can move in the gap, thereby varying the inductance of a coil wound on the C-core. The maximum inductance will occur when the centre line of the head coincides with a diameter of a pin and the minimum inductance when the centre line of the head coincides with the 'valley' between pins. In fact 4 heads are used arranged in a bridge. Accuracy can be as high as .0001". A circular version is also available. Manufactured by I.G.E. as the 'Accupin', it is comparable for cost and accuracy with linear resolvers (vi) and diffraction gratings.

viii) Encoders - usually, particularly for high resolution, in circular form they comprise non-conducting discs printed with conducting segments in such a manner as to read off a binary number proportional to the angular position. Cyclic binary coes are used and in some cases optical systems, having less contact problems, are employed. Somewhat similar to synchros in application except that position is indicated directly by the binary number. Digital shaft encoders such as the Rotax or Hilger and Watts products are available with up to $2\frac{1}{2}$ seconds of arc digit size and 3 seconds of arc error. They are as yet limited in application due to high cost; a Renault machine tool fitted with such a system being one of the first examples. Almost comparable in accuracy with radial diffraction gratings, although far more expensive, they can be used in incremental or absolute systems.

Rotary or Linear Transducers?

Control of lead screw displacement has so far proved the cheapest of machine tool control systems. Its disadvantage is that leadscrew errors and wind up are not allowed for. Leadscrew control of 10 minutes of arc equivalent to a nominal .0001" in practice results in about .001" accuracy, particularly as the machine ages. Linear transducers are more expensive but more accurate, particularly in that they include leadscrew errors etc. inside the control loop. This complicates the servo somewhat but reduces deteriation due to machine wear. Finally the use of hydraulic ram drives means that rotary transducers must have a special leadscrew or precision rack.

PART 2

The theory of crossed diffraction gratings

It has so far been explained what diffraction gratings are and how they are used by thinking of light as travelling in beams or rays.

It is now necessary to explain why the spacing of two gratings is critical and also the effects of angle of incidence of the light source, at which stage the beam of light theory falls down. Also, gratings other than line and space types can be considered.

Consider a piece of glass with a regular surface pattern of pitch w struck by a plane wave of light of wave length λ at an angle of incidence i (figure 8). Light waves refracted in a given direction will be equal in amplitude but different in phase. In the figure, the waves shown originating from neighbouring rulings will differ in phase due to a difference in path length AB + BC. In fact:

$$AB + BC = w \sin i + w \sin \theta$$

$$= w (\sin i + \sin \theta)$$

A difference in path length of λ would cause a difference in phase of 2π so that the phase difference between the waves refracted from similar parts of neighbouring rulings shown is

$$\frac{2\pi(AB + BC)}{\lambda}$$

In general then, the mean value of light in any direction θ is zero due to random phase differences, except where the value of θ is such that the phase differences are multiples of 2π, i.e. effectively in phase. Thus waves will be produced at values of θ where

$$\frac{2\pi(AB + BC)}{\lambda} = n.2\pi$$

$$\text{i.e. } w(\sin i + \sin \theta) = n\lambda$$

This is the basic equation of all plane gratings and is known as the 'grating equation'. 'n' is known as the order of the refracted wave. The zero order wave is that which has $\theta = - i$, i.e. in line with the incident wave. The light energy is distributed into the various orders dependent obviously upon the detailed structure of the ruling and upon i, θ and λ. This is known as the Blaze distribution.

For example, the grating of figure 9a, can be made to concentrate most of its energy in the zero and 1st orders in about equal quantities. If for this grating λ is halved, most of the energy goes into the first order.

'Blazing' of the light then can be achieved in two different ways. One, by modifying the phase of the light, which is known as a phase grating, i.e. the prismatic form shown in figure 9a. Two, by modifying the amplitude of the light, i.e. absorption gratings of the line and space type we have considered previously. (Figure 9b). This type concentrates most energy into the zero and two first orders. Note the plus and minus orders.

To produce Moiré fringes then, two parallel similar gratings are required. Each wave is split into its n orders at the first grating and each order wave is split into m orders at the second grating. If the gratings are identical n = m and we have a resultant of n^2 composite order waves, travelling in n directions, i.e. each final wave will have n components. Figure 10 shows the components of the resultant 1st order wave. Note that the algebraic sum of the orders of diffraction add up to 1.

Now consider a point P on the first grating surface and let the phase of the zero, 1st, 2nd etc. order waves with respect to the incident wave be $\phi_{\cdot}$, ϕ_1, ϕ_2, etc., at this point. At a point Q distant x along the grating surface from P then the phases are ϕ_o, $(\phi_1 + \frac{x}{w}.2\pi)$, $+ (\phi_2 + 2.\frac{x}{w}.2\pi)$, etc. This is obvious from the derivation of the grating equation, i.e. the second order wave is produced by 2 complete cycles of phase shift at spacing x = w.

Thus the phase of the m order wave relative to the incident wave is $\phi m + \frac{2\pi mx}{w}$.

If the relative displacement between the two gratings is ℓ then the displacement from P is $\ell + x$ and the phase of the n order wave leaving the second grating with respect to its parent, the m order wave from the first grating, is $\phi n + \frac{2\pi n(x + \ell)}{w}$ assuming the two gratings to lie in the same plane. Therefore, the phase of m + n order component produced with respect to the incident wave is

$$\phi n + \phi m + \frac{2\pi(n + m)}{w}x + \frac{2\pi n\ell}{w}$$

Another component of the m + n order wave will be the n - p order at the second grating produced from the m + p order at the first, the relative phase of which will be

$$\phi(n - p) + \phi(m + p) + \frac{2\pi(n + m)}{w}x + \frac{2\pi(n - p)}{w}\ell$$

The relative phase of these two components of the n + m order wave is

$$\phi n + \phi m - \phi(n - p) - \phi(m + p) + \frac{2\pi p\ell}{w}$$

Thus if p = 1, relative displacement of the gratings by a distance $\ell = w$ will cause a complete cycle of phase shift. The resultant signal therefore goes through one complete cycle for a relative movement of 1 ruled pitch.

For p bigger than 1, harmonics will be produced in the output. Blazing of the grating to produce a minimum of orders will therefore be desirable. Fringe blurring can also be introduced to reduce harmonics but this also reduces the depth of modulation.

This system is shown in figure 11, where a slit is used to isolate the m + n order wave. This is called a 'Spectroscopic' system. The useful range of gratings for such a system is about 1000 to 7000 lines per inch; the lower limit set by low angular spread of the various order waves making them inseperable, and the upper limit by fringe blurring due to the finite size of the source.

Normal incidence system

The general case of the spectroscopic system shown above has a somewhat limited application to machine tools for the following reasons:

i) The system is complex to set up since parallelity of the gratings, angle of incidence and the relative angle and size of the slit are all critical.

ii) the limit set by the number of lines per inch available is unacceptable.

The 'normal incidence' system (i.e. angle of incidence = 0) is suitable for all pitches up to 1000 lines per inch and with certain refinements up to 2500 lines per inch, the desirability of which range we have already discussed.

Here the photocell is placed up against the index grating, therefore receiving all transmitted orders (figure 12). In fact the zero order wave represents the fundamental output with harmonics introduced by the 1, -1; 2, -2, etc. waves. To limit the magnitude of the waves of orders other than zero the blaze distribution of the grating is important and for this reason, line and space gratings are employed as being most suitable as well as most easily manufactured.

At this stage, while developing an expression for the zero order, it will be convenient to also consider the effect of gap width between the grating and index.

Optimum gap for normal incidence system

Refer to figure 13. Consider two points on a line of normal incidence distant g apart, then the m order wave front has travelled $g \cos \theta m$ resulting in a phase change

$$\frac{2\pi g}{\lambda} \cdot \cos \theta m$$

Also from the grating equation

$$w(\sin i + \sin \theta) = n\lambda$$

with $i = 0 \quad \sin\theta = \frac{n}{w}\lambda$

$$\therefore \quad \cos\theta m = \sqrt{1 - \frac{m^2\lambda^2}{w^2}} = k_m$$

The total phase of the m order wave at P relative to the incident wave is now

$$\phi m + \frac{2\pi xm}{w} + \frac{2\pi g}{\lambda} \cdot km$$

Let the second (index) grating be displaced ℓ with respect to the first grating. Then the n order wave derived from the m order wave from the first grating has a total phase relative to the incident wave of

$$\phi n + \phi m + \frac{2\pi xm}{w} \cdot + \frac{2\pi(x + \ell)}{w} n + \frac{2\pi g}{\lambda} \cdot k_m$$

$$= \phi n + \phi m + \frac{2\pi x}{w}(m+n) + \frac{2\pi \ell n}{w} + \frac{2\pi g}{\lambda} \cdot km$$

Consider the zero order resultant wave. Let the relevant waves be the 0, 0; 1, -1; and -1, 1 waves, all others being negligible. Let the amplitudes of the various waves be C_o, C_1, C_{-1}, etc. and of the resultant waves $A_{o,o}$; $A_{1,-1}$ etc.

Also since $km = \left(1 - \frac{m^2\lambda^2}{w^2}\right)^{\frac{1}{2}}$

$$ko = 1; \; k_1 = k_{-1} = \left(1 - \frac{\lambda^2}{w^2}\right)^{\frac{1}{2}}$$

then

$$Aoo = Co^2$$

$$\phi oo = 2\phi_o + \frac{2\pi g}{\lambda}$$

$$A_{1,-1} = A_{-1,1} = C_1 \cdot C_{-1}$$

$$\phi_{1,-1} = \phi_1 + \phi_{-1} - \frac{2\pi\ell}{w} + \frac{2\pi g}{\lambda} \cdot k_1$$

$$\phi_{-1,1} = \phi_{-1} + \phi_1 + \frac{2\pi\ell}{w} + \frac{2\pi g}{\lambda} \cdot k_{-1}$$

For a line and a space grating

$$C_o = \tfrac{1}{2} \; ; \quad \phi_o = 0.$$

$$C_1 = C_{-1} = \frac{1}{\pi}; \quad \phi_1 = \phi_{-1} = 0.$$

Hence

$$Aoo = \frac{1}{4} \qquad \phi oo = \frac{2\pi g}{\lambda}$$

$$A_{1,-1} = \frac{1}{\pi^2} \qquad \phi_{1,-1} = \frac{2\pi g}{\lambda}k_1 - \frac{2\pi \ell}{w}$$

$$A_{-1,1} = \frac{1}{\pi^2} \qquad \phi_{-1,1} = \frac{2\pi g}{\lambda}k_1 + \frac{2\pi \ell}{w}$$

put $\frac{2\pi g}{\lambda} = \alpha$

Hence the resultant zero order wave;

$$\overline{A_o} = \overline{A_{oo}} + \overline{A_{1,-1}} + \overline{A_{-1,1}} + \text{other negligible terms}$$

$$= \frac{1}{4}e^{j\alpha} + \frac{1}{\pi^2}e^{j(k_1\alpha - \frac{2\pi\ell}{w})} + \frac{1}{\pi^2}e^{j(k_1\alpha + \frac{2\pi\ell}{w})}$$

$$= \frac{1}{4}e^{j\alpha} + \frac{1}{\pi^2}\left(e^{-\frac{j2\pi\ell}{w}} + e^{\frac{j2\pi\ell}{w}}\right) e^{jk_1\alpha}$$

$$= \frac{1}{4}e^{j\alpha} + \frac{1}{\pi^2} \cdot 2\cos\frac{2\pi\ell}{w} \cdot e^{jk_1\alpha}$$

Refer to figure 14.

The angle α is a constant for a given gap. However, as the relative position of the two gratings, ℓ, varies $\bar{A}_o$ describes a locus as shown in figure 14. It can be seen then, that for maximum modulation,

$$\alpha = k_1\alpha + n\pi$$

Thus the optimum gap is given by

$$\alpha = \frac{n\pi}{1 - k_1}$$

since

$$\alpha = \frac{2\pi g}{\lambda}$$

and

$$k_1 = \left(1 - \frac{\lambda^2}{w^2}\right)^{\frac{1}{2}}$$

$$= 1 - \frac{1}{2}\frac{\lambda^2}{w^2} \qquad \text{if } \lambda \ll w$$

then

$$\frac{2\pi g}{\lambda} = \frac{n\pi}{\frac{1}{2} \cdot \frac{\lambda^2}{w^2}}$$

$\therefore \quad g = \frac{nw^2}{\lambda}$ for maximum modulation

thus $\quad A_o' max. = \frac{1}{4} + \frac{2}{\pi^2}$

and $\quad A_o\ min. = \frac{1}{4} - \frac{2}{\pi^2}$

Since the photocell output is proportional to light intensity, which is proportional to amplitude squared.

$$\text{P.C. output max.} \ \alpha \ \left(\frac{1}{4} + \frac{2}{\pi^2}\right)^2 = .205$$

$$\text{P.C. output min.} \ \alpha \ \left(\frac{1}{4} - \frac{2}{\pi^2}\right) = .0022$$

$$\therefore \ \text{Depth of modulation} = \frac{I\ max - I\ min}{I\ max + I\ min}$$

$$= 99\%$$

Practical tolerance on gap:

Further calculations from the preceeding work will show that the theoretical modulation is always bigger than 90% provided that:

$$g = \frac{(n \pm 0.1)\ w^2}{\lambda}$$

Thus for light of wave length $= 5{,}000\ \mathring{A}$

$$= 2 \times 10^{-5} \text{ ins.}$$

and a grating of 1000 lines per inch

$$w = .001 \text{ ins.}$$

then

$$g = \frac{n\ .\ .001^2}{2 \times 10^{-5}} = 5 \times 10^{-2}\ .\ n \text{ ins.}$$

and the tolerance of $0.1 \times 5 \times 10^{-2} = 5 \times 10^{-3}$ ins.

Thus gaps of 0 to 5×10^{-3} or 4.5×10^{-2} to 5.5×10^{-2} inches can be used. In practice the narrower gap is normally used.

Theoretically larger gaps corresponding to higher values of n than 0 and 1 could be used but the results get progressively poorer, due to the neglected terms, i.e. $A_{2,-2}$ etc. and due to the fact that the more usual photographic line and bar grating, due to emulsion shrinkage, has some phase grating properties which alters the blaze distribution.

Fringe blurring

If the light waves are so arranged as not to be collimated perfectly but to have some degree of spread (figure 15) fringe blurring will be introduced which can eliminate certain harmonics in the photocell output. The disadvantage will be a reduction in the depth of modulation.

The degree of blurring is best expressed as an equivalent range of phase change, α, such that 1 fringe of blurring gives $\alpha = \pm \pi$, for the fundamental distribution over a ruling pitch.

Let the light signal be $1 + a \sin \theta$

where θ depends upon ℓ and a the depth of modulation. Then with blurring θ lines in the range $\theta + \alpha$ to $\theta - \alpha$

$$\therefore \text{ mean signal } = \frac{1}{2\alpha} \int_{-\alpha}^{\alpha} 1 + a \sin(\theta + \beta)\, d\beta$$

$$= 1 + \frac{a}{2\alpha} \Big[-\cos(\theta + \beta) \Big]_{-\alpha}^{\alpha}$$

$$= 1 + \frac{a}{2\alpha} [\cos(\theta - \alpha) - \cos(\theta + \alpha)]$$

$$= 1 + \frac{a}{2\alpha} \cdot 2 \sin\theta \sin\alpha$$

$$= 1 + \frac{\sin\alpha}{\alpha} \cdot a \sin\theta$$

Thus the depth of modulation is reduced due to blurring by $\frac{\sin\alpha}{\alpha}$. For the nth harmonic, i.e. n complete cycles distributed over one ruled pitch, reduction is by:

$$\frac{\sin n\alpha}{n\alpha}$$

Thus if we have $\frac{1}{2}$ fringe of blurring

$$\alpha = \pm \frac{\pi}{2}$$

$$\therefore \quad \frac{\sin \alpha}{\alpha} = \frac{1}{\frac{\pi}{2}} = .64$$

reducing the depth of modulation to $.64 \times 99 = 63\%$

But for the 2nd harmonic $\dfrac{\sin 2\frac{\pi}{2}}{2\frac{\pi}{2}} = 0$

In fact all even harmonics are eliminated.

References

1. McIlraith, Dr. A.H. — An introduction to the theory of crossed diffraction gratings and their applications to linear measurement. Machine Shop Mag., April 1962.

2. Guild, J. — Diffraction gratings as measuring scales. Oxford University Press.

3. - — Measurement and machine control by means of gratings. D.S.I.R. pamphlet.

4. Blake, D.V., and Sayce, L.A. — Machine tool monitor using diffraction gratings. Trans. Soc. Instr. Tech. Vol. 10, No. 4, Dec. 1958.

5. Barber, B.L.A., and Atkinson, M.P. — Method of measuring displacement using optical gratings. Journal of Scientific Instruments, Vol. 36, Dec. 1959.

6. Davies, B.J., Robbins, R.C., Wallis, C., and Wilde, R.W. — A high resolution measuring system using coarse optical gratings. Proc. I.E.E. Vol. 107 B, No. 36, Nov. 1960.

7. Williamson, D.T.N. — Computer control of machine tools. Control. July and August 1958.

8. Ogden, H. — Numerical control of machine tools. Industrial Electronics. November 1962.

9. Walker, D.F. — Ferranti Mk IV continuous path phase control system. Internat. Journal of M/c Tool Des. and Res. Vol. 3, pp. 61-73.

10. Oster, G., and Nishijima, V. — Moiré patterns. Scientific American, May 1963.

11. Stanley, V.W., and Hall, R.G.N. — The production and performance of metrological diffraction gratings. The Prod. Engineer. January 1961.

12. Burch, J.M. — Photographic production of gratings for measurement. Research Vol. 13, January 1960.

13. Russell, A. — Electronic digital ratio equipment. D.S.I.R. - N.E.L. Report No. 94.

14. Leslie, W.H.P. — Widening the applications of diffraction gratings. Int. Journal of M/c Tool Des. and Res. Vol. 2, No. 4.

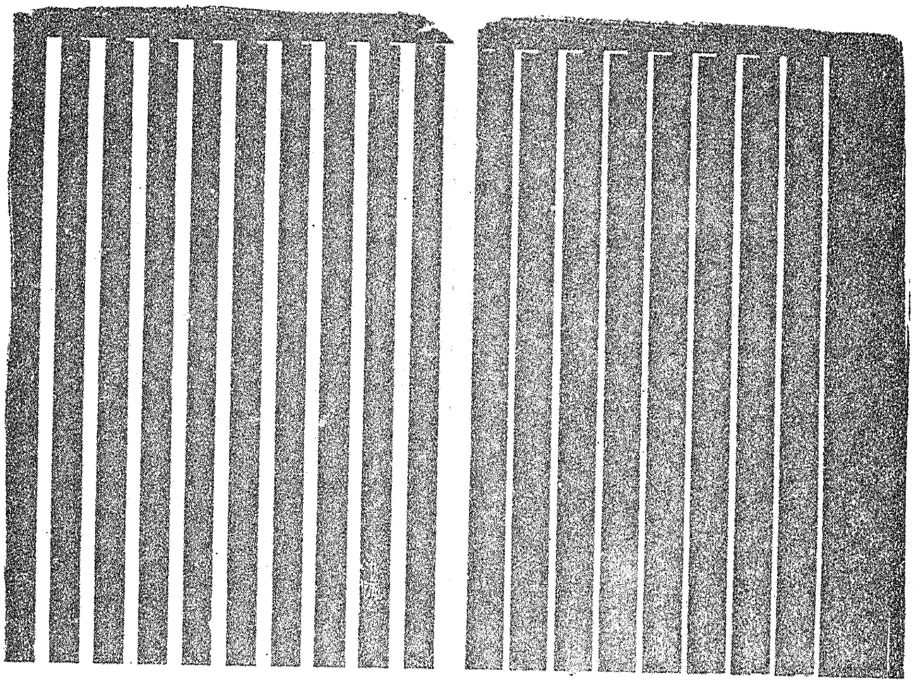

FIG. 1

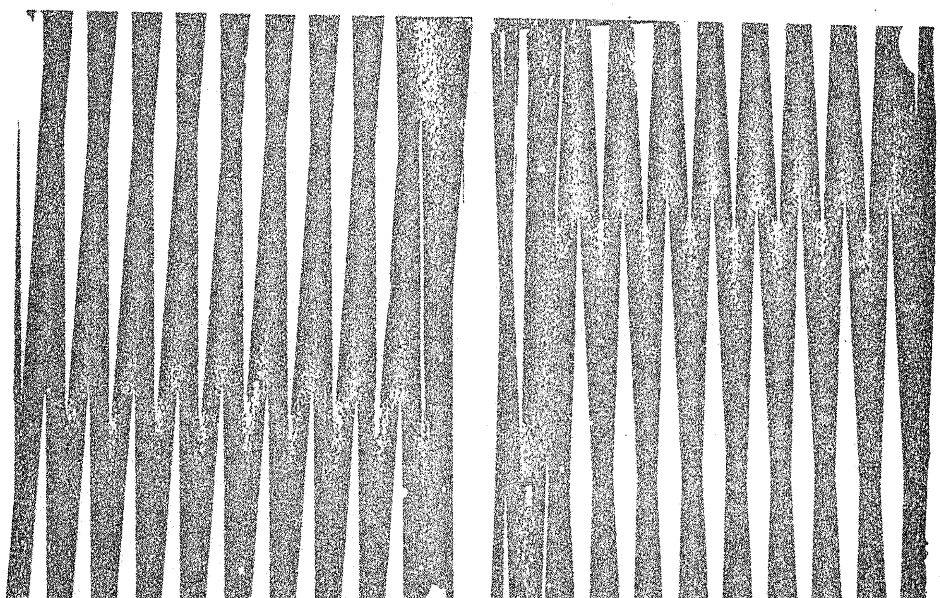

FIG. 2

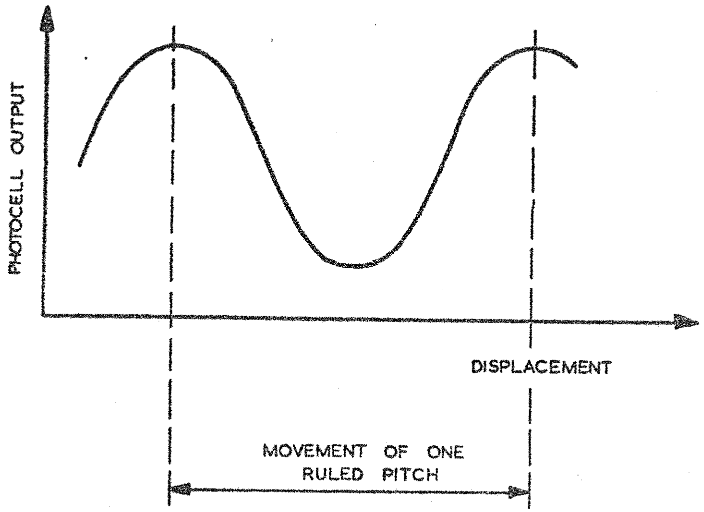

FIG. 3. PHOTOCELL OUTPUT

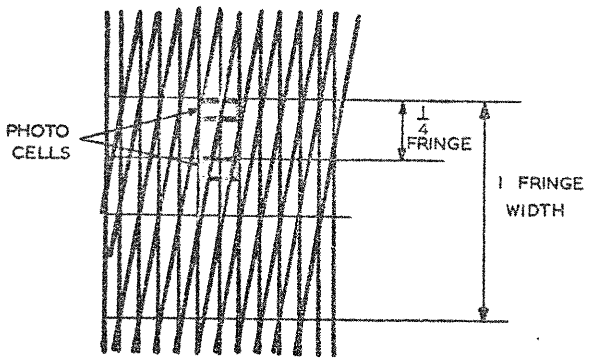

FIG. 4. PLACING OF PHOTOCELLS TO OBTAIN DIRECTION SENSING.

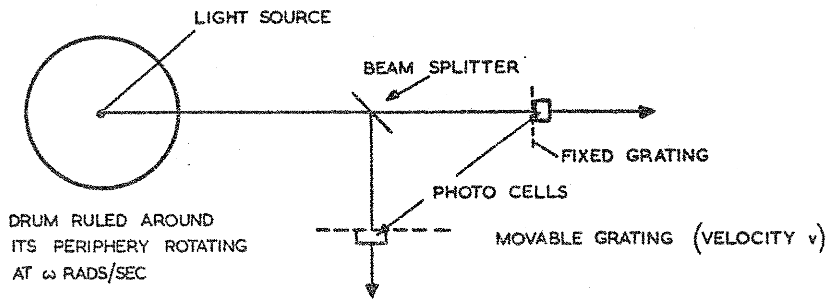

FIG. 5. PHASE MODULATED SYSTEM.

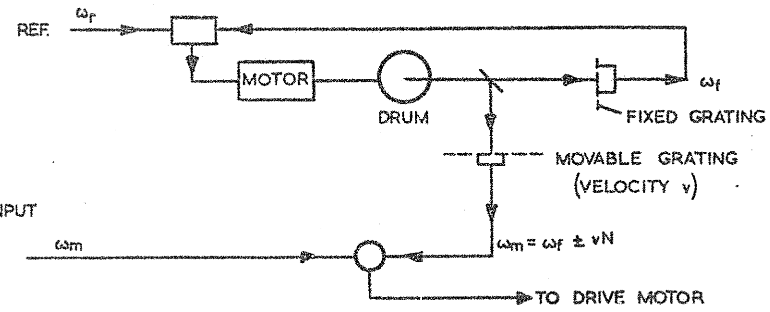

FIG. 6. THE FERRANTI SYSTEM.

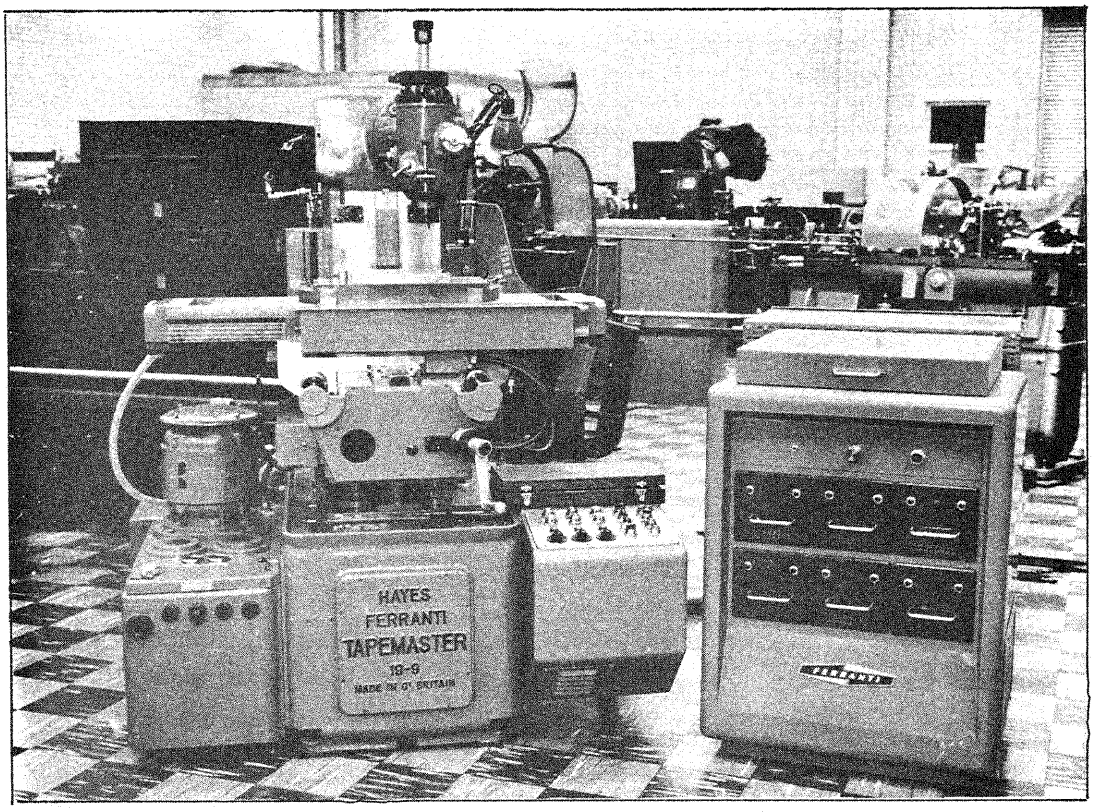

FIG 7 (b) HAYES-FERRANTI NUMERICALLY CONTROLLED MILLING MACHINE

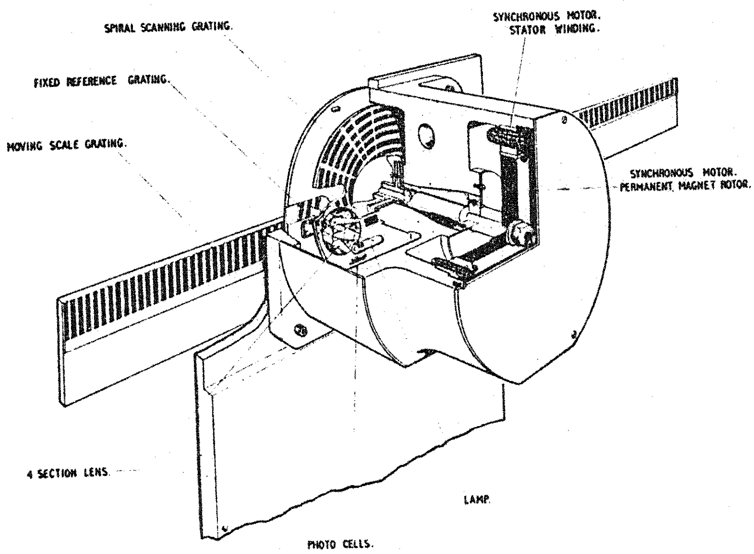

FIG 7 (c) FERRANTI SCANNING READING HEAD

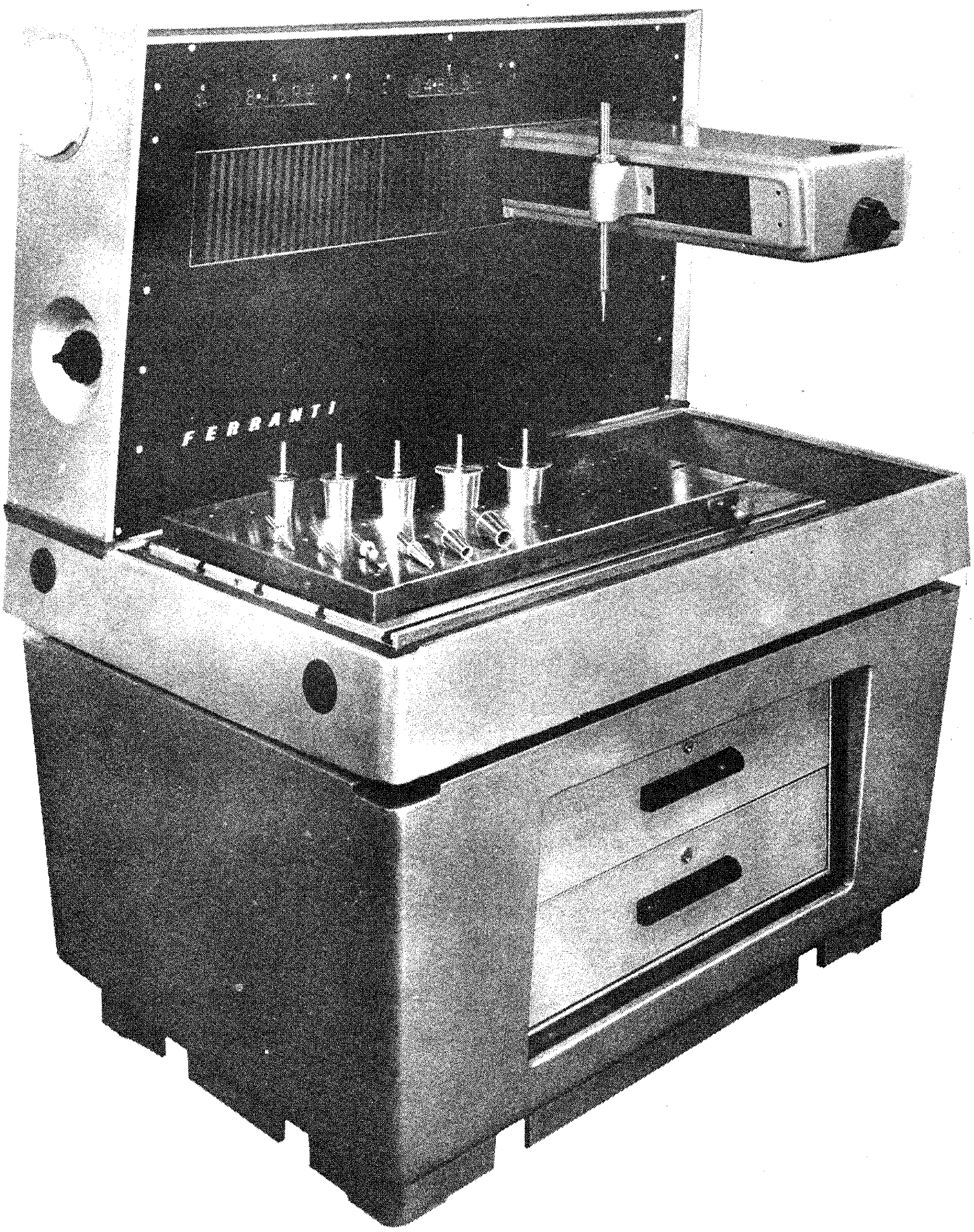

FIG 7 (a) FERRANTI CO-ORDINATE MEASURING MACHINE

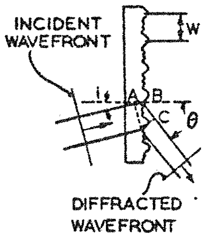

FIG. 8. EFFECT OF REPEATED ROUGHNESS PATTERNS.

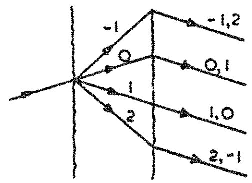

FIG. 10. PATHS OF COMPONENT WAVES THROUGH A PAIR OF GRATINGS.

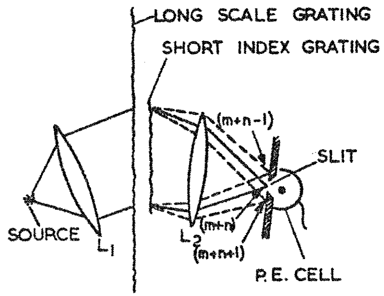

FIG. 11. THE SPECTROSCOPIC SYSTEM.

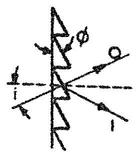

FIG. 9a. PHASE GRATING WITH PRISMATIC GROOVES.

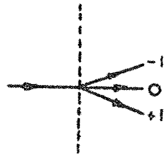

FIG. 9b. LINE AND SPACE AMPLITUDE GRATING.

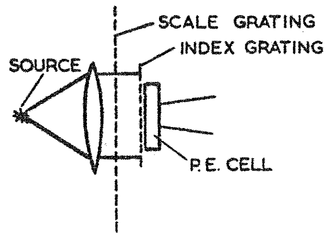

FIG. 12. COARSE AMPLITUDE GRATINGS FOR MEASUREMENT OF DISPLACEMENT.

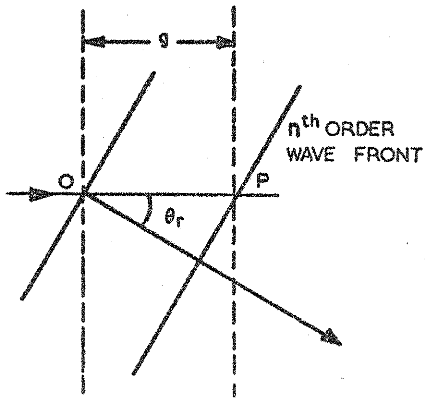

FIG. 13.

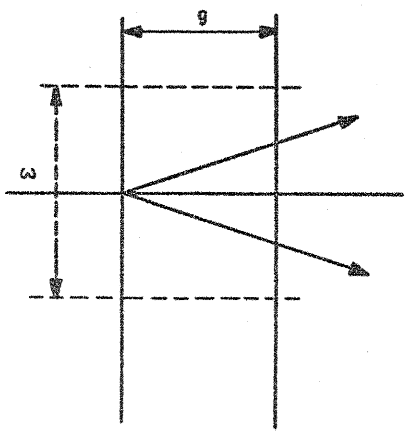

FIG. 15. FRINGE BLURRING.
(AS SHOWN $\alpha = \pm\frac{\bar{\Lambda}}{2}$)

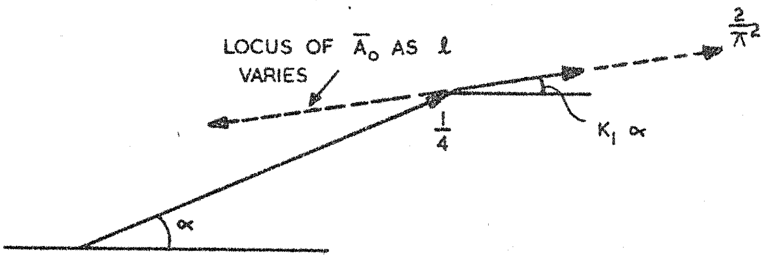

FIG. 14. VECTOR DIAGRAM OF RESULTANT ZERO ORDER WAVE.