A Thesis on

"An Achromatic Wavefront Folding Interferometer"

submitted for the PhD degree

George Richard Wloch B. Sc (Eng) ACGI

Electrical Engineering Department
Imperial College, London.

December 1980

Acknowledgement

This work would not have been accomplished had it not been for the unending gentle help and guidance of my Tutor, Dr. John Cozens. For his understanding and patience I am truly grateful.

I would also like to express my gratitude to Professor J. Brown and his staff for the facilities provided and for countless discussions and suggestions.

Finally, I wish to thank Yvonne and Anita without whose help this Thesis would not have been produced.

Dedicated to
my Mother and Father,
Yvonne and Richard,

ABSTRACT

This thesis describes the development of a passive optical system for the formation, in real time, of the two dimensional cosine Fourier transform of a polychromatic, incoherent, two dimensional intensity distribution. The generated transform is itself an intensity distribution and as such can be either processed by an optical filter, or be directly recorded, by a conventional technique, for subsequent electronic processing.

The system is based on the wavefront folding and shearing interferometer which conventionally produces an interference pattern whose spatially varying visibility is directly related to the input object's Fourier transform. The interferometer requires that the two dimensional input object be spatially incoherent and quasi-monochromatic. The particular development described here is the extension of the interferometer to accommodate source emission over the full visible spectrum. This is achieved by introducing spatially and chromatically varying degrees of shear to create an achromatic interference pattern. The necessary shearing function has been realised with a pair of mutually inclined diffraction gratings.

A theoretical analysis of the instrument is presented for two dimensional input objects and their transforms, comparing the spectrally compensated interferometer with other techniques for producing Fourier transforms, and with other achromatic fringe systems. The design and realisation of a one dimensional implementation of the instrument is discussed.

An experimental investigation of the one dimensional interferometer is detailed, showing the manner in which the Fourier transform was obtained from a 1 cm wide object illuminated with white light. The sinc function transform of the slit modulates the visibility of over 400 essentially white carrier fringes. The 5th side lobe of the transform was detectable. Finally, the effects of the bias intensity level on a subsequent reconstruction of the original distribution are discussed.

TABLE OF CONTENTS

CHAPTER I

Two Dimensional Information

Many two variable functions describing disturbances of interest may be presented as two dimensional light distributions. Typical examples of such light distributions would range from astronomical sources to microscopic images and from C.R.T., displays to laser signals. Not all two variable functions when displayed in two dimensional space are in such a form that the information about the event of interest is readily discernable. Hence some form of two dimensional signal processing is necessary in order to extract the required information. This processing may involve recognising characters, either alphanumerics or some predetermined patterns,enhancing the edges of the given function and removing image degradations.

The various two dimensional signal processors that have been devised to date can be divided into two groups; those that operate in the image plane where the function is displayed, and those which produce one of several possible transforms of the function and then operate on this transform. The main advantage of the former group is that the information is processed directly without the need to pass through several intermediate stages, where some of the information could be lost or distorted. Against this, the fundamental disadvantage is that any processing of a given feature first requires that the entire image plane be scanned to locate the given feature. In contrast, using a judicious transform, the information about the given feature can be located in a specified area regardless of the exact position of the feature in the image plane. The apriori knowledge of the character's transform greatly facilitates processing the original function.

1.1. General Signal Processing

The two most successful techniques developed to date to implement two dimensional signal processing are based on digital electronic systems and on optical analogue systems. The actual form of processing performed by the digital systems will depend to a large extent on the computing facilities available. The general purpose computers are better suited for image plane filtering due to their ability to utilise a large range of flexible computer programs. This enables the computers to implement a wide range of mathematical operations designed for digital filtering (Crouchiere and Oppenheim 1975).

The most widely used concept in this type of processing is correlation for feature recognition (Wood and Treitel 1975, Stockham et al 1975) and deconvolution for removing image degradations (Sondhi 1972). However, these general purpose computers are usually much slower in generating various transforms of the input data than special, purpose built, computer systems (Allen 1975, Freeny 1975) which have been designed to produce and then filter only one specific transform.

In general, digital signal processing, based on computers has the merits of flexibility in design and in languages, and hence in programs, giving the processor a wide and powerful operating range. Secondly, the operations involved are virtually noise free and hence are infinitely repeatable. The two major disadvantages stem from the fact that the operations of all electronic systems are limited to one dimension. Primarily this implies that the input image must be sequentially sampled at an appropriate rate and a binary number associated with the magnitude of the light signal at the particular sample. This digitisation of the input must inevitably introduce a certain amount of error. Secondly, due to the limitation of operating on a function of one variable at a given instant, the two dimensional input signal has to be considered as a sequence of one dimensional signals. Clearly, this is time consuming and presents a finite time delay between the arrival of the signal and the presentation of the processed information. Hence, processing signals in real time, that is with a time delay which is of no consequence in a practical system, may be very difficult.

Two dimensional signal processing can also be performed by optical analogue systems. A two dimensional wave propagating through an optical system will be modified by the two dimensional transfer function of the system. Hence, if the transfer function can be made to represent some form of filter, then the two dimensional information on the wave will be processed instantaneously. This ability to handle two dimensional information in real time gives the optical signal processors a distinct potential advantage over the digital systems.

Optical analogue systems can also operate directly on the image or produce a transform of the image. In the first instance, the transfer function could be some form of matched filter recognising a given character from a known alphabet. Or, secondly, and more usually, the transfer function of the optical system gives rise to an output which can be related to the Fourier transform of the original input function. Different optical processors generate outputs that have diversified

relationships with the exact Fourier transform of the input. Thus before proceeding with a discussion of the various processors it is useful to clarify the meaning of a Fourier transform of a spatial function.

1.2. Fourier Transforms

A two dimensional Fourier transform can be associated with a two dimensional energy distribution, $g(x,y)$ provided that this input function has only a finite number of discontinuities and that it is integrable over its spatial extent. If these conditions are satisfied then the Fourier transform is defined by

$$G(p,q) = \iint_{-\infty}^{\infty} g(x,y) \exp - j2\Pi(xp + yq)\, dxdy \qquad 1.1$$

where p and q represent the Fourier frequencies. Since $g(x,y)$ is a spatial distribution, $G(p,q)$ must also represent an energy distribution in two dimensional space. Thus the Fourier spatial frequencies p and q are directly proportional to the distances along an orthogonal set of axes in the co-ordinate frame defining the transform plane. The above integral can also be considered as representing the transfer of energy, or information from a particular spatial periodicity in the image plane, to a point whose precise position in the transform plane, is determined by the wavelength of the periodicity. Thus the input object can be described in terms of a set of spatial periodicities whose complex amplitudes are given by $G(p,q)$.

Perhaps the most significant advantage of the Fourier plane over the image plane in signal processing is that the magnitude of the spatial frequency spectrum is independent of the input function position in the image plane. Thus if equation 1.1. represents the spatial spectrum of the function $g(x,y)$ then for a shift in the position of the function of $(+a,+b)$ to $g(x+a,y+b)$ the transform becomes

$$G'(p,q) = \iint_{-\infty}^{\infty} g(x+a\ ,\ y+b) \exp - j2\Pi(xp+yq)\, dxdy$$

$$= \exp j2\Pi(pa + qb) \iint_{-\infty}^{\infty} g(x+a,\ y+b)\ \exp - j2\Pi\{p(x+a) + q\ (y+b)\}\, d(x+a)\, d(y+b)$$

$$= \exp j2\Pi(pa + qb)\ G(p,q) \qquad 1.2$$

Which is clearly the transform of the original undisplaced function multiplied by a phase term. Hence, the location of the spatial spectrum remains invariant for input function displacements. Knowing that the transform of a given feature will always appear in the same area, regardless of the feature's position in the image plane, greatly facilitates the processing of the given feature. One of the simplest forms of processing in this Fourier plane is to place a binary mask at the centre of the transform, thus removing the low spatial frequencies in order to enhance the edges or differentiate the input function. Clearly, more complicated masks, even holograms containing phase information, can be placed in the Fourier plane to realise other forms of processing.

This form of filtering can be considered as a product, in the transform plane, of the filter function $H(p,q)$ and the transform of the function, $G(p,q)$. This product can be related to an operation in the image plane by taking the Fourier transforms, thus

$$O(x,y) = \iint_{-\infty}^{\infty} H(p,q)\ G(p,q)\ \exp j2\Pi(px + qy)\ dpdq$$

$$= \iint_{-\infty}^{\infty} H(p,q) \left\{ \iint_{-\infty}^{\infty} g(r,s) \exp -j2\Pi(pr + qs)\, drds \right\} \exp j2\Pi(px + qy)\, dpdq$$

$$= \iint_{-\infty}^{\infty} g(r,s) \left\{ \iint_{-\infty}^{\infty} H(p,q) \exp j2\Pi(p\{x-r\} + q\{y-s\})\ dpdq \right\} drds$$

$$= \iint_{-\infty}^{\infty} g(r,s)\ h(x-r,\ y-s)\ drds \qquad 1.3$$

which can be recognised as a convolution integral. Hence, complicated convolution operations in the image plane reduce to simple multiplications in the Fourier plane.

The Fourier transform as defined in equation 1.1., is a linear and unique mapping from the image plane x, y to the spatial frequency plane p, q and hence the inverse transform can be readily written as

$$g(x,y) = \iint_{-\infty}^{\infty} G(p,q) \exp j2\Pi(px+qy)\, dpdq \qquad 1.4$$

Thus the information in the processed transform can be readily transferred back to the image plane by performing the inverse Fourier transform as defined above. However, a result of far greater practical significance is the consequence of performing two successive positive transformations on a given signal. Thus, the result of such an operation $g'(x',y')$ is given by

$$g'(x',y') = \iint_{-\infty}^{\infty} G(p,q) \exp - j2\Pi(px' + qy')\, dpdq$$

$$= \iint_{-\infty}^{\infty} \left\{ \iint_{-\infty}^{\infty} g(x,y) \exp - j2\Pi(px+qy) dxdy \right\} \exp -j2\Pi(px'+qy')\, dpdq$$

$$= \iiint\!\!\iint_{-\infty}^{\infty} g(x,y) \exp - j2\Pi\left\{p(x+x') + q(y+y')\right\}\, dxdy\, dpdq$$

$$= \iint_{-\infty}^{\infty} g(x,y) \left\{ \iint_{-\infty}^{\infty} \exp -j2\Pi\left\{p(x+x') + q(y+y')\right\}\, dpdq \right\} dxdy$$

The integral inside the brackets may be expressed as follows:-

$$= \lim_{a \to \infty} \int_{-a}^{+a} \int \exp -j2\Pi\left\{p(x+x') + q(y+y')\right\}\, dpdq$$

$$= \lim_{a\to\infty} \left\{ \frac{\sin 2\Pi a\ (x+x')}{2\Pi(x+x')} \cdot \frac{\sin 2\Pi a\ (y+y')}{2\Pi(y+y')} \right\}$$

$$= \delta(x+x')\,\delta(y+y')$$

$$g'(x',y') = \int_{-\infty}^{\infty}\int g(x,y)\ \delta(x+x')\ \delta(y+y')\ dxdy$$

$$g'(x',y') = g(-x', -y') \qquad 1.5$$

Thus a real system producing a transform of an input function can be used to reconstruct the function. The only difference between this reconstruction and that obtained with an inverse transform is that this reconstruction is inverted. However, for all practical situations this is of little or no consequence.

1.3. Propagation of Light

The mode of operation of the optical system as suggested in section 1.1., depends on the propagation of light waves, hence it is necessary to consider the nature of optical sources and the fields that are generated by them. Although information when displayed as a light distribution is always perceived as an intensity function, a single point radiating at a discrete wavelength generates an electro-magnetic field over which the amplitude and phase are well defined. Therefore, when considering the effect of an optical system on such an elementary point radiator, that point must be described in terms of a complex amplitude, or even more precisely in terms of its electric and magnetic vectors.

The rigorous Maxwellian electro-magnetic approach does not readily lend itself to complete solutions in all but the simplest situation (Somerfeld 1954). However, provided certain restrictions are admitted, a simpler scalar technique, based on the amplitude of the electric vector can be successfully used. These restrictions are significant when the behaviour of electric and magnetic components of the light waves is important, notably in the vicinity of objects with dimensions comparable

to the wavelength of light. Therefore, small apertures, field distributions within several wavelengths of the source and polarisation effects must be treated with caution. However, in practical circumstances these restrictions are rarely an imposition.

The propagation of a light wave can be described by the general wave equation

$$\nabla^2\Psi = \frac{1}{c^2}\frac{\partial^2\Psi}{\partial t^2}$$

Where ∇^2 is the Laplacian differential operator and c is the speed of light in vacuum. For an isotropic point source situated at x', y', o in a cartesian co-ordinate system, this wave equation will be satisfied by a propagating function of the form

$$\Psi = \frac{a}{2}\cos\left\{k\left\{(x-x')^2 + (y-y')^2 + z^2\right\}^{\frac{1}{2}} - \omega t - \Phi\right\}$$

Where a and Φ are positive real constants and k is the particular radiation wavenumber defined by

$$k = \frac{2\Pi}{\lambda}$$

Although this propagation function is not defined at the source x, y, o it does not lose its validity in the general field more than several wavelengths from the source where the scalar theory is applicable. For ease of mathematical manipulation, the above function can be represented by an exponential

$$\Psi = \tfrac{1}{2} V \exp j\left\{k\left\{(x-x')^2 + (y-y')^2 + z\right\}^{\frac{1}{2}} - \omega t\right\} \qquad 1.6$$

Where V is a complex constant.

Due to the ease with which the propagation of a plane wave through a system can be analysed, it is convenient to describe this spherical wave emanating from the elementary point source in terms of a set of plane waves propagating in directions normal to the spherical wavefront. Thus if each of these plane waves is assigned an appropriate amplitude and initial phase, the complex amplitude distribution across the source will be exactly represented. A typical uniform plane wave, of complex amplitude U propagating in a direction such that its normal makes angles of ϑ, χ, ψ with the x, y, z axes respectively, can be represented by

$$U \exp j \left\{ 2\Pi(xu + yv + zw) - \omega t \right\}$$

where u, v, w are spatial frequencies defined by

$$u = \frac{\cos \vartheta}{\lambda} \qquad v = \frac{\cos \chi}{\lambda} \qquad w = \frac{\cos \psi}{\lambda}$$

This equation can be simplified by omitting the propagation term ωt which remains invariant whilst the wave propagates through an isotropic medium, and by omitting the spatial frequency w as it is determined by the values of the other two spatial frequencies in the relationship

$$u^2 + v^2 + w^2 = \frac{1}{\lambda^2} \qquad 1.7$$

Hence, the typical plane wave can be written as

$$U_s(u, v) \exp j2\Pi(ux + vy)$$

with u, v determining the direction of propagation of the plane wave and $U_s(u, v)$ its amplitude. As suggested earlier, the source complex amplitude can be expressed by the sum of a number of plane waves each of which propagates in a different direction. Therefore, the complex

amplitude distribution $V(x,y)$ in the source plane is given by

$$V(x,y) = \iint_{-\infty}^{\infty} U_s(u,v) \exp j\, 2\Pi (ux + vy)\, du dv \qquad 1.8$$

This description is known as the angular spectrum of plane waves. The limits of integration are taken from $-\infty$ to $+\infty$ in order to have the integral in the standard Fourier form, however, they must be treated with caution. Clearly for values of u and v greater than $\left|1/\lambda\right|$ the direction cosines are larger than unity, in which case propagation can only be defined in terms of imaginary angles. The plane waves produced under such conditions are evanescent waves which oscillate in planes parallel to the source plane, but are attenuated exponentially in the normal direction. Thus their contribution to the field beyond the immediate vicinity of the source plane is negligible and the practical integration in equation 1.8., is performed between $\pm 1/\lambda$. With this equation it is possible to associate an inverse, that is, the positive Fourier integral

$$U_s(u,v) = \iint_{-\infty}^{\infty} V(x,y) \exp - j2\Pi(ux + vy)\, dx dy \qquad 1.9$$

which clearly defines the amplitudes of all the individual plane waves representing the radiation from the source.

A plane wave propagating from the source plane to the plane of interest ξ, η through an isotropic medium will be altered in phase only by an amount ϕ given by

$$\phi = \exp j\, 2\Pi z\, w$$

Which on expressing w from equation 1.7 gives

$$\phi = \exp j\, 2\Pi z\, (1/\lambda^2 - u^2 - v^2)^{1/2}$$

This phase term applies equally to all the plane waves in the spectrum propagating from the source. Thus it is possible to define an

angular spectrum in the plane ξ, η in terms of the source spectrum as follows

$$U'(u,v) = U_s(u,v) \exp j 2\Pi \left\{ 1/\lambda^2 - u^2 - v^2 \right\}^{1/2} z$$

For directions of propagation limited to small angles with respect to the z axis, the square root in the exponent can be simplified by reducing it to the first two terms of the Binomial expansion

$$U'(u,v) = U_s(u,v) \exp jkz. \exp - j\Pi\lambda z \ (u^2 + v^2)$$

The amplitude distribution in this plane is given by the inverse Fourier transform of the angular spectrum in the plane ξ, η thus

$$V(\xi,\eta) = \exp jkz \ \mathcal{F}^{-1}_{\xi n} \left\{ U_s(u,v) \exp - j\Pi\lambda z \ (u^2 + v^2) \right\}$$

where $\mathcal{F}^{-1}_{\xi n} \{ \cdots \}$ denotes the inverse Fourier transform of the expression inside the brackets, related to the ξ, η plane. This expression can be broken up by the convolution theorem into

$$V(\xi,\eta) = \exp jkz \ \mathcal{F}^{-1}_{\xi n} \left\{ U_s(u,v) \right\} \otimes \mathcal{F}^{-1}_{\xi n} \left\{ \exp - j\Pi\lambda z \ (u^2 + v^2) \right\} \qquad 1.10$$

The second Fourier transform is a standard result which is

$$\mathcal{F}^{-1}_{\xi n} \left\{ \exp j\Pi\lambda z \ (u^2 + v^2) \right\} = \frac{1}{j\lambda z} \exp j \frac{\Pi}{\lambda z} (\xi^2 + \eta^2)$$

The first expression can at the moment, be expressed as a function $F(\xi, \eta)$ which clearly is related to $V(x,y)$ through the change of variables. Equation 1.10., can be written as

$$V(\xi, \eta) = \frac{\exp jkz\ F(\xi, \eta)}{j \lambda z} \otimes \exp \frac{j \Pi}{\lambda z} (\xi^2 + \eta^2)$$

On performing the convolution and noting that

$$F(x,y) = V(x,y)$$

the amplitude distribution at a given point in the field is given by

$$V(\xi, \eta) = \frac{\exp jkz}{j\lambda z} \iint_{-\infty}^{\infty} V(x,y) \exp jk \left\{ \frac{(\xi - x)^2 + (\eta - y)^2}{2z} \right\} dxdy \quad 1.11$$

This is a general expression for the light distribution over a limited range of the field generated by a monochromatic light source. It is valid so long as $V(x,y)$ is a true description of the complex amplitude of the monochromatic light in the source plane. If the field is restricted still further by the Fraunhofer condition (Goodman 1968), that is

$$\frac{2z}{k} \gg (x^2 + \xi^2), \quad (y^2 + \eta^2)$$

then equation 1.11., can be written as

$$V(\xi, \eta) = \frac{\exp jkz}{j\lambda z} \iint_{-\infty}^{\infty} V(x,y) \exp \frac{-j2\Pi}{\lambda z} (\xi x + \eta y)\, dxdy \quad 1.12$$

which when compared with equation 1.1., can be recognised as a product of a constant phase term with a scaled Fourier transform of the original light distribution. The real dimensions in the plane of interest are ξ, η

hence the scaling factor which determines the amplitudes and positions of the spatial frequencies is $1/\lambda z$. Although this is not a perfect Fourier transform relationship, the phase being multiplied by $\exp jkz$, it does offer a very simple first stage of an optical signal processor.

There are, however, two fundamental restrictions regarding the light distribution in the image plane. Primarily, it must be monochromatic as the wavelength is an important factor in determining the scaling of the transform. A spread of wavelengths would introduce a range of transforms with different scalings which would result in a pattern too complex to be useful. Secondly, in defining the Fourier transform integral it was stated, as a condition, that the function must have only a limited number of discontinuities. However, in some light sources, say a C.R.T., screen, a radiator will emit a photon independently of its neighbouring radiator. Thus a discontinuity, which varies randomly in time, exists between every pair of radiators and as such it is impossible to describe the complex amplitude distribution in such a manner as to satisfy the condition for the existance of the Fourier integral.

Although such sources do not naturally produce their Fourier transforms in the far field, it is possible to deduce the light distribution due to such a source by considering it as a composition of point radiators, each of which can take on a different complex amplitude. The amplitude and phase at a point in the field generated by a single point situated at x', y' of complex amplitude V', is given from equation 1.11., as

$$V(\xi,\eta) = \frac{V'}{j\lambda z} \exp j \frac{k}{2z} \left\{ 2z^2 + (\xi - x')^2 + (\eta - y')^2 \right\} \qquad 1.13$$

On comparing this expression with equation 1.6., it can be seen that the complex amplitude V is given by

$$V = \frac{2V'}{j\lambda}$$

Thus equation 1.13., will satisfy the general wave equation. Clearly, so will the propagation from a neighbouring point with a different complex amplitude. Invoking the linearity of the wave equation, then their sum will also satisfy the wave equation. Clearly, this argument

can be extended to an infinite number of source points giving an expression for the light in the field as

$$V(\xi,\eta) = \sum_{n=-\infty}^{\infty}\sum_{m=-\infty}^{\infty} \frac{V(x_n, y_m)}{j\lambda z} \exp j\frac{k}{2z} \left\{ 2z^2 + (\xi - x_n)^2 + (\eta - y_m)^2 \right\} \qquad 1.14$$

If the complex amplitude $V(x_n, y_m)$ of the point sources is expressed in terms of an amplitude and phase distribution

$$V(x_n, y_m) = A(x_n, y_m) \exp - j k \Phi(x_n, y_m)$$

then equation 1.14., becomes

$$V(\xi,\eta) = \sum_{n=-\infty}^{\infty}\sum_{m=-\infty}^{\infty} \frac{A(x_n, y_m)}{j\lambda z} \exp j\frac{k}{2z} \left\{ 2z^2 + (\xi - x_n)^2 + (\eta - y_m)^2 - 2z\Phi(x_n, y_m) \right\} \qquad 1.15$$

Thus even with $\Phi(x_n, y_m)$ as a random function the instantaneous complex amplitude could be evaluated although it does not resemble any defined transform and hence is not very useful from the signal processing viewpoint. If $\Phi(x_n, y_m)$ were a well behaved function then the double summation could be replaced by a double integral and equation 1.15., would be identical to equation 1.11.

Hence equation 1.14., defines the amplitude and phase at any point in a limited field which has been generated by a monochromatic light source. The extension to a polychromatic source is achieved by simply summing a set of expressions like 1.14., one for each wavelength. However, a point of greater significance is that the nature of the source, through the function $\Phi(x_n, y_m)$ determines the exact form of the far field light distribution.

1.4. Real Light Sources

In observing any light pattern, detectors need to react throughout a finite time interval, typically at least one nanosecond in order to record the light distribution. Such periods of time are very long compared with the fundamental period of light wave, hence the records are of the average light distribution during the recording period. Hence if a light distribution related to

the power spectrum of the source is to be faithfully recorded, then the original light distribution must remain constant during the recording process. Apart from imposing a maximum temporal frequency band on the input function, it also demands that the light displaying the function must have a constant phasefront. Clearly, for light sources where individual point radiators emit photons independently of each other the phase across the screen is constantly changing, making it impossible to record a discernable pattern related to the Fourier transform in the far field.

Thus the conditions for producing and recording the Fourier transform in real time of a light source, are that it must be monochromatic and that all the individual radiators must be coupled to each other. The most common way of satisfying these requirements is to present the information on a photographic transparency which is then illuminated by a spatially filtered laser wave.

Other forms of light sources displaying two dimensional functions usually consist of C.R.T., displays for radar, oscilloscope and computer consoles, or arrays of L.E.D.s. These sources are in general presented with the information as an electrical signal. Individual elemental radiators absorb some energy from the electrons carrying the information and emit photons. As the elemental radiators are, in the main, independent of each other, so are the photons that they emit. Hence, there is no coupling between photons from various radiators and as such these sources will not produce a recognisable Fourier transform in the far field.

There is a third way of displaying two dimensional information which does not fall into either of the two above groups. There are light distributions which have a limited degree of coupling between the point radiators. Such distributions occur when an uncoupled source is imaged through an optical system onto the input image plane. Due to the finite size of the optical components and also due to some possible component defects, a point in the original source is mapped onto a finite area. Clearly, secondary point radiators in this area are coupled as the light originated from one point. Although these light distributions presented as real or virtual images are partially coupled, the degree of coupling is not usually sufficient to enable the source to produce a wide spectrum transform.

It is possible to derive expressions for a light distribution in the far field due to various sources in terms of the cross-correlation of the light at any two points in the field. However, such an approach based on the Mutual Coherence Function does not contribute a great deal in this case except mathematical rigour to the above discussion. The main result, that only information presented on a plane monochromatic wave readily produces its Fourier transform appears an inescapable fact at this stage.

CHAPTER 2

Optical Signal Processing

In the previous chapter it was shown, albeit only qualitatively, that optical signal processors appear to have a fundamental advantage, that of an extra dimension over their electronic counterparts. Inevitably, the optical signal processors have several inherent restrictions which must be understood and surmounted if the potential of these optical techniques is to be realised. One common obstacle, developed towards the end of the last chapter, is the constraint that the light source must produce a well defined phasefront. Coupled with the need to have a readily detectable amount of light in the observation plane, thus precluding highly filtered broadband sources, this restriction limits useable sources to lasers.

Various schemes have been proposed and developed in attempts to overcome the dependence on the laser, and these schemes will now be discussed. However, prior to this appraisal, coherent light signal processors will be examined in order to provide an insight into the potential of optical signal processing.

2.1. Coherent Light Signal Processors

It was shown in chapter 1, specifically by equation 1.12 that in a plane at a large distance from the coherent source, the complex light amplitude distribution will represent the Fourier transform of the complex amplitude at the source. However, the condition defining this distance between the plane of observation and the source plane is rather fearsome. For typical values of object size 25 mm. and a light wavelength of 0.63 μm., the required distance was given by

$$z \gg \frac{k\,(x^2 + y^2)}{2}$$

$$\therefore\ z \gg 1.5 \text{ km}$$

which clearly is a prohibitive criterion for a practical signal processor.

This difficulty was overcome by noting that a convex lens will display in the back focal plane the Fourier Transform of the complex amplitude distribution present in the front focal plane. If $V(x,y)$ denotes the complex amplitude in the input plane, then the output plane complex amplitude can be written by approximating the Fresnel-Kirchoff diffraction formula as

$$V(u,v) = \frac{1}{j\lambda r} \int_{-\infty}^{\infty} V(x,y) \exp jkr\,dx\,dy$$

Where r is the distance between any point in the input plane and any point in the output plane. This distance can be evaluated by considering a plane wave in the front focal plane with its phasefront at angles θ and γ relative to the x,y axes. Such a wave will be focussed to a point (u,v) in the output plane, where

$$u = F \sin\theta$$

$$v = F \sin\gamma$$

implying that the distance between any point on the plane wave and the point (u,v) is constant and equal to 2F, where F is the lens focal length. The distance from a point (x,y) on the plane wave to the input plane is

$$-(x \sin\theta + y \sin\gamma)$$

which can be re-written as

$$-\left(\frac{xu}{F} + \frac{yv}{F}\right)$$

giving the distance r as

$$r = 2F - \frac{xu}{F} - \frac{yv}{F}$$

and hence the amplitude distribution in the back focal plane is

$$V(u,v) = \frac{\exp jkF}{j\lambda F} \int_{-\infty}^{\infty} V(x,y) \exp -j \frac{k}{F} (ux + vy)\,dx\,dy \qquad 2.1.$$

which, when compared with equation 1.1. can be recognised, to within a relatively unimportant constant phase term, as a scaled Fourier transform of the complex amplitude in the front focal plane. This result was utilised (Rhodes 1953, Cutrona et al. 1960) to produce a coherent light signal processor, as shown in Figure 2.1.

An input function g (x, y) is presented on a photographic transparency in the first focal plane and is illuminated by a plane monochromatic wavefront. The first lens displays the scaled Fourier transform G(u, v) in the back focal plane, where a complex filter H(u, v) on the second transparency operates simultaneously on all the Fourier frequencies present in the signal. This plane also constitutes the input to the second lens which produces in its back focal plane, the modified and processed information.

The main forms of processing information by this coherent optical technique have been concerned with extracting a given feature from background noise. In this application noise implies all the signals present in the input function apart from the desired information. The desired feature may be either multiplied by the noise, added to the noise or convolved with the noise; each case requiring a different form of filtering.

In the first case where the signal is multiplied by the noise, the required information can, under certain restrictions be extracted by a simple binary mask. A practical situation demonstrating this form of filtering is in removing the half tone dot structure from photographs printed in newspapers. Such a printed photograph can be considered as the original scene s (x, y) multiplied by the convolution of the dot size n(x,y) with the two dimensional lattice delta function COMB $(\frac{x}{X}, \frac{y}{Y})$. Thus with subscript c denoting dummy convolution variables

$$g(x,y) = s(x,y) \left\{ n(x_c, y_c) \otimes \text{COMB}\left(\frac{x_c}{X}, \frac{y_c}{Y}\right) \right\}$$

which when presented on a transparency in the input plane of Figure 2.1., becomes in the transform plane

$$G(u,v) = S(u_c, v_c) \otimes \left\{ N(u_c, v_c).\ \text{COMB}\left(\frac{Xu_c}{\lambda F}, \frac{Yv_c}{\lambda F}\right) \right\}$$

If the original scene has a limited resolution, its transform S (u, v) will be space limited in a region U, V . In order to preserve the information in the original scene the dots must be closely spaced with their size being smaller than the finest detail. Hence the spatial extent of the transform of a single dot N(u,v) is larger than the transform of the original scene. Thus the product of the transform of the original scene with the transform of the dot will be non-zero only within the region U, V. . However, the spacing between successive delta functions in the transform of the COMB function is inversely proportional to the spacing of the dots. Hence if the distance between adjacent delta functions in the COMB transform is greater than the size of the region U, V then successive products will not overlap each other. Thus if a simple binary mask which will transmit only one such product and exclude the others is placed in the transform plane, then the input to the second lens will be

$$G(u,v) = S(u,v)\, N(u,v)$$

which will be transformed to produce in the output plane

$$g(x',y') = s(x'_c, y'_c) \otimes n(x'_c, y'_c)$$

This can be considered as the original scene imaged through an optical system with a point spread function n(x',y') which implies that the half tone dot structure has been removed.

In the second case, where the noise in the input function is additive, the input to the optical processor is of the form

$$g(x,y) = s(x,y) + n(x,y)$$

which becomes in the transform plane

$$G(u,v) = S(u,v) + N(u,v)$$

If this is now multiplied by a filter function which is the complex conjugate of the transform of the sought for signal, then the input to the second lens becomes

$$G(u,v) = S(u,v).S^*(u,v) + N(u,v).S^*(u,v)$$

Noting that

$$\mathcal{F}_{x'y'}\left\{S^*(u,v)\right\} = s(-x', -y')$$

the output of the signal processor is

$$g(x',y') = s(x'_c,y'_c) \otimes s(-x'_c,-y'_c) + n(x'_c,y'_c) \otimes s(-x',-y'_c)$$

This can be recognised as the auto-correlation of the original signal added to the cross-correlation of the signal and the noise. The auto-correlation term compresses all the signal energy into one well defined point, while the cross-correlation term spreads the energy over an appreciable area. Hence the presence of a high intensity point in the output indicates the existance of the signal in the input function. This technique has been widely used in recognising characters and determining their positions in the input plane. From equation 1.2. a character positioned at (x,y) will be transformed to

$$G(u,v) = S(u,v) \exp - j \frac{k}{F} (xu + yv)$$

which when multiplied by the filter and transformed by the second lens becomes

$$g(x',y') = s(x'_c, y'_c) \otimes s(x'_c, y'_c) \otimes \delta(x'_c - x, y'_c - y)$$

which is quite clearly the auto-correlation of the signal centred on (x, y) . Therefore, the presence of the requisite character and its position relative to the input plane centre are well defined.

The last case of filtering to be considered, that of identifying a signal convolved with noise is shown in clearing a blurred photograph. In this case the input function to the signal processor is

$$g(x,y) = s(x_c, y_c) \otimes n(x_c, y_c)$$

which will be transformed to

$$G(u,v) = S(u,v)\, N(u,v)$$

Clearly a filter function $1/N(u,v)$ will remove the noise and isolate the transform of the signal which can then be readily retransformed.

One of the first problems in utilising the above concept is to physically realise the required filters. It has been shown that the phase content of the filter function can be recorded in the standard holographic fashion (Vander Lugt 1964), thus enabling the types of filter necessary in character recognition to be realised. However, a filter for true deconvolution or division in the transform plane is in general impossible. The transform of the noise function will have its normalised amplitude varying between 0 and 1.0, thus its inverse, the filter function amplitude must vary from ∞ to 1.0, which for a passive system is physically unrealisable. A limited degree of deconvolution is possible with filters that only compensate for the phase of the noise transform.

A second difficulty arises from the fact that the filters and the input function are usually recorded on photographic emulsions. During development the emulsion and the acetate base shrink and distort, thus introducing a spatial phase distribution across the transparency, which increases the ambient noise level, hence reducing the signal to noise ratio. This can be overcome by presenting the transparency, whether filter or input function, in a liquid gate where the index matching fluid between parallel, optically flat, glass plates effectively eliminates any film irregularities.

Another source of error stems from the inaccurate positioning of the filter in the transform plane. The physical dimensions in this plane represent Fourier frequencies, hence a misaligned filter modifies a given frequency by an inappropriate amount. It has been shown that for a typical system with a focal length of 1m and an input function size of 25 mm, a reduction in the signal to noise ratio by a factor of 2 will result if the filter position is in error by 7μm for uniform noise and by 3μm for nonuniform noise (Vander Lugt 1967).

Finally, perhaps the most important limitation of the coherent light optical signal processor is its inability to operate in real time. As discussed in section 1.4, the majority of real light sources do not have sufficient coherence to generate their Fourier transforms in the far field. Hence the information displayed by such a source must be reproduced on a photographic film, which in turn must be developed, before being presented to the signal processor. This inability to operate in real time all but nullifies the advantages of optical signal processors over their electronic counterparts.

There are two possible approaches by which this limitation may be overcome; one is to produce a real time two dimensional coherent light modulator, the other is to develop a process that will operate directly on incoherent light signals.

2.2. Real Time Spatial Modulators

Real time spatial modulators are active units intended to transfer the two-dimensional image presented as an incoherent light distribution onto a monochromatic plane wave in real time. One such spatial modulating technique has been to impose the input function onto a continuously moving roll of film. The exposed frames are passed through a rapid developer and onto the input plane of a coherent processor. However, this only reduces the magnitude of the problem without solving it.

A more complete solution has been based on composite crystal devices utilising the crystal's birefringence dependence on electric fields (Vohl et al. 1973). A typical device consists of an anisotropic, uniaxial, electro-optic crystal with a transparent composite insulator-electrode on one face, and only a transparent electrode on the opposite face, as shown in Figure 2.2a. The crystal is also a semiconductor, with a wide energy gap, displaying photo-conductive characteristics.

In the usual mode of operation, a uniform constant electric field is applied across the device. In the absence of light the crystal acts as an insulator which results in the electric field remaining uniformly distributed across it. The writing incoherent light pattern is focused onto the crystal through the transparent electrode and insulator. On reaching the crystal, the photons with an energy greater than the crystal band gap are absorbed, generating electron hole pairs. The electrons are swept away across the crystal leaving the holes trapped at the crystal-insulator interface. These trapped holes reduce the field across the crystal in proportion to their density, as shown in Figure 2.2b. As this charge density pattern is a direct consequence of the input photon density distribution, the resulting, modified field across the crystal represents the initial image intensity.

The crystal, apart from being a semiconductor, is also electro-optic and anisotropic, which means that the dielectric tensor and hence the refractive index in different directions, depends on the external electric field vector. The crystal is cut and aligned in such a manner that the direction of the applied field and the direction of the reading light propagation are parallel to one of the three crystal axes. Then the field dependent refractive index variations along the two remaining crystal axes modify plane waves, polarised in a direction which bisects the angle between the two axes, to elliptically polarised waves. The magnitude of the minor axis of the polarisation ellipse is directly proportional to the strength of the external field.

Thus if the reading light is a monochromatic plane wave polarised in the required direction, the initial image will exist unambiguously as an amplitude distribution present on the minor axis of the polarisation ellipse. This amplitude can be readily isolated by a polariser crossed with the original direction of polarisation. However, it must be noted that the original intensity distribution is now an amplitude distribution and care must be taken if the contrast levels are to be maintained.

The real time transfer of information from incoherent light to coherent light is made possible by placing a band pass interference filter on the insulating layer. The incoherent writing wave with its spectral distribution centred on this band is transmitted and hence focussed onto the crystal. However, the coherent reading wave, at a wavelength outside this transmission band, entering the device at the opposite face, passes through the crystal and is reflected back through it by the interference filter. The information carrying reflected wave leaving the crystal can be readily isolated from the incoming reading wave by means of a beam splitter. Thus the information can be written in on one side and read out through the other in real time.

The successful operation of these crystal devices depends on several important crystal characteristics, which also introduce several problems. The energy gap in the semi-conductor determines which photons will be absorbed and which will be transmitted. Clearly, it must be so chosen as to divide the visible spectrum; absorbing the short wavelength, and hence the more energetic photons, while transmitting the longer wavelength photons. However, the absorbtion characteristics at room temperature of typically used crystals (Oliver et al. 1971) do not have sharp cut-offs. Hence for maximum photon absorption the writing wavelength must be well into the absorbtion spectrum. Similarly, to prevent reading photons from being absorbed and thus obliterating the image information, the reading wavelength must be well clear of the absorbtion spectrum.

This implies, and is bourne out by reported experiments (Oliver et al. 1970 & 1971, Hou et al. 1971) that writing must be achieved with ultra-violet or blue light while reading is performed with red light. Such division is rather unfortunate as the majority of C.R.T. phosphors and L.E.D.s emit green or red light, but not blue (leverenz 1950), which limits the applicability of the device. Clearly there is a trade-off between reducing the band gap to accomodate longer wavelength writing light and preserving the information when reading it.

There exists a second trade-off, namely that between reading intensity and reading time, which is shown by recently reported devices (Lipson, Nienson 1974) having a band gap of 3.25 eV which corresponds to a wavelength of 0.382 μ m,and need a writing energy at 0.366 μ m of 2.5 $\mu J/cm^2$. On readout at 0.633μ m the information is destroyed by 1.2 mJ/cm Another important characteristic, the ability to retain the electric field pattern in the absence of reading light is limited by the finite resistivity of the semiconductor and is limited to several hours (Feinlab,Oliver 1972).

The polarisation of the reading wave is mainly modified by the birefringence variation across the crystal, thus transfering the original information from an electric field distribution to an amplitude variation. However, the polarisation is also perturbed by impurities and strains in the crystal as well as piezo-electric bending caused by the applied field, and thickness variations. Some of the crystals used also display optical activity which rotates the direction of polarisation, but this is automatically overcome when the reading wave travels forward and back through the crystal, the second passage compensating for the rotation incurred in the first, (Nisenson,Iwasa 1972). The other phase perturbations are not so easy to remove (Feinlab,Oliver 1972) and tend to degrade the final image. The efficiency of the Pockels effect, measured as the ratio of light power in the transmitted image to the amount of reading light power falling on the crystal, is approximately 10%, which when taken together with the maximum permissable reading energy gives rather low intensity images.

Finally, the devices compare well with T.V. system quality as regards resolution, typically 100 lines per mm, and frame rates, again a typical write and erase cycle takes 5m sec.

2.3. Incoherent Light Signal Processors

The other possible technique of achieving real time two dimensional signal processing, is to process directly the incoherent light distribution displaying the signal. Optical systems which operate on incoherent light information displays can be divided into the two general signal processing groups described in Chapter 1. The first group, operating in the image plane, utilises the ray properties of light and generally produces correlations

of the input signal with some other predetermined function. The second group based on the wave nature of light produces interference patterns which are related to the input intensity distribution by a recognisable transform.

One of the earliest needs for an incoherent optical correlator arose out of the work performed by X-ray crystalographers. A crystal illuminated by X-rays produced a far field diffraction pattern which was the magnitude of the Fourier transform of its lattice configuration. Taking a transform of this intensity distribution produced the auto-correlation of the lattice configuration. As a direct analysis of the auto-correlation was impossible, an iterative technique was suggested, where the auto-correlation of a proposed structure was compared with the experimental auto-correlation. Obviously a fast and simple correlator was necessary if the iterative technique was to be successful. Robertson (1932) and later others (Haag 1944, Philips 1954, McLochlen 1962) proposed the correlator shown in Figure 2.3.

The two functions g and h to be correlated are presented as transmission variations in two separate planes. A uniform incadescent source is focussed onto the first function, so that any radiating point with a strength determined by the value of the function at that point, will illuminate the entire second function. A ray of intensity dg(x,y) from a typical point (x,y) in the first plane will pass through a point (x',y') in the second plane, where it will be reduced in intensity according to the local transmission factor dh(x',y'). Hence the incremental intensity at P (u, v) in the output plane is given by

$$dP(u,v) = dg(x,y)\, dh(x',y') \qquad 2.2.$$

If the second plane is midway between the output plane, and the first plane, the system will naturally be convergent giving

$$x - u = 2(x' - u)$$
$$x' = \tfrac{1}{2}(x + u) \qquad 2.3.$$

similarly

$$y' = \tfrac{1}{2}(y + v) \qquad 2.4.$$

which on substituting in equation 2.2. gives

$$dP(u,v) = dg(x,y)\, dh\left\{ \tfrac{1}{2}(x+u), \tfrac{1}{2}(y+v) \right\}$$

Clearly the total intensity at P(u,v) is given by the summation of all rays leaving the (x,y) plane, passing through the second plane and falling on the point P

$$P(u,v) = \iint_{-\infty}^{\infty} g(x,y)\, h\left\{ \tfrac{1}{2}(x+u),\ \tfrac{1}{2}(y+v) \right\} dxdy \qquad 2.5.$$

Thus if the function h is to half the scale of g , then the true correlation will be presented in the output plane to the same scale as g . Thus a reasonably fast and very simple technique was developed for displaying, in two dimensional space the cross-correlation of a pair of two dimensional functions. Although the initial incentive came from X-ray crystalography, this correlator does have other potential applications. Several other techniques for producing correlations of two functions have been proposed (Kovasznay 1957, Felstead 1967, Schneider 1975) but they do not offer any real advantages over the system just discussed.

This correlator has been used as the basis of a character recognition system which determines the presence of specific Fourier coefficients in the input signal (Leifer et al. 1969). The Fourier coefficient at a specified spatial frequency, of a given signal, is obtained by presenting the signal g(x, y) in the first plane, and a cosinusoidal function h(x',y') at twice the required spatial frequency in the second plane. If

$$h(x',y') = 1 + \cos 2\Pi(\omega_x x' + \omega_y y')$$

then on substituting into equation 2.5. with appropriate co-ordinate changes, the intensity distribution in the observation plane becomes

$$P(u,v) = \iint_{-\infty}^{\infty} g(x,y) \left\{ 1 + \cos 2\Pi \left\{ \frac{\omega_x}{2}(x+u) + \frac{\omega_y}{2}(y+v) \right\} \right\} dxdy$$

which on using a standard trigonometric identity can be written as

$$P(u,v) = \iint_{-\infty}^{\infty} g(x,y)\, dxdy$$

$$+ \cos 2\Pi \left\{ \frac{\omega_x}{2}u + \frac{\omega_y}{2}v \right\} \iint_{-\infty}^{\infty} g(x,y) \cos 2\Pi \left\{ \frac{\omega_x}{2}x + \frac{\omega_y}{2}y \right\} dxdy$$

$$- \sin 2\Pi \left\{ \frac{\omega_x}{2}u + \frac{\omega_y}{2}v \right\} \iint_{-\infty}^{\infty} g(x,y) \sin 2\Pi \left\{ \frac{\omega_x}{2}x + \frac{\omega_y}{2}y \right\} dxdy$$

where the value of the integrals determines the amplitude and phase of the Fourier coefficient at a frequency $\frac{\omega_x}{2}, \frac{\omega_y}{2}$. This expression can be re-written as

$$P(u,v) = \text{const.} + c.\ \sin 2\Pi \left\{ \frac{\omega_x}{2}u + \frac{\omega_y}{2}v + \Phi \right\}$$

where $c =$

$$\left\{ \left\{ \iint g(x,y) \cos 2\Pi \left\{ \frac{\omega_x}{2}x + \frac{\omega_y}{2}y \right\} dxdy \right\}^2 + \left\{ \iint_{-\infty}^{\infty} g(x,y) \sin 2\Pi \left\{ \frac{\omega_x}{2}x + \frac{\omega_y}{2}y \right\} dxdy \right\}^2 \right\}^{1/2}$$

and $\Phi =$

$$\tan^{-1} \left\{ \frac{\iint_{\infty}^{\infty} g(x,y) \cos 2\Pi \left\{ \frac{\omega_x}{2}x + \frac{\omega_y}{2}y \right\} dxdy}{\iint_{\infty}^{\infty} g(x,y) \sin 2\Pi \left\{ \frac{\omega_x}{2}x + \frac{\omega_y}{2}y \right\} dxdy} \right\}$$

Thus the magnitude and phase of a Fourier coefficient is presented as the contrast and position of the shadow fringes in the observation plane.

To determine more than one Fourier coefficient of a given signal, a composite filter consisting of several spatially separated periodicities is placed in the second plane. The signal is positioned eccentrically in the first plane and is rotated about the system's axis of symmetry, thus scanning the composite filter. As a particular spatial frequency in the filter is illuminated by the signal, shadow fringes at half that frequency, modulated by the Fourier coefficient will appear in the observation plane. A photodetector in this plane will produce a frequency and amplitude modulated signal from which the Fourier coefficients at given frequencies can be decoded by electronic techniques. In order to determine the phase of the coefficient an appropriate reference signal is required.

In view of the more established Fourier systems reviewed in section 2.1., this sampled transform at first appears to be rather limited. However, Leifer et al. have shown that some Fourier coefficients are more useful in differentiating between characters than others, hence only several coefficients need to be known. This can be interpreted as, in Fourier space, highlighting the tail in the letter Q to distinguish it from the letter O. This concept, coupled with the relative ease with which the sampled transform of incoherent light patterns is obtained, has made this character reading technique a practical

system. However, as a more general signal processor it does have several limitations.

Primarily the transform is not presented in two dimensional space, but has to be electronically decoded and hence appears as a time dependent signal. This sequential generation of Fourier coefficients inhibits true real time two dimensional parallel processing. Secondly, only a sampled transform is produced with the number of samples being rather low, typically about four (Leifer et al. 1969, Rogers 1974). Notwithstanding that this may be adequate for recognising characters, it is clearly insufficient to filter and subsequently reconstruct all but the simplest objects.

A different approach to character recognition and matched filtering with incoherent light, centred on optical correlators has been suggested (Lohmann 1968). It was shown that an optical system with a predetermined pupil function will produce an image similar to the output of a matched filter when the input object contains the desired information. An image can be considered as the convolution of the object and the impulse response of the optical system

$$\text{IMAGE} = \text{OBJECT} \otimes \text{IMPULSE RESPONSE}$$

However the impulse response is no more than the Fourier transform of the pupil function, hence the image can be expressed as

$$\text{IMAGE} = \text{OBJECT} \otimes \mathcal{F}\left\{\text{PUPIL FUNCTION}\right\}$$

Clearly, if the pupil function is the inverse Fourier transform of the required character, then the appearance of that character in the object distribution will generate an image comprising a delta function positioned in accordance with the location of the character in the object. This output is identical to that obtained from the more conventional matched filters. The requisite pupil function, containing amplitude and phase information can be readily obtained in the form of a Fraunhofer hologram, recorded either in the standard fashion or generated by a computer. The fact that a Fourier transform, rather than the required inverse is obtained will only invert the information in the image plane, which once noted, is of no consequence.

This system was designed for character recognition only and as such does not readily lend itself to general processing. However, even as a matched filter it does have several limitations. In order to recognise more than one character, several Fraunhofer holograms must be recorded on the same emulsion. Although they can be recorded with different reference waves, great care would have to be exercised when deciding whether the output represents a particular character, or a different character in another position.

Secondly, because the complex amplitude of the impulse response must be recorded accurately, this filtering technique suffers from positioning problems similar to those associated with coherent light processors. Its main advantage over the coherent light systems is that quasi-monochromatic inputs can be used, or if the holograms are recorded on thick emulsions, nominally white light objects can be recognised.

In a further development (Maloney 1971a & 1971b), this system was improved by having a set of spatially independent subholograms as the pupil function. Each subhologram was the Fraunhofer distribution of the required character's distinctive features. However, the improvements achieved appear to be marginal.

A different approach to incoherent light signal processing is to generate an interference pattern from the input signal, where the two are related through a Fresnel or Fraunhofer transform. Such interference can be produced from incoherent sources by wavefront shearing interferometers. Although there are many types of wavefront shearing interferometers, they differ only in the construction and not in the principle of operation.

The underlying concept consists of producing two mutually sheared virtual objects in the virtual input plane from the one real input light source presented in the real input plane. Light from these two virtual sources is brought together in a specified plane where interference occurs and can be observed. Hence a real input object g (x, y) , a function bounded by the region (X, Y) is amplitude divided and sheared by an amount (x_o, y_o) such that a typical point (x, y) becomes situated at ($x_j=x_o+x, y_j=y_o+y$) in the first quadrant of the virtual input plane, and at ($x_3=x_o-x, y_3=y_o-y$) in the corresponding third quadrant, as shown in Figure 2.4. For a quasi-monochromatic input object, the amplitude distribution in the observation plane due to a single point (x_1, y_1) in the first quadrant is given by equation 1.13 as

$$dA_1(u,v) = \frac{dV(x_1,y_1)}{j\lambda F} \exp - j k \Phi(x_1, y_1) \exp \frac{jk}{2F} \left\{ 2F^2 + (x_1-u)^2 + (y_1-u)^2 \right\} \qquad 2.6.$$

where $dV(x_1, y_1)$ is the amplitude of the virtual point and $\exp -jk (x_1,y_1)$ is its initial phase.

Similarly, the amplitude distribution in the observation plane due to a different point in the third quadrant will be

$$d A_3(u,v) = \frac{dV(x_3, y_3)}{j\lambda F} \exp - j k \Phi(x_3, y_3) \exp \frac{jk}{2F} \left\{ 2F^2 + (x_3+u)^2 + (y_3+v)^2 \right\} \qquad 2.7.$$

These two amplitudes will add to produce an interference pattern whose intensity will be

$$dI(u,v) = Re\left\{\left\{dA_1(u,v) + dA_3(u,v)\right\} \cdot \left\{dA_1(u,v) + dA_3(u,v)\right\}^*\right\}$$

which on substituting from equations 2.6 and 2.7. and noting that

$$Re\left\{Z_1 Z_2^*\right\} = Re\left\{Z_1^* Z_2\right\}$$

gives the incremental intensity as

$$dI(u,v) = Re\left\{dI'(x_1,y_1:x_3,y_3) + dI''(x_1,y_1:x_3,y_3)\exp\frac{-jk}{2F}\left\{2u(x_1+x_3) + 2v(y_1+y_3) + (x_3^2-x_1^2) + (y_3^2-y_1^2) + \Phi(x_3,y_3) - \Phi(x_1,y_1)\right\}\right\} \quad 2.8.$$

where $$dI'(x_1,y_1:x_3,y_3) = \frac{dV(x_1,y_1)\,dV^*(x_1,y_1)}{\lambda^2F^2} + \frac{dV(x_3,y_3)\,dV^*(x_3,y_3)}{\lambda^2F^2}$$

$$dI''(x_1,y_1:x_3,y_3) = \frac{2\,dV(x_1,y_1)\,dV^*(x_3,y_3)}{\lambda^2F^2}$$

If the two points at (x_1,y_1) and (x_3,y_3) do not originate from the same physical point on the real incoherent input object, then the difference of the two initial phase terms $\Phi(x_1,y_1)$ and $\Phi(x_3,y_3)$ in the above expression will take on a random value. Thus the interference pattern between any two virtual object points, describing different real object points is not defined analytically.

However, if the two virtual points at (x_1,y_1) and (x_3,y_3) do originate from the same source point, then clearly the difference of the two initial phase terms will be identically zero. Also, (x_1,y_1) become related to (x_3,y_3) through the relationships

$$x_3 = x_1 - 2x_o \qquad y_3 = y_1 - 2y_o$$

Under these conditions equation 2.8 becomes

$$dI(u,v) = Re\left\{dI(x,y) + dI(x,y)\exp\frac{-2jk}{F}\left\{u(x+x_o) + x_o^2 - \frac{xx_o}{2} + v(y+y_o) + y_o^2 - \frac{yy_o}{2}\right\}\right\} \quad 2.9.$$

where

$$dI(x,y) = \frac{2\,dV(x,y)\,dV^*(x,y)}{\lambda^2F^2}$$

Due to the random phase difference between any two real points in the incoherent object, the incremental fringe patterns can add only intensity. Thus the total intensity distribution in the observation plane is given by.

$$I(u,v) = Re\left\{\iint_{-\infty}^{\infty}\frac{I(x,y)}{\lambda^2F^2}\left\{1 + \exp\frac{-2jk}{F}\left\{u(x+x_o) + x_o\left(\frac{x}{2}-x_o\right) + v(y+y_o) + y_o\left(\frac{y}{2}-y_o\right)\right\}\right\}dxdy\right\} \quad 2.10$$

which is a definite, albeit slightly cumbersome, description of the interference intensity in terms of the various input and interferometer parameters. A closer examination of equation 2.10 will reveal that if x_o, y_o are made equal to zero, the interference pattern reduces to a Fourier transform relationship.

The terms in x_o, y_o can be removed from equation 2.10 by noting that they originate from the assymmetric light distribution in the virtual input plane. Hence if one of the virtual objects is given a two dimensional fold with respect to the other, the light distribution in the virtual input plane will be symmetrical and the output interference pattern will be related to the Fourier transform of the input object. It is possible that a wavefront folding and shearing interferometer producing such an output from an incoherent light distribution could realise the potential of optical, real time, two dimensional signal processing.

CHAPTER 3

Wavefront Folding Interferometers

In the previous chapter it was shown that one of the more promising techniques of achieving real time, two dimensional signal processing is based on the wavefront folding and shearing interferometer. The fundamental difference between this interferometer and other shearing interferometers is that the light distribution in the virtual input plane is symmetrical about the interferometer axis. This symmetery, realised by folding one virtual object with respect to the other, is achieved by having an odd number of reflections in one arm and an even number in the other.

3.1. Theoretical Appraisal

It was shown in section 2.3., that the intensity distribution in the observation plane of a wavefront shearing interferometer is given by

$$I(u,v) = \mathrm{Re}\left\{ \frac{1}{\lambda^2 F^2} \iint_{-\infty}^{\infty} I(x,y) \left\{ 1 + \exp \frac{-jk}{2F} \left\{ 4u(x_o + x) + 4x_o(x - x_o) + 4v(y_o + y) + 4y_o(y_o - y) \right\} \right\} dxdy \right\}$$

which, for a folding interferometer, reduces to

$$I(u,v) = \mathrm{Re}\left\{ \frac{1}{\lambda^2 F^2} \iint_{-\infty}^{\infty} I(x,y) \left\{ 1 + \exp \frac{-jk}{2F} \left\{ 4ux + 4vy \right\} \right\} dxdy \right\} \qquad 3.1.$$

Thus the interference pattern in the output plane of the interferometer is related to the input intensity distribution through a scaled Fourier transform.

In order to relate this output pattern directly to the input function, the virtual plane function $I\ (x,y)$ can be replaced by the sheared real input plane function $g\ (x_o+x_r,\ y_o+y_r)$. Hence equation 3.1. becomes

$$I\ (u,v) = Re\left\{ \frac{1}{\lambda^2 F^2} \int_{-\infty}^{\infty} g\ (x_o+x_r,\ y_o+y_r) \left\{ 1 + \exp \frac{-j2k}{F} \left\{ u\ (x_o+x_r) + v\ (y_o+y_r) \right\} \right\} d(x_o+x_r)\, d(y_o+y_r) \right\}$$

3.2

which on using the Fourier shift theorem becomes

$$I\ (u,v) = Re \left\{ \exp\ j\ \frac{2k}{F}\ (ux_o + vy_o) \int_{-\infty}^{\infty} \frac{g(x_r, y_r)}{\lambda^2 F^2} \exp -j\ \frac{2k}{F}\ (ux_r + vy_r)\ dx_r dy_r \right\}$$

$$+ \text{const.}$$

$$= Re \left\{ \exp j\ \frac{2k}{F}\ (ux_o + vy_o)\ \mathcal{F}_{\xi n} \left\{ g(x_r, y_r) \right\} \right\} + \text{const.}$$

3.3

The original input function can be written as the sum of its even and odd parts, hence the transform can also be divided into the sum of the transforms of the even and odd input functions. Thus equation 3.3 can be written as

$$I\ (u,v) = Re \left\{ \exp j\ \frac{2k}{F}\ (ux_o + vy_o) \left\{ G_e(u,v) + G_o(u,v) \right\} \right\} + \text{const.}$$

However, the transforms of purely odd functions are imaginary, and of even functions are real, hence the above expression can be written as

$$I\ (u,v) = G_e(u,v) \cos \frac{2k}{F}\ (ux_o + vy_o) + G_o(u,v) \sin \frac{2k}{F}\ (ux_o + vy_o) + \text{const.}$$

3.4.

The constant in the above expression is given, from equation 3.2. as

$$\text{const.} = \text{Re}\left\{ \iint_{-\infty}^{\infty} g(x_r, y_r)\, dx_r\, dy_r \right\}$$

The magnitude of this constant in relation to the magnitude of the transforms $G_e(u,v)$ and $G_o(u,v)$ can be obtained from the Schwartz inequality. Applied to this instance it states

$$\left| G_e(u,v) + G_o(u,v) \right|^2 \leq \iint_{-\infty}^{\infty} \left| g(x_r, y_r) \right|^2 dx_r dy_r \cdot \iint_{-\infty}^{\infty} \left| \exp - j \frac{2k}{F} (ux_r + vy_r) \right|^2 dx_r dy_r$$

Noting that the second integral on the right hand side tends to unity and that both $G_e(u,v)$ and $G_o(u,v)$ as defined in equation 3.4 are positive and real, the inequality may be written as

$$G_e(u,v) + G_o(u,v) \leq \iint_{-\infty}^{\infty} g(x_r, y_r)\, dxdy \qquad 3.5$$

with the equality occuring at $u=v=o$. Thus the magnitude of the constant is at least equal to, or larger than the magnitude of the Fourier transform. The output pattern presented by equation 3.4 can be interpreted as the Fourier transform of the even part of the input function amplitude modulating cosine carrier fringes, while the transform of the odd part modulates sine carrier fringes. This degree of modulation is always less than unity except at zero spatial frequency when it equals unity. Thus the magnitude of the transform is directly present as a visibility variation of the fringes, and as such it is an intensity distribution which can be directly observed and recorded.

The periodicity of the fringes is in the first instance determined by the interferometer parameters. Thus, within the region where the transform is valid, the periodicity is a well defined constant value. However, in general the frequency of the fringes in the output plane will be modified by two extra factors. Primarily, the positioning of the input function in the real input plane will introduce a constant frequency shift. If an input function $g(x,y)$ is centred on its co-ordinates, then the output pattern will be given by equation 3.4, repeated here for convenience

$$I(u,v) = G_e(u,v) \cos \frac{2k}{F} (ux_o + vy_o) + G_o(u,v) \sin \frac{2k}{F} (ux_o + vy_o) + \text{const.}$$

However, if the input function is now repositioned so that it is centred at (x_ρ, y_ρ) then the transform is given from equation 1.2 as

$$G'(u,v) = G(u,v) \exp - j \frac{2k}{F} (ux_\rho + vy_\rho)$$

hence equation 3.4 becomes

$$I(u,v) = G_e(u,v) \cos \frac{2k}{F} \left\{ u(x_o + x_\rho) + v(y_o + y_\rho) \right\}$$

$$+ G_o(u,v) \sin \frac{2k}{F} \left\{ u(x_o + x_\rho) + v(y_o + y_\rho) \right\} + \text{const.}$$

Thus the position of the input function is quite clearly portrayed in the output plane as a constant fringe frequency shift.

Secondly, the fringe frequency is varied across the output plane by the ratio of the transforms of the odd and even parts of the input function. Equation 3.4 can be expressed, using a standard trigonometric identity in the form

$$I(u,v) = \left\{ G_e^2(u,v) + G_o^2(u,v) \right\}^{\frac{1}{2}} \cos \left\{ \frac{2k}{F} \left\{ u(x_o + x_\rho) + v(y_o + y_\rho) \right\} + \tan^{-1} \left\{ \frac{G_o(u,v)}{G_e(u,v)} \right\} \right\} + \text{const.} \qquad 3.6$$

Where the relative phase angle and depth of modulation at a given point in the output plane uniquely define the Fourier transform of the input function.

Finally, it is worth noting that the transform of a purely odd function is identically zero at zero spatial frequency. Hence the peak of a fringe will always be centred on the output plane origin. Secondly, from equation 3.5 it can be seen that the depth of modulation is a maximum at the origin. Thus the centre of the output plane will always be defined by the brightest fringe.

3.2. Practical Range

The first limit to the validity and extent of the Fourier transform stems directly from the assumption, made in Chapter 1 in deriving equation 1.13, that the plane waves propagating from the source made small angles with the axis. This enables the square root phase term representing the direction cosine in the z direction to be approximated by the first two terms of a binomial expansion. The error involved in such a truncation is

$$E(r) = \frac{r^2}{2!} \cdot \frac{d^2 f(r\theta)}{dr^2} \qquad 0 \leqslant \theta \leqslant 1$$

Therefore, the error in the phase term of equation 2.6 describing the amplitude distribution in the observation plane due to a point in one virtual object is

$$\frac{-\left\{(x-u)^2 + (y-v)^2\right\}^2}{32F\left\{F^2 + (x-u)^2 + (y-v)^2\right\}^{3/2}} \leqslant E(r_{2.6}) \leqslant \frac{\left\{(x-u)^2 + (y-v)^2\right\}^2}{8F^4} \qquad 3.7a$$

Similarly, the error in the phase of equation 2.7 describing the amplitude distribution in the observation plane due to the same real point, but presented in the other virtual object is

$$\frac{-\left\{(x+u)^2 + (y+v)^2\right\}^2}{32F\left\{F^2 + (x+u)^2 + (y+v)^2\right\}^{3/2}} \leqslant E(r_{2.7}) \leqslant \frac{\left\{(x+u)^2 + (y+v)^2\right\}^2}{8F^4} \qquad 3.7b$$

In determining the intensity of the interference pattern produced by the two amplitude distributions of equations 2.6 and 2.7 the difference of the two phase terms becomes the dominant factor.

Hence, the largest possible error in equation 3.1 describing the intensity distribution, will be given by the difference between the largest and smallest errors expressed by equation 3.7. This resulting error will be given by

$$E(r_{2\cdot6} - r_{2\cdot7}) = \frac{\left\{(x-u)^2 + (y-v)^2\right\}^2}{32F\left\{F^2 + (x-u)^2 + (y-v)^2\right\}^{3/2}} + \frac{\left\{(x+u)^2 + (y+v)^2\right\}^2}{8F^4} \qquad 3.8$$

This error must be sufficiently small so that it does not change the phase in equation 3.1 by more than, say $\pi/10$ This implies that

$$kFE(r_1 - r_2) \leqslant \frac{\Pi}{10}$$

$$FE(r_1 - r_2) = \frac{\lambda}{20} \qquad 3.9$$

On substituting equation 3.8 into the above expression and using typical values, $F = 10^3$ mm, $\lambda = 0.55\,\mu$m, $x = y = 10$mm we have condition 3.9 satisfied by $u^2 + v^2 = 110\text{mm}^2$ 3.10

From a more qualitative approach it can be seen that the preceeding approximations made in describing the intensity distribution in the observation plane are those that are usually associated with Fresnel diffraction. However, due to the symmetrical disposition of the virtual objects, the quadratic phase term used to distinguish Fresnel from Fraunhofer regions disappears. Hence the output distribution is the Fraunhofer diffraction pattern within the region bounded by the less stringent Fresnel limits.

A second limitation on the accuracy of the scaled Fourier transform generated by the interferometer stems from a finite degree of coherence in an otherwise incoherent source. In deriving equation 3.1 it was assumed that the time average object intensity due to light from two different, but possibly neighbouring points was identically zero. This implies that any two point radiators in the object are statistically independent and hence that the object is perfectly incoherent.

This need not be always true, as in the case of a source being imaged onto the interferometer input plane by a lens. Clearly a point in the source will become, in the interferometer input plane, an area of light whose distribution will be determined by the point spread function of the lens. For any real optical system this area of light will be significant when compared with the original point radiator in the source. It must also be noted that the light in this imaged area originates from one point and as such will be partially coherent. Thus the interference pattern in the observation plane can no longer be adequately described by equation 3. 4.

The imaged input function in the interferometer input plane can be considered as a convolution of the initial object and the point spread function of the imaging optical system. Hence the distribution in the observation plane can be anticipated to be the product of the far field interference pattern due to the object and that due to the point spread function. The interference pattern due to the point spread function is obtained under conditions of coherent illumination and its significance must be considered separately.

The two most important aspects of the interference pattern produced by the point spread function are that waves from any given secondary point in one virtual object will interfere with all the waves emanating from the other virtual point spread function. Secondly, all the interference patterns between any pair of secondary points in the two virtual objects will add in amplitude, and not in intensity, with all the other interference patterns generated by the remaining secondary points.

The amplitude distribution in the observation plane due to a point radiator in the virtual point spread function present in the first quadrant as shown in Figure 3. 1 is, from equation 1.13, given by

$$A_i(u,v) = \frac{V(x_i, y_i)}{F\lambda} \exp \frac{-jk}{2F} \left\{ 2F^2 + (u-x_i)^2 + (v-y_i)^2 \right\} \qquad 3.11$$

Similarly, the amplitude distribution in the observation plane due to a different point in the same point spread function in the third quadrant will be given by

$$A_3(u,v) = \frac{V(x_3,y_3)}{\lambda F} \exp \frac{-jk}{2F} \left\{ 2F^2 + (u+x_3)^2 + (v+y_3)^2 \right\} \qquad 3.12$$

However, all such waves from the point spread function in the third quadrant will interfere with the wave from the first quadrant, as expressed by equation 3.11. Hence the amplitude of the resulting interference pattern is

$$A(u,v) = A_1(u,v) + \iint_{-\infty}^{\infty} A_3(u,v)\, dx_3 dy_3$$

and the intensity distribution is given by

$$I(u,v) = Re \left\{ A_1(u,v)\, A_1^*(u,v) + \iint_{-\infty}^{\infty} A_3(u,v)\, A_3^*(u,v)\, dx_3 dy_3 \right.$$

$$\left. + 2 \iint_{-\infty}^{\infty} A_1(u,v)\, A_3^*(u,v)\, dx_3 dy_3 \right\}$$

However, such patterns will be generated for every point in the virtual point spread function in the first quadrant. Thus the complete intensity distribution in the observation plane will be given by

$$I(u,v) = Re \left\{ \iint_{-\infty}^{\infty} \left\{ A_1(u,v)\, A_1^*(u,v) + \iint_{-\infty}^{\infty} A_3(u,v)\, A_3^*(u,v)\, dx_3 dy_3 \right.\right.$$

$$\left.\left. + 2 \iint_{-\infty}^{\infty} A_1(u,v)\, A_3^*(u,v)\, dx_3 dy_3 \right\} dx_1 dy_1 \right\}$$

On substituting for $A_1(u,v)$ and $A_3(u,v)$ from equation 3.11 and 3.12 respectively and noting that $x_3 = x - x_1$ the intensity distribution becomes

$$I(u,v) = \text{const} + \text{Re}\left\{ \frac{2}{\lambda^2 F^2} \int_{-\infty}^{\infty}\int\int\int V(x,y) V(x-x_1, y-y_1) \exp\frac{-j2k}{F}(ux_1 + vy_1) \cdot \exp\frac{jk}{F}(ux + vy)\, d(x-x_1)\, d(y-y_1)\, dxdy \right\}$$

The integral in the above expression can be espressed as

$$\int_{-\infty}^{\infty}\int V(x,y)\left\{ \int_{-\infty}^{\infty}\int V^*(x-x_1, y-y_1) \exp\frac{jk}{F}(ux + vy)\, d(x-x_1)\, d(y-y_1) \right\} \cdot \exp\frac{-j2k}{F}(ux+vy)\, dxdy$$

The term inside the brackets can be recognised as the transform of $V^*(x,y)$ shifted by $(-x, -y)$, which on using the Fourier shift theorem, allows the integral to be written as

$$\int_{-\infty}^{\infty}\int V(x,y)\left\{ g^*(u,v) \exp\frac{jk}{F}(ux+vy) \right\} \exp\frac{-j2k}{F}(ux+vy)\, dxdy \qquad 3.13$$

$$= g^*(u,v) \int_{-\infty}^{\infty}\int V(x,y) \exp\frac{jk}{F}(ux+vy)\, dxdy$$

$$= g(u,v)\, g^*(u,v)$$

Therefore, the intensity distribution in the observation plane due to a single point imaged onto the interferometer input by an optical system with a point spread function $P(x,y)$ will be

$$I(u,v) = \text{Re}\left\{ \text{const.} + p(u,v)\, p^*(u,v) \right\}$$

It can be shown from an argument similar to that used in deriving equation 3.4 from equation 3.1 that the above expression may be written as

$$I(u,v) = \text{const.} + \cos\frac{2k}{F}(ux+vy)\left|g_e(u,v)\right|^2 + \sin\frac{2k}{F}(ux+vy)\left|g_o(u,v)\right|^2 \quad 3.14$$

where $g_e(u,v)$ and $g_o(u,v)$ are the even and odd transforms of the amplitude of the point spread function. However, it must be noted that the transform of the point spread function as described in equation 3.13 is half the scale of the transforms obtained for incoherently illuminated objects. This difference in scaling arises because in the former case the transform is taken of a coherent amplitude distribution, while in the later, the transform taken is that of an incoherent intensity distribution.

This consideration does not affect the periodicity of the carrier fringes as they are dependent only on the various interferometer parameters and not on the coherence of the virtual objects. Thus the interference pattern in the observation plane due to a coherently illuminated distribution is the spatial power spectrum of that distribution, amplitude modulating carrier fringes.

The intensity distribution expressed by equation 3.14 can be interpreted as the far field interference pattern produced by a point imaged onto the interferometer input plane by a lens with a given impulse response. If the lens were perfect and of infinite extent, the interferometer would produce two virtual points situated at (x,y) and at $(-x,-y)$. Therefore, if a more general object is considered, consisting of a continuum of points, each of which is imaged onto the input plane, the far field interference pattern for this general object can be obtained by integrating equation 3.14 over the virtual object plane, hence

$$I(u,v) = \text{const.} + \left|g_e(u,v)\right|^2 \iint_{-\infty}^{\infty} I_e(x,y)\cos\frac{2k}{F}(ux+vy)\,dxdy$$

$$+ \left|g_o(u,v)\right|^2 \iint_{-\infty}^{\infty} I_o(x,y)\sin\frac{2k}{F}(ux+vy)\,dxdy$$

Clearly, this can be recognised as the scaled Fourier transform of the original object multiplied by the power spectrum of the point spread function of the lens imaging the object onto the interferometer input plane. It should be noted that both the transform and the power spectrum are centred on the co-ordinate origin of the observation plane. Hence for the typical point spread functions obtained with reasonably well corrected lenses (Born and Wolf 1970) the transform will be affected by the power spectrum only at the higher spatial frequencies. This implies a loss of resolution in the object for lenses with large point spread functions, which is consistent with standard imaging properties of lenses.

Finally, the most crippling limitation of this wavefront folding interferometer is the restriction of input objects to those with narrow spectral widths, in order that the Fourier transform may be displayed unambiguously. The effect of a wide spectral range on the far field interference pattern can be readily appreciated by considering the special case of a polychromatic point situated at $(\pm x, \pm y)$ in the virtual object plane. From equation 3.4 the intensity distribution in the observation plane due to a single wavelength will be

$$dI(u,v) = I(k)\left\{1 + \cos\frac{2k}{F}(ux + vy)\right\} dk$$

A similar pattern will be produced by every wavelength present in the source, however, as the waves at different wavelengths, emanating from the point have a random initial phase, these patterns in the output plane will add only in intensity. Thus the intensity distribution in the observation plane will be

$$I(u,v) = \iint_0^\infty I(k)\left\{1 + \cos\frac{2k}{F}(ux + vy)\right\} dk \qquad 3.15$$

which is the scaled cosine Fourier transform of the spectral distribution.

Thus if the spectrum is a rect function of the form

$$I(k) = I_o \,\mathrm{rect}\left(\frac{k}{\delta k}\right)$$

then the intensity distribution in the observation plane, as given by equation 3.15 becomes

$$I(u,v) = I_o + \frac{I_o \sin {}^{2}/_{F}\,\delta k\,(ux + vy)}{{}^{2}/_{F}\,(ux + vy)} \cos {}^{2k}/_{F}\,(ux + vy)$$

Thus the necessary constant visibility fringes defining the Fourier transform of a point, are modified by the transform of the source spectrum. Clearly the larger the bandwidth, the greater is the departure from the true spatial transform. A working limit on the spectral width can be imposed by requiring that the fringe visibility is reduced by no more than half over the practical region as defined by equation 3.10. Thus

$$\frac{\sin \psi}{\psi} = \frac{1}{2} \qquad \psi = 0.6\,\Pi$$

$$\frac{2\,\delta k}{F}(ux + vy) = 0.6\,\Pi$$

$$\delta\lambda = 0.4\,\mathring{A}$$

Clearly, such a narrow bandwidth excludes most of the standard systems for displaying or generating two-dimensional signals.

Thus, although the interferometer is capable of producing from an extended incoherent source, two-dimensional patterns that are directly related to the Fourier transform of the spatial extent of the source, the inability to accommodate a large spectral bandwidth frustrates its potential for real-time processing of two dimensional signals.

3.3. Realisations

Although Mertz and Young suggested this wavefront folding interferometer, they themselves do not appear to have reported the construction of such a device. It seems that the first successful experiment with this type of interferometer was performed by Murty (1964) who was primarily concerned with generating a high contrast interference pattern.

His interferometer consisted of a Twymann Green with a roof prism in one arm, by means of which a single-fold rotation was introduced. The source used, was a filtered, high pressure mercury discharge lamp focussed onto a narrow slit. The far field interference pattern obtained was the scaled cosine Fourier transform of the slit and this is clearly demonstrated in his paper.

A two-dimensional folding interferometer was built (Stroke and Restrick 1965, Worthington 1966) in an attempt to produce holograms of two-dimensional information illuminated by incoherent light. The former overcame the spectral bandwidth problem by using a gas laser as the source. The required degree of spatial incoherence was obtained by moving a ground glass plate across the laser wave, thus introducing a time varying, random phase variation across the wavefront. Worthington solved the problem by using a highly filtered low pressure mercury discharge lamp.

Both accounts are concerned with developing holography with incoherent light, however, with such stringent restrictions on permissable spectral bandwidths, they can be, at best, considered as monochromatic demonstrations of the interferometer's potential.

The most successful application to date of the interferometer has been in astronomy. In one branch it was used to measure the perturbation of light through turbulent air along a given path (Wendy 1970, Bertolotti 1970). A laser source was located at one end of the path of interest, while the interferometer was at the other. The laser light entering the interferometer had its phase modulated by the air turbulence. This turbulence was presented as a visibility variation of the fringe pattern in the observation plane. However, as a laser was essential, the experiment had to be limited to light propagation parallel to the ground.

The interferometer has also been used (Breckinridge 1972, 1974, Dainty and Scaddon 1974) to estimate the diameter of stars. The former used a complex, solid glass prism which sheared and folded the wavefront in two dimensions, while the latter used a modified Michleson interferometer. However, both interferometers have produced similar results. A star can be considered as consisting of two separate functions; one representing its spatial extent and the other its spectral distribution. As an approximation it can be assumed that the interference pattern generated by the interferometer will be the product of the individual transforms of these two functions.

The transform of the spatial extent of the star will be a circular Bessel function which has well defined zeros. The exact position of these zeros is dependent upon known interferometric parameters and on the star's diameter. If a Gaussian or Lorentzian distribution is assumed for the star's temporal spectrum, the corresponding transform will not have any zeros within the range of the interferometer. Thus the distance between

zeros in the interference pattern can be measured and the star's diameter calculated. Clearly, with a broad spectrum it is difficult to ascertain the exact location of these zeros and a filter may be necessary to reduce the spectral width.

The major advantage of this interferometer over Michleson's Stellar interferometer is that the transform of the star is displayed in space which simplifies the task of determining the exact position of a zero.

Finally, interest has been shown (Lacourt et al 1972) in using this interferometer to obtain the Optical Transfer Function of any optical system. From linear systems theory it is clear that the frequency response of a system is given by the Fourier transform of its impulse response. Therefore, if the point spread function produced by an optical system is focused onto the input plane of the interferometer, the Optical Transfer Function will be presented in the observation plane as a visibility variation. However, this system is limited to monochromatic illumination and hence no information about the chromatic dependence of the optical system is available.

CHAPTER 4

Spectrally Compensated Wavefront Folding Interferometer

In the previous chapter it was shown that the inability to accomodate polychromatic input functions severely limited the potential of the wavefront folding interferometer. The trivial solution of restricting the spectral range of the source by a highly selective interference filter is very inefficient, as only a very small amount of the source energy is made available to form the interference pattern. A more practical solution would be to make the patterns produced by different wavelengths identical, and hence independent of the wavelength.

The initial attempts at overcoming the problems caused by the broad optical spectrum were directed towards producing holograms with incoherent light. One of the first reasonably successful incoherent light holograms produced (Mertz and Young 1961), was based on the concept of considering a hologram as a superposition of Fresnel zone plates produced by individual points in the object (Rogers 1952). Thus the recorded pattern, or hologram, is a convolution of the object distribution with a Fresnel zone plate.

Reconstruction can be achieved by illuminating the hologram with a plane monochromatic wave, in which case the various zone plates refocus their appropriate object points. A second possibility is to correlate the hologram with a zone plate (Silva and Rogers 1975), which will produce the original object convolved with the autocorrelation of the zone plate. As this autocorrelation approaches a delta function, so the object can be observed unambiguously.

Although this technique for overcoming the problems associated with a broad band source has been used reasonably successfully in X-ray and in γ-ray holography (Barret et al. 1972, Wilson et al. 1973, Barret et al. 1973) it does not appear to have the same potential for real time two dimensional signal processing, as did the wavefront folding interferometer. Thus in order to overcome the effects of the temporal bandwidth, it is necessary to look for a form of spectral compensation which will enable the object to generate identical interference patterns for all wavelengths present.

4.1. Required Compensation

Without any spectral compensation, the interferometer produces similar patterns, but with different scalings as shown in section 3.2. Thus a typical virtual point, situated at (x, y) in the virtual object plane, radiating at a wavelength λ will interfere with its corresponding point in the other virtual object to produce a fringe pattern given by

$$I(u,v) = A\left\{1 + \cos\frac{4\Pi}{\lambda F}(ux + vy)\right\} \qquad 4.1$$

The same point also radiates with a wavelength $\lambda + \delta\lambda$ and will produce an interference pattern described by

$$I(u,v) = A\left\{1 + \cos\frac{4\Pi}{(\lambda+\delta\lambda)F}(ux + vy)\right\} \qquad 4.2$$

The two fringe patterns described by equations 4.1. and 4.2., have different periodicities. However, as they both have a maximum at the centre, or coordinate origin, of the observation plane, they will become progressively out of step for increasing positive and negative observation plane coordinates. This phase angle between the two fringe patterns will equally limit the extent of positive and negative Fourier frequencies that can be recognised. If an arbitrary, but working condition restrains the relative phase angle to be no more than π in order to maintain fringe visibility, then the maximum range U, V over which fringes can be said to exist can be readily defined. However, such a description of the fringe pattern depends not only on the spectral distribution of the source, but also on various interferometer parameters. A much more useful description would be in terms of the numbers of fringes that exist within the working range. If within this range U,V, equation 4.1., describes N_+ fringes, then equation 4.2., will describe $N_+ - \frac{1}{2}$ fringes. Therefore

$$\frac{4\Pi}{\lambda F}(Ux + Vy) = N_+ 2\Pi$$

$$\frac{4\Pi}{(\lambda+\delta\lambda)F}(Ux + Vy) = (N_+ - \tfrac{1}{2})\,2\Pi$$

from which

$$N_+ = \frac{\lambda+\delta\lambda}{2\delta\lambda}$$

An expression similar to equation 4.2. can be written for the object point radiating at the other extreme of the source spectrum, namely at $\lambda - \delta\lambda$ and on proceeding through an argument similar to that above, the number of visible fringes will be limited to

$$N_- = \frac{\lambda - \delta\lambda}{2\,\delta\lambda}$$

This merely states that the fringes at the shorter wavelength will fall out of phase sooner than the longer wavelength fringes. However, as the question of whether a fringe is visible or not is rather subjective, for convenience the number of visible fringes produced by a point source of spectral width $2\,\delta\lambda$ centred on λ can be defined as the mean of N_+ and N_-, namely

$$N = \frac{\lambda}{2\,\delta\lambda} \qquad 4.3.$$

This defines the number of fringes in the positive half range of the observation plane, thus the total number of visible fringes produced by a source will be 2N. This gives, for a white light source, the number of detectable fringes as approximately 3.

This number of fringes could be increased if a parameter within the cosine argument in equation 4.2. could be made a function of wavelength, such that the argument became independent of the wavelength. Clearly, there are only two possible parameters; F the distance between the virtual input plane and the observation plane, and (x,y) the virtual input plane coordinates. A brief look at equation 4.2. will reveal that for a maximum bandwidth $2\,\delta\lambda$ approximately equal to λ, F must be halved, or (x,y) must be doubled. For a more practical interferometer F would be 10^3 mm while (x,y) would be 20mm, hence it is easiest, with the readily available dispersive elements, to achieve a controlled display of the spectrum in the virtual input plane. Therefore, if the virtual input plane coordinates are made wavelength dependent, as shown in Figure 4.1., then equation 4.2. will become

$$I(u,v) = A\left\{1 + \cos\frac{4\,\Pi}{(\lambda + \delta\lambda)F}\left\{(x + \delta x)u + (y + \delta y)v\right\}\right\} \qquad 4.4$$

For ideal compensation, the fringe periodicities expressed in equations 4.1. and 4.4. should be identical

$$\frac{4\,\Pi}{\lambda\,F}\,(ux + vy) = \frac{4\,\Pi}{(\lambda+\delta\lambda)F}\left((x+\delta x)\,u + (y+\delta y)\,v\right)$$

which on re-arranging gives

$$\frac{\delta\lambda}{\lambda} = \frac{u\,\delta x + v\,\delta y}{ux + vy} \qquad 4.5.$$

This clearly defines the required displacements in the x and y directions of a point radiating at a wavelength $\lambda + \delta\lambda$ from the same original point radiating at a wavelength λ. An identical expression for negative displacements can be obtained for the object point radiating at a wavelength $\lambda - \delta\lambda$ From now on where such symmetry prevails only the positive half of the spectrum will be considered.

This form of compensation is fundamentally a technique for generating achromatic fringe patterns, and various different systems have been suggested for producing such fringes, mainly with incoherent light holography as a goal. However, as will be shown in the next section, not one of these systems can be successfully applied to wavefront folding interferometers.

4.2. Achromatic Fringe Systems

It would appear that the first successful demonstration of achromatic fringes was achieved by Leith and Upatnieks (1967) although the possibility had been previously discussed (Lohmann 1962). Leith and Upatnieks proposed a system where a collimated beam, from a broad spectrum source illuminated a diffraction grating, as shown in Figure 4.2.

A lens, imaging the grating with unity magnification onto the observation plane, produced in the back focal plane well defined areas of light, consisting of a spatial display of the source spectrum, positioned at each of the diffracted orders. The grating lines, as observed in the image plane can be thought of as fringes

formed by the interference of two or more of the areas of light in the back focal plane, acting as secondary sources. However, in order to calculate the spatial extent and location of these secondary sources, it is convenient and accurate to a first order, to consider the effect of a diffraction grating in terms of light rays. Thus with reference to Figure 4.2., the undiffracted beam will form a polychromatic point on the axis. The extent of the secondary source at the first diffracted order can be found by first considering its size in the plane of the lens, thus

$$\delta x_L = \frac{2F\delta\alpha}{\cos\alpha}$$

where α is defined by the diffraction equation for normal incidence on a grating of N lines per mm.

$$\sin\alpha = \lambda N$$

$$\cos\alpha . \delta\alpha = N\delta\lambda$$

$$\delta x_L = \frac{2FN\delta\lambda}{\cos^2\alpha}$$

The secondary source at the first diffracted order, in the back focal plane will have an extent half that at the lens. Noting that $\tan\alpha$ can be expressed in terms of the focal length and of the distance d between the secondary source and the polychromatic point, we have

$$\delta x = \frac{d\,\delta\lambda}{\lambda\cos\alpha} \qquad 4.6.$$

which defines the spatial extent in the back focal plane of the displayed source spectrum.

The system can be considered as an interferometer with the back focal plane of the lens serving as the input plane, and the image plane as the observation plane, as shown in Figure 4.3. The complex amplitude distribution in the observation plane produced by the polychromatic point situated at $(-x, y)$ radiating at a wavelength λ is given from equation 1.13 as

$$A_0(u,v) = V_0(\lambda) \exp \frac{jk}{2F} \left(2F^2 + (u+x)^2 + (v-y)^2 \right)$$

where $V_0(\lambda)$ is the magnitude of the spectral composition of the zero order diffraction term. Similarly, the amplitude distribution due to a point radiating at the same wavelength, situated in the first order diffraction term at (x, y), is given by

$$A_1(u,v) = V_1(\lambda) \exp \frac{jk}{2F} \left(2F^2 + (u-x)^2 + (v-y)^2 \right) \qquad 4.7.$$

These two distributions produce a fringe pattern whose intensity is given by

$$I(u,v) = \frac{V_0^2(\lambda)}{\lambda^2 F^2} + \frac{V_1^2(\lambda)}{\lambda^2 F^2} + \frac{2 V_0^2(\lambda) V_1(\lambda)}{\lambda^2 F^2} \cos \frac{4 \Pi u x}{\lambda F} \qquad 4.8.$$

Comparable amplitude distributions can be written for the radiation at $\lambda + \delta\lambda$, noting that it now originates at $x + \delta x$ in the dispersed first order, and not from x. Thus the intensity of the fringe pattern produced by points radiating at a wavelength $\lambda + \delta\lambda$ will be given by

$$I(u,v) = \frac{V_0^2(\lambda+\delta\lambda)}{(\lambda+\delta\lambda)^2 F^2} + \frac{V_1^2(\lambda+\delta\lambda)}{(\lambda+\delta\lambda)^2 F^2} + \frac{2 V_0(\lambda+\delta\lambda) V_1(\lambda+\delta\lambda)}{F^2(\lambda+\delta\lambda)^2} \cos \frac{4\Pi}{F(\lambda+\delta\lambda)} \left(u\left(x + \frac{\delta x}{2}\right) - \frac{\delta x}{4}(\delta x + 2x) \right) \qquad 4.9.$$

If achromatic fringes are to be produced, the periodicities expressed in equations 4.8. and 4.9., must be equal, which, on ignoring the constant phase shift $\delta x(\delta x + 2x)/4$ in equation 4.9., gives the condition δx must satisfy, namely

$$\frac{x + \frac{\delta x}{2}}{\lambda + \delta\lambda} = \frac{x}{\lambda}$$

$$\delta x = \frac{2 x \delta\lambda}{\lambda} \qquad 4.10$$

On comparing this condition with equation 4.5. it is clear, that for small diffraction angles, this system will produce a high contrast fringe pattern. Leith and Upatnieks suggested that this system could be used for producing holograms by using the polychromatic point source as the wave illuminating the object, while one of the first order diffracted terms would be the reference wave. Clearly, the object is limited to a transmittance distribution and must be placed in a plane between the back focal plane of the lens and the observation plane. If this object plane is denoted by the coordinates x', y' and is at a distance D from the back focal plane, then the amplitude distribution across the transmittance object $T(x', y')$ will be

$$A(x', y') = \frac{V_o(\lambda)}{\lambda D} \exp \frac{jk}{2D} \left(2D^2 + (x + x')^2 + (y + y')^2 \right)$$

Therefore, a typical point on the object, with a transmittance $T(x'_i y') A(x'_i y')$ will produce an amplitude distribution in the observation plane given by

$$A(u, v) = \frac{T(x', y') A(x' y')}{(F - D)} \exp \frac{-jk}{2(F - D)} \left((F-D)^2 + (u + x')^2 + (v + y')^2 \right)$$

If the object plane is midway between the back focal plane and the observation plane, then

$$2 D = F \qquad \text{and} \qquad \frac{x u}{F} = \frac{2 x x'}{F}$$

On substituting for $A(x', y')$, the amplitude distribution in the observation plane becomes

$$A(u, v) = \frac{T(x', y') V_o(\lambda)}{F \lambda} \exp \frac{jk}{2F} \left(2 F^2 + (x + u)^2 + (y + v)^2 \right)$$

On combining this amplitude distribution with that produced by the reference wave, expressed by equation 4.7., an interference pattern arises, whose intensity is

$$I(u,v) = \frac{V_i^2(\lambda)}{\lambda^2 F^2} + \frac{V_o^2(\lambda)\, T^2(x', y')}{\lambda^2 F^2} + \frac{2\, V_i(\lambda)\, V_o(\lambda)\, T(x',y')}{\lambda^2 F^2} \cos \frac{4\pi u x}{F}$$

4.11.

Thus, the fringes, or the grating lines, are amplitude modulated by the transmittance function $T(x',y')$ and hence the information is recorded holographically with incoherent light. There are however, several fundamental limitations of this technique when applied to real time signal processing, or even to practical holography.

Primarily, the constant phase term in equation 4.9. is a function of wavelength and is only equal to zero for one well defined wavelength. Hence, to justify its omission from the condition stated by equation 4.10., the spectral range of the source must be restricted, typically to about 100 Å.

Secondly, due to the fixed relationship between the polychromatic point and the dispersed first order reference source, the possible objects are limited to a point or to a transparency placed in a well defined plane. Due to the geometry of the system, Fourier holograms are produced only for objects situated on the polychromatic point source, hence these Fourier patterns are limited to single point objects.

An attempt was made to produce Fourier transform holograms from larger objects, by using a Fresnel zone plate, instead of a diffraction grating, as the dispersive element (Kato and Suzuki 1969). The zone plate focussed the undiffracted beam in one plane, and the first order diffracted term in a second plane which became the object plane. The transparency was placed in the object plane such that it would be illuminated by the diverging undiffracted wave. Hence with the reference source and the object present in the same plane a Fourier transform hologram, to within a quadratic phase term, was produced in the far field. However, with the object limited to a transparency, the advantages of incoherent light signal processing rapidly diminish.

Although for image plane holography the object size has been increased by placing it in contact with the diffraction grating (Bryngdahl and Lohmann 1970), the fundamental problem of a restricted spectral range and the need to present the information on a transparency persist. If precise Fourier holograms are desired then the object is limited to a single point in a well defined position.

4.3. Spectral Compensation for Extended Sources

A first possible step to improve the spectral range would be to remove the constant phase term present in equation 4.9. This can be achieved by having two symmetrically dispersed sources producing the interference fringes, in preference to the asymmetric light distribution in Figure 4.3. The required dispersion can be produced by a pair of parallel gratings as shown in Figure 4.4a, from which a wavefront folding interferometer can readily produce two virtual, sheared images as shown in Figure 4.4b.

The first grating sends the light at different wavelengths into different directions, while the second grating translates this angular dispersion into a spatial dispersion. Thus the amount of dispersion δx for a spectral range $\delta \lambda$ introduced is dependent solely on the grating periodicities and the distance separating the gratings. From Figure 4.4a, if P_λ is the path length of the radiation at λ between the gratings, then

$$P_\lambda = \frac{r}{\cos \alpha}$$

where α the diffraction angle is given by

$$\lambda N_1 = \sin \alpha$$

$$\frac{\delta x}{\delta \alpha} = \frac{P_\lambda}{\cos (\alpha + \delta \alpha)}$$

$$\delta x = \frac{r N_1 \delta \lambda}{\cos^2 \alpha \cos (\alpha + \delta \alpha)}$$

which for small diffraction angles reduces to

$$\delta x = \frac{r N_1 \delta \lambda}{\cos^3 \alpha} \qquad 4.12$$

Thus δx is directly proportional to $\delta \lambda$. From Figure 4.4b and section 3.2 it is clear that the pair of points positioned at $\pm x$ radiating

at a wavelength λ will produce a fringe pattern whose intensity will be

$$I(u) = A\left(1 + \cos \frac{4\Pi u x}{\lambda F}\right)$$

Similarly, the pair of points positioned at $\pm(x + \delta x)$ radiating at a wavelength $(\lambda + \delta\lambda)$ will produce a fringe pattern whose intensity is

$$I(u) = A\left(1 + \cos \frac{4\Pi u (x + \delta x)}{F(\lambda + \delta\lambda)}\right)$$

These two periodicities will be equal if

$$\delta x = \frac{x\,\delta\lambda}{\lambda} \qquad 4.13$$

However, δx is fixed, by the geometry of the grating, for a given $\delta\lambda$; thus equation 4.13., defines a unique value of object position x at which the compensation is ideal. Hence, this form of spectral compensation is limited to an object consisting of a single point in a very well defined position. Although this is clearly of limited value for signal processing, an identical, in principle, form of compensation has been used for holography (Leith and Chang 1973, Chang 1973) and for stellar interferometry (Cutler and Lohmann 1974).

However, if in equation 4.12 the dispersion δx could be made linearly dependent on the object plane dimension x then clearly equation 4.13 would be satisfied for all x and thus presenting ideal spectral compensation across the entire object plane. On combining equations 4.12 and 4.13 we have

$$\frac{r N_1 \lambda}{\cos^3 \alpha} = x \qquad 4.14$$

which defines a relationship between the compensating system parameters and

the object plane dimensions. This is the fundamental requisite condition that the gratings must satisfy in order to achieve spectral compensation for extended sources.

One possible way of achieving this is to vary N_1 the number of lines per mm on the first grating. However, on expressing $\cos^3\alpha$ in terms of N_1 equation 4.14 becomes

$$x = \frac{r\, N_1 \lambda}{\left\{1 - N_1^2 \lambda^2\right\}^{3/2}}$$

Varying the number of lines per mm on a grating according to the above relationship is clearly not easy. There is, however, another parameter in equation 4.14, namely r the distance between the gratings which can be made proportional to x. This can be achieved by inclining the second grating relative to the first, as shown in Figure 4.5.

Thus a polychromatic point situated at x along the first grating will be dispersed into a solid angle bounded by α and $\alpha + \delta\alpha$ for wavelengths λ and $\lambda + \delta\lambda$ respectively. From Figure 4.5.,

$$\delta x = \delta x_g \cos\theta$$

$$\frac{\delta x_g}{\delta \alpha} = \frac{P_\lambda}{\cos(\alpha + \delta\alpha + \beta)}$$

$$P_\lambda = \frac{x \sin\beta}{\cos(\alpha + \beta)}$$

$$\delta x = \frac{x \sin\beta \cos\theta\, \delta\alpha}{\cos(\alpha + \beta)\cos(\alpha + \delta\alpha + \beta)} \qquad 4.15$$

The angles in the above expressions are also dependent upon three physical parameters in the dispersive system. These three parameters, the periodicities of the two gratings and the angle between them, are related by the grating equations and the appropriate derivatives, which for the first

grating are

$$\lambda N_1 = \sin \alpha \qquad 4.16\ a$$

$$\delta \lambda N_1 = \cos \alpha \, \delta \alpha \qquad 4.16\ b$$

and for the second grating

$$\lambda N_2 = \sin (\alpha + \beta) - \sin \theta \qquad 4.17\ a$$

$$\delta \lambda N_2 = \cos (\alpha + \beta) \, \delta \alpha - \cos \theta \, \delta \theta \qquad 4.17\ b$$

Substituting for $\delta \alpha$ from equation 4.16b into equation 4.15 gives

$$\delta x = \frac{x \delta \lambda}{\lambda} \left\{ \frac{\sin \beta \cos \theta \tan \alpha}{\cos (\alpha + \beta) \cos (\alpha + \delta \alpha + \beta)} \right\}$$

On comparing the above equation, which describes the wavelength dependent displacement as a function of the grating parameters, with the equation 4.13 which defines the ideal displacements, we have that

$$\frac{\sin \beta \cos \theta \tan \alpha}{\cos (\alpha + \beta) \cos (\alpha + \delta \alpha + \beta)} = 1.0 \qquad 4.18$$

which assuming that $\delta \alpha$ is small compared with $\alpha + \beta$, gives

$$\frac{\sin \beta \cos \theta \tan \alpha}{\cos^2 (\alpha + \beta)} = 1.0 \qquad 4.19$$

This condition remains constant for all wavelengths and hence can be readily satisfied by a suitable combination of the three available parameters. A second condition which was implied at the beginning of this analysis, namely that

light at all wavelengths leaves the second grating at the same angle, gives a second equation. Thus from equation 4.17 $\delta\theta = 0$, hence

$$N_1 \cos(\alpha + \beta) = N_2 \cos\alpha \qquad 4.20$$

These two conditions can be satisfied by a continuous range of the three variables. However, due to the complicated nature of these conditions an analytic expression for the possible range of solutions would be far too complex to be useful. Hence, the range of solutions was derived by numerical methods on a computer, and is shown in Figure 4.6. Thus any pair of gratings read off from the curve and inclined at the appropriate angle will produce a nominally infinite number of fringes.

However, in deriving equation 4.19 it was assumed that $\delta\alpha$ was small compared with $\alpha + \beta$. The effect of this assumption can be observed by estimating the limit imposed by the omitted term on the number of visible fringes. Thus with equation 4.19 satisfied we have

$$\delta x = \frac{x\,\delta\lambda}{\lambda}$$

However, if equation 4.19 is replaced by equation 4.18., but keeping equation 4.19., satisfied, then we have

$$\delta x = \frac{x\,\delta\lambda}{\lambda} \left\{ 1 + \varepsilon\left\{\frac{\delta\lambda}{\lambda}\right\} \right\} \qquad 4.21$$

where $\varepsilon\left\{\frac{\delta\lambda}{\lambda}\right\}$ represents the departure from ideal displacement caused by the difference in equation 4.18., and 4.19.

If the number of visible fringes in the positive half range of the observation plane is defined as in section 4.1. then a pair of points situated at $\pm x$ radiating at a wavelength λ will produce N_+ fringes such that

$$2\pi N_+ = \frac{4\pi u x}{F\lambda}$$

and

$$2\pi (N_+ - \tfrac{1}{2}) = \frac{4\pi u (x + \delta x)}{F(\lambda + \delta\lambda)}$$

which on combining and re-arranging gives

$$N_+ = \frac{x(\lambda + \delta\lambda)}{2(x\delta\lambda - \lambda\delta x)}$$

Substituting for δx from equation 4.21 and eliminating x gives

$$N_+ = \frac{\lambda + \delta\lambda}{2\delta\lambda\, \varepsilon\left\{\frac{\delta\lambda}{\lambda}\right\}}$$

Similarly

$$N_- = \frac{\lambda - \delta\lambda}{2\delta\lambda\, \varepsilon\left\{\frac{\delta\lambda}{\lambda}\right\}}$$

hence

$$N = \frac{\lambda}{2\delta\lambda\, \varepsilon\left\{\frac{\delta\lambda}{\lambda}\right\}} \qquad 4.22$$

Clearly if the error term $\varepsilon\left\{\frac{\delta\lambda}{\lambda}\right\}$ is equal to zero then N tends to infinity and the spectral compensation is perfect. On the other hand, if this error term equals unity then equation 4.22 reduces to equation 4.3 which defines the number of fringes for an uncompensated polychromatic source, and the effect of the spectral compensation has been eliminated.

This error term $\varepsilon\left\{\frac{\delta\lambda}{\lambda}\right\}$ has been evaluated for the ideal values of the three grating parameters as defined by Figure 4.6. The results for the positive half of the spectral width are shown in Figure 4.7. Those for the negative half of the spectrum are not shown as they are similar in form to those shown, differing only by a negative sign and a slight reduction in magnitude.

The main point to note is that for the full visible spectrum, $\delta\lambda/\lambda = 0.3$ the error is far from insignificant offering at best only five times as many fringes than for the uncompensated source. The main reson for such a large error at small diffraction angles is that although α, and hence $\delta\alpha$ in equation 4.18 are small, the angle between the gratings β needs to be large for grating pairs with low numbers of lines per mm. Thus within the argument of a cosine even a small $\delta\alpha$ when compared with a large angle can cause a significant change in the resulting value of the cosine.

The physical significance of this error term $\varepsilon\left\{\frac{\delta\lambda}{\lambda}\right\}$ is that the spectral components of the source are not displaced linearly as required by equation 4.18., hence they give rise to an incremental fringe pattern which after a while falls out of step with the other incremental fringe pattern thus limiting the number of constructive, hence visible fringes. However, it should be noted that although the spectral compensation is not ideal it is independent of the position of the polychromatic source and hence independent of source size.

4.4 Practical Limits and Aberrations

The error term $\varepsilon\left\{\frac{\delta\lambda}{\lambda}\right\}$ had, in the preceeding section, been determined only by a theoretical consideration, however, a departure by any one or a combination of all of the three grating parameters from their ideal values will also introduce a similar error term which will influence the performance of the compensating system. From the previous section, the ideal relationship between the displacement and bandwidth is

$$\delta x_i = \frac{x\,\delta\lambda}{\lambda} \left\{ \frac{\sin\beta\cos\theta\tan\alpha}{\cos^2(\alpha+\beta)} \right\} \qquad 4.23$$

provided that

$$\delta\theta = \delta\lambda \left\{ \frac{N_1\cos(\alpha+\beta) - N_2\cos\alpha}{\cos\theta\cos\alpha} \right\} = 0 \qquad 4.24$$

Clearly, a small variation in any of the three grating parameters will not only upset the coefficient of x in equation 4.23., but will also introduce a similar error term which will in turn further perturb the relationship between δx and x. Thus the real δx_r becomes

$$\delta x_r = \frac{x\,\delta\lambda}{\lambda} \left\{ 1 + \varepsilon(\delta N_1, \delta N_2, \beta) \right\}$$

where as previously, $\varepsilon(\delta N_1, \delta N_2, \beta)$ represents the departure of the coefficient of x from the ideal. As suggested earlier, it consists of two parts; ε_1, due to a non zero $\delta\theta$ and ε_2, a direct result of variation of the grating parameters in equation 4.23.

From Figure 4.8 it can be seen that a non zero $\delta\theta$ decreases δx_i by an amount δx_ε

where $\delta x_\varepsilon = P_{\lambda+\delta\lambda}\,\delta\theta$

and hence

$$\varepsilon_1 = \frac{P_{\lambda+\delta\lambda}\,\delta\theta}{\delta x_i}$$

From Figure 4.8

$$P_{\lambda+\delta\lambda} = \frac{x \sin\beta}{\cos(\alpha+\delta\alpha+\beta)}$$

while from equation 4.24

$$\delta\theta = \frac{\partial\delta\theta}{\partial N_1}\Delta N_1 + \frac{\partial\delta\theta}{\partial N_2}\Delta N_2 + \frac{\partial\delta\theta}{\partial\beta}\Delta\beta$$

Hence ε_1 can be evaluated. The second contribution ε_2 to the overall error is given by

$$\varepsilon_2 = \frac{\partial\delta x_i}{\partial N_1}\Delta N_1 + \frac{\partial\delta x_i}{\partial N_2}\Delta N_2 + \frac{\partial\delta x_i}{\partial\beta}\Delta\beta$$

The total error term $\varepsilon_1 + \varepsilon_2$ has been evaluated and is shown in Figure 4.9.

The main point to note about these error curves is that it is possible to compensate for the fixed errors in N_1 and N_2 by introducing a controlled error into β. Thus within the assumption of small errors it is possible to tune the compensating system to an optimum level.

This process, in fact, optimises the equation

$$N = \frac{\lambda}{2\,\delta\lambda\,\varepsilon}$$

Thus for ε less than unity, the compensating system enables more fringes to be formed than would have been possible when using the source by itself. As ε approaches unity the number of extra fringes decreased until at ε equal to unity the number of fringes produced by the compensated and uncompensated systems is the same.

However, an interesting point now arises, in that if ϵ increases beyond unity, then this system will produce fewer fringes than would have been produced by the source alone.

The readily available commercial gratings are obtained by cutting a master grating and using it to press out the grating pattern on a suitable deformable material. Hence two types of errors can arise. Primarily, local variations in the numbers of lines per mm due to microscopic instability of the master cutting tool and secondly due to pockets of impurity in the replica material. The second type of error, an overall variation in the number of lines per mm is due to the contraction of the replica material after the pressing operation.

However, with existing control methods, manufacturers do not expect the number of lines per mm over a given area to vary more than one or two per cent.

Finally, the gratings, even idealised ones, will produce image aberrations when used in an optical focusing system. Thus a real, monochromatic point source situated at S when viewed through a grating will appear as an extended point situated at V as in Figure 4.10. In order to calculate the wavefront aberrations present, the method suggested by Beuther (1945) and developed by Gillieson (1949) of expanding, as a power series, the path length difference between the principle ray and any typical ray will be used. Thus the path length difference W is given by

$$W = SB - VB - (SA - VA)$$

Expressions for the four lengths can be readily obtained by simple geometry from Figure 4.10 to give

$$W = (U^2 + L^2)^{\frac{1}{2}} - L + \frac{U}{\sin(\alpha - \alpha')} \left\{ \cos\alpha - \cos\alpha' \right\}$$

Expanding the square root term and the trigonometric function as power series gives

$$W = L \left\{ 1 + \frac{U^2}{2L^2} - \frac{U^4}{8L^4} \ldots \right\} - L + \frac{U \left\{ 1 - \frac{\alpha'^2}{2} + \frac{\alpha'^4}{4} \ldots - 1 + \frac{\alpha^2}{2} - \frac{\alpha^4}{4} \ldots \right\}}{\left\{ \alpha - \alpha' - \frac{(\alpha\ \alpha')}{3!} - \frac{(\alpha\ \alpha')}{5!} - \ldots \right\}}$$

However, in this particular circumstance the wavefront folding interferometer requires paraxial rays, implying that $L \gg U$ and that ψ in Figure 4.10 is small. If a further condition is added namely that the number of lines per mm on the grating is small, giving rise to small diffraction angles, then the aberration can be written as

$$W = \frac{U^2}{2L^2} + \frac{U}{2}(\alpha + \alpha')$$

while the grating equation can be written as

$$\lambda N_1 = \alpha$$

$$\lambda N_1 = \alpha' + \psi$$

Substituting for α and α' gives the aberrations as

$$W = \frac{U^2}{2L^2} + \frac{U}{2}(2\lambda N_1 - \psi)$$

From Figure 4.10 can be expressed as

$$\psi = \frac{U}{L}$$

$$W = \lambda N_1 U$$

Thus for paraxial propagation the one dimensional aberrations introduced by a low spatial frequency grating are limited to a linear phase shift which signifies that a non zero diffraction order is being considered. This is a result similar to that obtained by Bose (1977) who investigated the effect of a diffraction grating on a point source in terms of its angular spectrum of plane waves.

CHAPTER 5

The Construction of a White Light Interferometer

In the two preceeding chapters the theory behind the wavefront folding interferometer and the spectrum compensating system, has been developed. It is now necessary to unite and test these concepts in a practical interferometer. After passing through several designs, the experimental interferometer eventually built is shown in Figure 5.1. The interferometer was one dimensional as it was felt that this would adequately test the overall concepts while keeping constructional and alignment problems to a minimum. The interferometer components were mounted on a large steel table, which rested on six pneumatic tyre tubes to isolate it from the building's vibrations.

5.1. The Design of the Interferometer

The light source and the spectrum compensating diffraction gratings were mounted on one platform which could be moved relative to the remainder of the interferometer. The unwanted diffracted orders were removed by a spatial filter; thus the effective input to the wavefront folding part of the interferometer was a unique bundle of chromatically dispersed rays. Wavefront division was achieved by a glass block coated with a 50% transmitting - reflecting layer of aluminium on the front surface. This block was aligned at Brewster's angle in the horizontal plane relative to the incoming rays so that reflections from the rear glass-air interface would be kept to a minimum.

The wavefront rotation in one dimension, the horizontal, was achieved by having an extra reflection in one of the arms of the interferometer. This requirement, together with the wavefront division could be satisfied by a single reflecting surface placed in the manner of Lloyd's mirror. However, such a design does not offer any control over the relative path lengths, while the positioning of the output plane is dependent upon the relative separation of the real and virtual sources. For experimental purposes a more flexible design was necessary.

In order to control the path length in one arm and hence, the path difference between the two arms, a pair of mirrors was mounted orthogonally on a common platform. Thus light incident on the first mirror will leave the second mirror in a direction parallel to that of the incident light. On moving the platform in a direction parallel to the incoming light, the path length can be varied without upsetting the symmetery of the interferometer.

The separation of the two virtual sources can be introduced by one of several possible techniques. However, for a flexible interferometer two conditions must be satisfied; the separation must have a positive and negative range without a discontinuity at zero, and secondly the output plane must remain independent of the separation. The first condition can be satisfied by bringing the light waves in the two arms together with a semi-silvered beam splitter. The second requirement can be fulfilled by noting that the interferometer produces from one real source two virtual sources which are symmetrically displayed in the virtual input plane. Thus a motion of the real source in a direction orthogonal to the principle ray will cause the two virtual sources to move in opposite directions. This motion of the virtual sources can be towards each other or apart from each other depending upon the direction of folding introduced by the interferometer and upon the direction of motion of the real source.

Thus virtual object separation was achieved by mounting the real source on an input table which could be driven perpendicularly to the interferometer optic axis.

5.1.1. The Components Used

Ideally the input to the interferometer should be a self-luminous, spatially distributed scene; for example a C.R.T., screen. However, the light has to pass through two diffraction gratings and radiate through approximately one metre of space, hence in order to have an easily observable interference pattern in the output plane a more powerful input source was necessary. Thus the ideal input was simulated by focussing the filament of a tungsten halogen projector lamp onto a photographic transparency located in the effective input plane. The lamp used was a standard 250 watt projector lamp, type Al/223 with a filament size of 7mm by 3.5mm producing a nominal $8.5\ 10^3$ lumens. Under an optimum filament operating temperature of $3200^{\circ}K$ it had a spectral power distribution as shown in the manufacturer's data, reproduced in Figure 5.2.

The photographic transparency was mounted on a frame which had a pair of horizontally and vertically adjustable jaws defining the input function aperture. The filament of the projector lamp was focussed, with unit magnification by a 4cm diameter lens, onto the centre of this aperture. Thus adjusting the jaws controlled the spatial extent of the input object to a maximum determined by the width of the filament.

It appeared at first that choosing a pair of diffraction gratings to optimise the various conditions expounded towards the end of the last chapter would be quite difficult. However, the problem was reduced by the limited range of readily available gratings, to finding a pair that were close to the curve of Figure 4.6. The gratings eventually chosen had .590 lines per μm for the first grating and .250 lines per μm for the second.

From Figure 4.6. it can be seen that for

$$N_1 = 0.590$$

$$N_2 = 0.248 \qquad \text{and} \qquad \beta = 47.63$$

But N_2 is fixed at .250, thus introducing an error of $\delta N_2 = -0.002$ lines per μm, which can be compensated for by an intentional error in β. Thus from equation 4.26 we have for zero residual error and for $T_1 = 0$.

$$\delta\beta = - \frac{T_2 \lambda \delta N}{T_3}$$

which on using the data in Figure 4.9., gives

$$\delta\beta = - 0.01$$

Thus the initial grating parameters that were used in the experiments are:

$N_1 = 0.59$ $\qquad$ $N_2 = 0.25$ $\qquad$ $\beta = 47.62$

The first grating, mounted in a frame, was attached securely to the input platform with its plane orthogonal to the optic axis. The second grating was fixed to an arm which could rotate about a vertical axis located in the plane of the first grating. The arrangement is shown in Figure 5.3. The gratings were arranged such that the two diffracting surfaces were as close to each other as possible in order that the theoretical model in Figure 4.8., be maintained.

However, due to the finite thickness of the mounts and that the grating lines close to the edge will be distorted, the real object did not start at the origin of the object plane. This offset of the object from the junction of the gratings must be reproduced in the virtual object plane in order that the increased dispersion across the object will be correctly positioned.

The gratings were blazed such that an appreciable amount of the incident energy was transferred into one of the diffracted orders. Hence the gratings were also organised such that the desired optical path, as shown by the solid line in Figure 5.3., was along the enhanced diffracted orders. However, the amount of energy in the other diffraction terms emerging from the second grating was not insignificant and had to be removed by an appropriately positioned spatial filter.

Finally, the selected diffraction order passed through a polariser which transmitted only the vertically polarised components of the light waves. This was necessary to enable the reflections in the interferometer to be restricted to metal - air interfaces.

The beam splitters used were 51 mm square and 25.4 mm thick and were made from a glass of refractive index 1.5174 at a wavelength of .6328 μm. The surfaces were flat to 1 λ and parallel to 0.5 minutes of arc. The first beam splitter was mounted in a frame that could rotate about the vertical axis and could introduce small tilts about an orthogonal pair of axes in the plane of the steel table. The positioning of the second beam splitter was very critical as it affected both the separation of the virtual sources and the path lengths of both arms. Therefore, the mount for this beam splitter had the same controls as for the first mount together with a pair of micrometers that could move the beam splitter in a plane parallel to the table. One face of each beam splitter was coated with a semi-transparent layer of aluminium which gave an equal reflectance: transmission ratio for red light incident at Brewster's angle.

With these beam splitters orientated to the incoming light at such high angles of incidence, the dispersive properties of glass, coupled with the appreciable thickness of the glass will introduce a significant lateral spread of the spectrum. As can be seen from Figure 5.4a, each beam splitter affects the light in one arm only, however, in the output plane the dispersion in both arms is of equal magnitude and is in the same direction. The amount of shift δp for an increase in the wavelength from λ to $\lambda + \delta \lambda$ can be obtained by simple geometry from a consideration of the rays in either block. Thus

$$\delta p' = \frac{t \, \delta r}{\cos r \cos (r + \delta r)}$$

$$\delta p = \frac{t \cos i \, \delta r}{\cos (r + \delta r) \cos r}$$

The relationship between δr, the change in the refraction angle and $\delta \lambda$ the change in the wavelength can be obtained from Snell's law and from the Cauchy formula for the refractive index. From Snell's law we have

$$\sin r = \frac{\sin i}{n}$$

$$\delta r = - \frac{\sin i \, \delta n}{n \cos r}$$

From Cauchy's formula we have

$$n (\lambda) = A + \frac{B}{\lambda^2}$$

where A and B are constants characteristic of the glass. Hence

$$\delta n = - \frac{2B}{\lambda^3} \delta \lambda$$

which on substituting gives

$$\delta p = \frac{2 B t \sin r \cos i}{\lambda^3 \cos^2 r \cos (r + \delta r) n}$$

which can be approximated to

$$\delta p = \frac{2 B t \sin r \cos i}{(\lambda \cos r)^3 n}$$

For the beam splitters used, $B = 3.7219\ 10^{-3}$. Hence for polychromatic light incident at Brewster's angle

$$\delta p = 571.4\ \delta \lambda$$

This can be compared with the dispersion introduced by the spectrum compensating system. From equation 4.13 we have

$$\delta x = \frac{x}{\lambda} \delta \lambda$$

hence for $x = 1$ mm

$$\delta x = 1.8\ 10^3\ \delta \lambda$$

while for $x = 9$ mm

$$\delta x = 1.64\ 10^4\ \delta \lambda$$

Thus δp represents a change of 31.4% at the near edge of the object and 3.5% at the far edge. This will probably manifest itself as a space

dependent resolution variation across the object.

Finally, the mirrors in the interferometer were front silvered, covered with a transparent protective coating and were flat to $\lambda/10$ over the usable diameter of 45mm.

5.1.2. Limitations due to Components

Due to the finite errors present in the various components, the interference pattern observed will depart from the ideal. In the first instance the two phase fronts brought together by the final beam splitter will have small, but not necessarily insignificant, differences caused by the minute surface variation in the mirrors and beam splitters. This phase difference, which will be space variant, can be represented as a phase perturbation $\Phi(x)$ in one of the virtual objects. The magnitude of this phase perturbation will have a maximum contribution of $0.3\ \lambda$ from the three mirrors and $2\ \lambda$ from the beam splitters. A second perturbation of the wavefronts is caused by the dispersion introduced by the beam splitters. This can be modelled by assuming that the ideally displayed spectrum is shifted by an amount $-\delta p$ in both virtual objects.

Thus a point source, situated at x, radiating with a wavelength λ will combine with its corresponding point, positioned at $-x$ and entertaining a phase lag $\Phi(x)$ in the other virtual object, to produce an interference pattern whose intensity is given by

$$I_1(u) = 1 + \cos \frac{2\Pi}{\lambda} \left\{ \frac{2xu}{F} + \Phi(x) \right\}$$

The same point in the real object but radiating with a wavelength $\lambda + \delta\lambda$ will be positioned at $x + \delta x - \delta p$ in the first virtual object and at $-(x + \delta x + \delta p)$ in the second. Taking into account the relative phase difference $\Phi(x + \delta x)$ these two virtual points will produce an interference pattern whose intensity is

$$I_2(u) = 1 + \cos \frac{2\Pi}{(\lambda + \delta\lambda)} \left\{ \frac{2(x + \delta x)(u + \delta p)}{F} + \Phi(x + \delta x) \right\}$$

The maximum of the central fringe of the first pattern is located at

$$u_1 = \frac{F\Phi(x)}{2x}$$

while that due to the second pattern is located at

$$u_2 = \frac{F\Phi(x+\delta x)}{2(x+\delta x)} - \delta p$$

Thus the two interference patterns add with a relative shift Δu which when compared with the fringe wavelength Λ becomes

$$\frac{\Delta u}{\Lambda} = \frac{\Phi(x+\delta x) - \Phi(x)}{\lambda} - \frac{2x\,\delta p}{\lambda F}$$

The first term on the right hand side can take on a maximum value of approximately 2, however due to the random nature of the errors in the component surfaces, the overall shift should be small. The second term, however, is not a random distribution and although it has a maximum value of only 0.5 it will probably be the dominant term. The overall effect will be to reduce the visibility of the output interference pattern.

A problem which created difficulties in performing the experiment was the low light intensity in the observation plane. This situation arose out of the inefficient transfer of light by the gratings and the small solid angle that the boundaries of the observation plane imposed by the aperture of the mirrors made with the real input plane. If the projector lamp filament is considered to be a lambertian source with a total irradiance I_o and that ψ is the total solid angle subtended by the extent of the observation plane at the input plane, then the intensity at the observation plane will be

$$I(u,v) = \frac{2\int_{\varphi=0}^{\varphi=\pi}\int_{\rho=0}^{\rho=\psi/2} \varphi\cos\rho\,d\varphi\,d\rho}{2\int_{\varphi=0}^{\varphi=\pi}\int_{\rho=0}^{\rho=\pi/2} \varphi\cos\rho\,d\varphi\,d\rho} \qquad I_o = I_o \sin\frac{\psi}{2}$$

Although the diffraction gratings were blazed, the first grating transmitted only 34.3% of the incident light intensity into the dominant order, while the second grating was more efficient at 57.2%. Thus the spectrum compensating system passed only 19.6% of the incident power. The polariser removed another 50% and the final beam splitter removed 50% from each beam, thus the intensity in the observation plane became

$$I(u,v) = 4.9 \quad 10^{-2} \; I_o \sin \frac{\psi}{2}$$

The solid angle ψ was determined by mirror 3 in Figure 5.1., which had an aperture of 40mm and was typically 800mm from the input plane. Thus

$$\sin \frac{\psi}{2} = \frac{1}{40}$$

The area of the observation plane was typically $\Pi \; (25)^2 mm$ and hence the intensity per square mm. in the observation plane was

$$I(u,v) = 0.624 \; I_o \; \mu \text{lumens} \quad mm^{-2}$$

which for

$$I_o = 8.5 \quad 10^{-3} \text{ lumens}$$

gives

$$I(u,v) = 5.3m \text{ lumens } mm^{-2}$$

However, as was demonstrated in Chapter 1, the spatial extent of the Fourier transform is inversely related to the extent of the object. Thus

in order to observe a significent number of fringes a broad spatial transform is necessary, which implies a narrow object. Exact values will be given and discussed in the next chapter, however, suffice it to say that an expedient object size was several μ m rather than 8mm. Thus the intensity per square mm in the observation plane would be reduced by a factor of 10^3 to

$$I(u,v) = 5.3\ \mu \text{ lumens} \quad \text{mm}^{-2}$$

Unfortunately, the threshold of photoptic vision is 10 μ lumens mm^{-2}. This problem was deferred to a manageable state by using, as the input object, photographic transparencies with large apertures, containing structured transmission patterns.

5.2. Fringe Localisation

As shown in Chapter 3, the Fourier transform of the object is presented as a variation of fringe visibility. It has, however, been tacitly assumed that fringes do exist in the observation plane. Clearly, if the two arms of the interferometer are identical in all respects, then fringes must potentially exist in the observation plane. However, it is interesting to ask where else could the fringes exist and what determines the exact location of such interference fringes.

Consider two sources, either real or virtual, which have been sheared relative to each other by an amount x_s and rotated through an angle ϕ as shown in Figure 5.5. In order that a high contrast interference pattern can exist at a point P at a distance z from the source plane, the optical path lengths from a typical pair of complementary radiators in the sources must remain constant throughout the range of radiation angles. This implies that in Figure 5.5.

$$\frac{d}{d\psi}\left(\Delta r\right) = 0$$

From straightforward geometry we have

$$\Delta r = - \frac{x \sin\phi \sin(\varphi - \psi)}{\sin(\varphi - \phi)}$$

which on expressing ϕ as $\tan^{-1}(y_s/x_s)$ and re-arranging, gives

$$\Delta r = (x_s + x)\tan\phi\cos\psi - x_s\sin\psi$$

Differentiating Δr with respect to ψ and noting that x and ψ are related, we obtain

$$\frac{d}{d\psi}\left(\Delta r\right) = \cos\psi\tan\phi\frac{dx}{d\psi} - (x_s + x)\tan\phi\sin\psi - x_s\cos\psi \qquad 5.1.$$

From the geometry in Figure 5. 5. an incremental change of ψ can be seen to give rise to a change in x defined by

$$\delta x = \frac{z\,\delta\psi}{\cos^2\psi}$$

Thus equation 5.1 becomes

$$\frac{d}{d\psi}\left(\Delta r\right) = \frac{z\tan\phi}{\cos\psi} - (x_s + x)\tan\phi\sin\psi - x_s\cos\psi$$

which on equating with zero, gives the distance to the plane of highest contrast fringes, from the source plane, as

$$z = \tfrac{1}{2}x\sin 2\psi + x_s\cos\psi\left(\sin\psi - \frac{\cos\psi}{\tan\phi}\right) \qquad 5.2.$$

From the above equation it can be seen that for no relative tilt, the fringes will be localised at infinity, sensibly independent of the degree of shear x_s present. This is consistent with the requirement that the Fourier transform modulated fringe pattern should be observed at infinity. If finite tilt is introduced, then for non zero shear, z becomes finite and dependent upon ψ such that the surface of maximum fringe contrast is no longer a plane.

Hence, to observe and faithfully record the Fourier transform of a spatially extended source as generated by the wavefront folding interferometer, it is necessary to have the two virtual sources parallel and to observe the interference pattern at infinity.

5.3. Alignment of the Interferometer

The interferometer was assembled with the help of a low powered He-Ne laser, which was aligned such that its beam was in a plane parallel to the steel table. The first beam-splitter and the three mirrors of the wavefront folding section of the interferometer were placed one at a time in their approximate positions, ensuring that their reflecting surfaces were normal to the table. The final beam-splitter was then positioned such that it superimposed the two pencils of laser light. Thus at this stage the two arms of the interferometer were parallel but not equal in length.

The path lengths were equalised to within 2mm by placing a microscope objective in the path of the laser light, before the first beam-splitter such that circular interference fringes were formed in the observation plane, centred on the optic axis. Such fringes could only be generated by two spherical wavefronts, with different curvatures, originating from the two virtual images of the focal plane produced by the microscope objective. The straight line passing through these two virtual focal points also passes through the centre of the circular fringes. Thus requiring that the fringes be centred on the optic axis ensures that the microscope objective does not introduce any lateral displacements.

The periodicity of the fringes is inversely related, through the different radii of curvature of the wavefronts, to the separation of the two virtual points. Thus altering the path length of one arm such that the circular fringes appear to collapse into the centre, reduces the path length difference between the two arms. However, this technique could not reduce the path length difference to less than 2mm, because at such small separations, as compared with the interferometer length of 80cms., the size of the first circular fringe was comparable with the observation plane apperture.

At this stage the laser and the microscope objective were replaced by an extended white light source focussed onto approximately

the previous microscope objective focal plane enabling the interferometer to be described as in Figure 5.6. The intensity distribution in the observation plane can be described in a manner similar to that used in section 4.2. Thus the amplitude distribution due to a pair of elementary point radiators situated at $\pm x$ at distances F and $F + \delta F$ from the observation plane is

$$A(u) = A_1 \exp - j k F \left\{ 1 + \frac{(u+x)^2}{2F^2} \right\} + A_1 \exp - j k (F + \delta F) \left\{ 1 - \frac{(u+x)^2}{2(F+\delta F)^2} \right\}$$

This gives an intensity distribution due the entire source $I_s(x)$ as

$$I_s(u) = \int_{\text{source}} I_s(x) \left(1 + \cos k \left\{ \frac{2ux}{F} - \delta F \right\} \right) dx$$

which on assuming a simple rectangular function for the source spatial distribution gives

$$I(u) = I_1 + I_1 \sin \frac{kux}{F} \cos k \left\{ \frac{(s+x)u}{F} - \delta F \right\}$$

where s is the relative shear of the two virtual sources.

It can be clearly seen that whereas the transform is centred on the optic axis, $u = o$, the carrier fringes are centred on

$$u = \frac{F \delta F}{(s + x)}$$

which for typically

$$F = 10^3 \text{ mm}, \quad \delta F = 2 \text{ mm}, \quad (s + x) = 10 \text{ mm}$$

$$u = 200 \text{ mm}$$

However, as shown in section 4.1, for a broad band source interference fringes, for the typical parameters stated above, will only exist over several mms. Thus if at the end of the laser assisted alignment, δF remains finite, but of known sign, then on replacing the laser with a white source, the path length adjuster can be moved so as to reduce δF. Clearly, as δF approaches zero, the small band of fringes moves towards the optic axis, and as δF passes through zero, the fringes will be centred on the observation plane. It was shown in section 4. 3. , that the zero frequency component of the Fourier transform will always have the largest amplitude, hence the optic axis and thus zero path length difference can be located by noting the position of the brightest fringe. By this technique it was possible to equalise the two arms to within a fraction of a wavelength.

Having achieved this situation it was relatively easy to insert the two diffraction gratings and rotate the entire input table such that the principle ray leaving the second grating was parallel to the interferometer optic axis.

CHAPTER 6

Interference Pattern Produced by the Interferometer

The experiments performed on the interferometer described in the previous chapter, were designed to ascertain whether the number of fringes can be increased and whether the fringe visibility is related to the Fourier transform of the source. Hence, initial experiments were performed on the interferometer without any spectral compensation to obtain an estimate of the interferometer's performance, against which the results of subsequent experiments with spectral compensation could be compared. Finally, several attempts were made to retransform the interference pattern in order to reproduce the original input object.

6.1. Interferometer Without Gratings

The interferometer used throughout this set of experiments was arranged as shown in Figure 6.1., and had an effective object plane to image plane distance of 1 m.

In the first experiment, the source used was a high pressure mercury discharge lamp, rated at 250 watts with an arc length of 10mm. The arc was focused onto a 25mm wide and 7mm high slit, which constituted the input object. The lamp, lens and slit were all mounted on the common input table, as shown in Figure 6.1. The resulting interference pattern in the observation plane consisted of constant frequency fringes whose contrast was modulated by a simple periodic function. A typical interference pattern is shown in Figure 6.2.

In Chapter 3 it was shown that the intensity distribution in the observation plane of a spectrally uncompensated wavefront folding interferometer would be given by

$$I(u) = \text{const} + \iint_{-\infty}^{\infty} I(x,k) \cos \frac{2kux}{F} \, dx\,dk$$

As suggested in Chapter 3, the spatial and spectral transforms can be considered separately, thus the contribution from the spatial transform is

$$I_1(u) = \text{const} + I_o \int_s^{s+r} \cos \frac{rkux}{F} \, dx$$

where s is the distance of the slit in the virtual object plane to the interferometer axis, and r is the slit width

$$I_1(u) = \text{const} \left\{ 1 + \text{sinc} \frac{kur}{F} \cos \frac{ku(2s+r)}{F} \right\} \qquad 6.1.$$

The spectral transform contribution is given by

$$I_2(u) = \text{const} + \int_{k_1}^{k_2} I(k) \cos \frac{2kux}{F} \, dk$$

In the first instance the spectral composition of the mercury lamp can be approximated to a line at .546 μm and another of comparable strength at .578 μm. This accounts for 75% of the visible output, but ignores the fact that the lines will be broad. Within these approximations the spectral transform becomes

$$I_2(u) = \text{const} \left\{ 1 + \cos \frac{ux}{F} (k_1 - k_2) \cos \frac{ux}{F} (k_1 + k_2) \right\} \qquad 6.2.$$

On combining equations 6.1 and 6.2 gives, to a first order approximation, the intensity distribution as

$$I(u) = \text{const} \left\{ 1 + \text{sinc} \frac{k'ur}{F} \cos \frac{ux'}{F} (k_1 - k_2) \cos \frac{2k'ux'}{F} \right\}$$

where $k' = 4\Pi/(\lambda_1 + \lambda_2)$

$x' = s + r/2$

It can be seen from the above expression that while the spatial transform is independent of object position, the spectral transforms and fringe frequency are directly proportional to the object position relative to the interferometer axis. The extent of the central lobe of the spatial transform is given by

$$u = \frac{2F(\lambda_1 + \lambda_2)}{4r}$$

which for $r = 25.0\,\mu m$ gives

$$u = 22.48 \text{ mm}$$

However, in order to quantitatively observe the dependence of the spatial transform on object position, the observation plane had to be viewed through a cathetometer telescope with a graticule in the eye-piece focal plane. As the entrance pupil of this telescope was approximately 15 mm, the spatial transform can be safely ignored when measuring the spectral transform.

When the object position was varied, by moving the entire input table, the period of the modulating spectral transform was measured. The results are shown in Figure 6. 3., and as expected show an inverse relationship between the object shear and spectral transform period. The individual departures from the smooth curve result from the subjective estimates of nulls in the modulating function.

Although the pattern observed and shown in Figure 6 2., is reasonably accurately described by the preceeding discussion and associated assumptions, there remain two unanswered questions. Firstly, it will be noticed from Figure 6. 2., that as the modulating function approaches a null another higher frequency modulating function appears, and secondly the question of how many fringes have been formed.

In evaluating the spectral transform it was assumed that the spectrum could be approximated by a line at . 546 μm plus a combination of the two yellow lines at

. 577 μm and . 579μm to a line at . 578 μm of a strangth equal to the green. In fact the visible spectrum has 3 more lines (Harrison, Lord and Loafburrow 1957); a pair in the red at . 623 μm and . 615 μm of strength relative to the green of 0.1 and 0.2 respectively, and a blue line at . 435 μm of relative strength . 2. Thus assuming the red to be a single line at . 618 μm of relative strength 0.3 , equation 6.2. , describing the spectral transform becomes

$$I(u) = \text{const} \left\{ 1 + \left\{ 0.25 \cos \frac{u x'}{F} (k_3 - k_4) + \cos \frac{u x'}{F} (k_1 - k_2) \right\} \cdot \cos \frac{u x}{F} (k_1 + k_2) \right\}$$

where k_3 and k_4 represent the red and blue lines.

The first point to note is that the magnitude of the red and blue cosine is small and hence will only be noticeable when the main modulating cosine approaches a null. Secondly, the ratio of the periodicities of the two modulating functions is

$$R = \frac{P_{GY}}{P_{RB}} = \frac{k_3 - k_4}{k_1 - k_2} = 5.7$$

Comparing the number of fringes, as shown in Figure 6.2. , in the two periodicities gives a ratio of 5.2 . Although not in very close agreement, the ratios are sufficiently similar to accept the red and blue lines as the source of the higher frequency, lower amplitude, modulating function.

The second question, regarding the total number of fringes formed was not so easy to answer. The theoretical limit on the number of fringes is determined by the source spectrum. Thus if the spectrum is defined as a set of lines, each broadened, or convolved with a common broadening function, then the spectral transform will be the product of the cosine term due to the discrete lines, with the transform of the broadening function. A probable broadening function would have a Guassian distribution, hence its transform would also be Guassian with an inverse relationship between the widths of the two Guassian curves. The broadening function for the lamp used would be relatively narrow,

as the two yellow lines, separated by 20Å could be readily resolved, hence its transform was necessarily wide. Thus with a slowly varying transform and no zeros it was difficult to sensibly observe and place limits on the range over which fringes existed.

In order to have a wider spectrum and hence fewer fringes, and to give greater flexibility in the choice of spatial distributions, the mercury arc was replaced by the tungsten halogen projection lamp whose characteristics have been presented in Figure 5.2.

The object size was kept at 25 μm such that the interference pattern would be dominated by the transform of the lamp's spectrum. From Figure 5.2., this spectrum can be approximately described by

$$I(k) = a_0 + \frac{2\Pi a_1}{k}$$

where $a_0 = 0.1$

$a_1 = 2.67$ per μm

Thus the interference pattern is given by

$$I(u) = \text{const} + \int_{2\pi/0.7}^{2\pi/0.4} I(k) \cos \frac{2kux}{F} \, dk$$

The above integral does not readily lend itself to be expressed in a simple analytical form, hence it has been evaluated numerically and the resulting fringe pattern for a visibility greater than 10% is shown in Figure 6.4a. From this figure it can be seen that only 3 well defined fringes will be formed and that this interference pattern will be symmetrical about the optic axis.

However, taking into account the dispersive beam splitters, the interference pattern in the observation plane will be given by

$$I(u) = \text{const} + \int_{2\pi/0.7}^{2\pi/0.4} I(k) \cos \frac{2kx}{F} (u + \delta p) \, dk$$

where $\delta p = 571.4 \cdot 2\Pi \left(\frac{1}{k} - \frac{1}{k_1} \right)$

The above equation has also been evaluated by numerical techniques and the resulting fringe pattern is shown in Figure 6.4b. Although the three dominant fringes will exist, they now appear off centre and are not of a symmetrical disposition. It should also be noted that there is an increase in the number of low visibility fringes.

The interference pattern actually observed in the output plane produced by the tungsten halogen lamp, focussed onto a 25 μm hole is shown in Figure 6.5. The first interesting observation is that there are many more fringes than predicted from Figure 6.4b., however, this might be explained by a non-linear recording process enhancing some of the lower visibility fringes. This explanation is substantiated by the fact that only 3 black and white fringes were observed while the remainder were coloured.

These extra fringes were predominantly green on one side of the central black and white fringes, and red on the other side. Furthermore, there were approximately twice as many red fringes visible than green ones. The exact reason for this is not clear.

6.2. Interferometer with Gratings

The diffraction gratings were added to the interferometer as shown in Figure 6.6.

For the first experiment, the source was the same mercury lamp used in the previous section, focussed onto the same 25 μm slit. It was shown in Chapter 5 that due to the construction of the mount for the gratings, the virtual objects had to be sheared from the axis by an amount equal to the offset of the real object from the imaginary junction of the two gratings. It was further shown that a departure in the virtual plane from this exact shear would greatly reduce the number of fringes. Hence a simple, but powerful way in which to observe the performance of the spectrum compensating system as the wavelength dependent displacements approached ideal values, would be to vary the offset of the virtual objects from the interferometer axis.

Thus with the angle between the gratings set to 47.0 degrees, the virtual objects were gradually pulled apart from an initial position of zero shift. Initially, the fringe pattern observed was of the same format as in the previous section, with a dominant low frequency term

modulating higher frequency carrier fringes. However, the number of fringes observed between a pair of nulls was only 16, as shown in Figure 6.7a., compared with 26 in the previous section. As the virtual objects were moved away from the axis the periodicity of the fringes decreased, but contrary to the same experiment without the gratings, the periodicity of the modulating function increased as in Figure 6.7b. The increase of the modulating function periodicity slowed down and stopped while the fringe periodicity was decreasing, thus the number of fringes in a modulating function period was increasing, up to a maximum, as shown in Figure 6.7c. The corresponding optimum shear was 1.0mm. Finally, the modulating function periodicity began to decrease at a rate such that the number of fringes within a period decreased.

The rings visible in the centre of the photographs in Figure 6.7 are Newton's rings formed by the composite lens structure of the telescope objective. The two broad fringes on the left side of Figure 6.7b., are caused by an unwanted diffraction order passing through a spatial filter. A more detailed description of the observed dependence of the number of fringes on the virtual object separation is shown by the solid line in the graph of Figure 6.8. The theoretically expected number of fringes is shown in Figure 6.8 by the interrupted line.

The differences between the observed and expected numbers of fringes can be explained firstly by noting that the theoretical curve was evaluated on the assumption that the spectrum consists of a line of .546 μm and another at .578 μm. In the previous section it was shown that the region around a null is somewhat confused by the presence of the other lines in the spectrum, however, the fringes were counted between two successive nulls and no estimate was made of the exact position of the null of the dominant modulating function. Secondly, the extra dispersions introduced by the beam splitters will distort the dispersions produced by the gratings and thus reduce the number of fringes.

On the experimental evidence so far it is possible to say that the interferometer with the gratings behaves, to a first order approximation at least, as expected for a polychromatic point source. To extend the system's operation to larger sources, it is necessary to model such sources as the sum of several point sources. Thus if no interferometer parameter

is varied, positioning a point source at different locations in the input plane will give an indication of the performance of a continuous source sampled at these points. Finally, if there is no marked difference in the interference patterns produced by the various points, then it may be concluded that a continuous source will not limit the fringe forming potential of the interferometer.

Thus the 25 μ m slit was repositioned in 5 locations at .5mm intervals. The resulting number of fringes in a fundamental modulating period, as a function of virtual plane shear are shown, together with the previous scan in Figure 6.9.

It can be clearly seen that the number of fringes, produced by the point sources in different lateral positions, peaks at intervals of 0.5mm. Thus each of the five point sources has its spectrum correctly, or at least equally, compensated. Therefore it can be concluded that if the space covered by the point sources were occupied by a continuous source, then every point in that source would be spectrally compensated.

In the next set of experiments, the mercury arc lamp was replaced by the tungsten halogen lamp, with its filament focussed onto the variable slit. Due to the weaker nature of this source, in terms of lumens per source unit area, the smallest slit that could be accommodated consistent with having a detectable amount of light in the observation plane, was 0.11mm. This object size will produce a 5.5 mm central lobe of the sinc function in the transform, and hence, cannot be considered as a point source for a direct comparison with previous experiments. However, with the number of fringes maximised, the interference pattern should represent the Fourier transform of the object spatial distribution.

Thus the interference pattern was observed for different object slit widths and photographed by placing the photographic emulsion directly in the observation plane. The photographs, together with the object slit sizes are shown in Figure 6.10. The expected position of the nulls in the transform is given by equation 6.1., and has also been shown in the figure. The observed transform appears to be somewhat smaller than expected although it is of the correct shape. This would imply a degree of magnification of the object within the interferometer, however, this occurrence will be discussed later when accurate measurements can be made in the transform plane.

The depth of modulation of the carrier fringes for this particular object is determined by a sinc function namely

$$\frac{\sin\left(\frac{k\,u\,x}{F}\right)}{\frac{k\,u\,x}{F}}$$

In the photograph, although the central lobe can be seen very clearly only the first and sometimes the second lobes can be recognised. Therefore, assuming that the fringes do exist, it can be concluded that the magnitude of the third lobe in the modulating function is too small to be readily distinguished from the background noise. The magnitude of the third lobe of the sinc function is $9.1 \quad 10^{-2}$ which when compared with the central zero frequency term of unity implies that spatial frequencies of magnitudes less than 9% of the zero frequency will be lost.

However, this conclusion is based on simple direct observations and does not take into consideration any form of noise reduction, either optical of electronic. Clearly any such image enhancement technique, applied to the interference pattern will reduce the minimum necessary object contrast.

Finally, in Figure 6.10c a band of fringes is visible on either side of the centre, at a distance of 1.2cm. These bands are due to the fine structure of the filament, which appeared in the object aperture. The filament consisted of a wire 300 μm thick, looped round with 300 μm gaps in between successive loops. Thus the input object was a narrow rectangular function, the wire thickness convolved with a comb function, the wire spacing, all multiplied by a large rectangular function representing the object aperture.

The transform of such an input object would be a narrow sinc function, the transform of the object aperture, convolved with a comb function all multiplied by a wide sinc function being the transform of the wire thickness. For the measured filament sizes the comb function in the transform plane should have a spacing of 1.83mm, and the sinc function representing the wire thickness should have its first zero at 0.91 mm. These figures are in reasonable agreement with the measured distances in Figure 6.10, although again the observed transform appears to be smaller than expected.

In order to obtain an estimate of the number of fringes formed by this interferometer, it is necessary to produce a transform with a large amplitude at high spatial frequencies, and still have enough light in the observation plane to see the fringes. Therefore, the next input function, in an attempt to fulfill the above requirements was a computer generated binary grating of 10 lines per mm with a 6 to 1 mark to space ratio. This object grating was mounted as close as possible to the variable slit such that the overall width of the object could still be controlled. Hence, the input object can be considered as the convolution of a comb function defining the grating periodicity, with a rectangular function, defining the grating profile, all multiplied by the effective aperture. The expected interference pattern is shown schematically in Figure 6.11.

The interference pattern produced was recorded by placing the photographic emulsion in the observation plane, and a typical record is shown in Figure 6.12. The most important point to note is that eleven bands of fringes are clearly visible. These bands correspond to the zero frequency and the first five, positive and negative frequencies present in the grating. As the contrast of the carrier fringes in the absence of any Fourier transforms is constant, or at least slowly varying, it is safe to assume that if fringes exist at both extremities of a transform, then they will exist across the region bounded by the above extremities. The real distance in the observation plane between the positive and negative fifth orders is 33.5 mm., while the fringe periodicity in the observation plane is 89.3μm. Thus the number of fringes produced in the observation plane is

$$\frac{33.5}{.0893} = 375$$

which is convincingly in excess of the 3 fringes produced in section 6.1.

The width of these eleven bands is determined by the transform of the overall extent of the object, while their amplitudes are determined by the transforms of the individual grating slit size. Thus for an object size of 1mm the width of each band should be

$$w = \frac{\lambda F}{x} = \frac{.55 \quad 10^{6}}{10^{3}} = 0.55\text{mm}$$

while the visibility of the n^{th} band relative to the central band is given by

$$V = \frac{\sin N \frac{k u x}{F}}{\frac{N k u x}{F}}$$

which for successive bands becomes 95%, 83%, 63%, 41% and 19% It can be seen from Figure 6.12 that the values recorded correspond with the above calculated values, although the visibility estimate is necessarily subjective.

Finally, the separation of the bands is determined by the transform of the grating periodicity. Thus a grating of 10 lines per mm implies a band separation of 5.5mm . However, the measured band separation is sensibly constant at 3.35 mm. As this difference is a constant multiplicative factor some form of magnification can be assumed responsible. Such a magnification could result from the non-linear manner in which points across the input object are treated by the pair of gratings. Thus if a plane wavefront of width x is incident on the first grating, as shown in Figure 6.13 , then it will leave the second grating with a width w . From Figure 6.13.

$$A C = \frac{x \cos \alpha}{\cos (\alpha + \beta)}$$

$$w = A C \cos \theta$$

$$\frac{w}{x} = \frac{\cos \alpha \cos \theta}{\cos (\alpha + \beta)}$$

On substituting current values for the angles, the ratio w/x becomes equal to 1.48 , which corresponds closely to the ratio of the observed and expected band separations.

6.3. Object Reconstruction

It has been shown in the previous section, that it is possible to obtain a wide band Fourier transform of an extended polychromatic object.

Hence it is possible to consider the construction of a real time, two dimensional signal processor. Clearly the front part of such a processor would be the wavefront folding interferometer, producing the Fourier transform of the input object. Subsequently the processing of this transform may be performed either optically, electronically or a combination of the two.

However, retransformation of the filtered object from the processed transform is made difficult by the bias level in the transform plane. This can be appreciated by considering a straightforward system of an interferometer producing the transform which is recorded and presented to the input plane of another interferometer. Thus if the input object is a sinusoid of frequency ν and of finite extent r

$$I(x) = 1 + \cos \nu x \quad \text{for } |x| < r$$

illuminated with unit intensity per length. Then the interference pattern in the transform plane will be

$$I(u) = \int_{s}^{s+r} (1 + \cos \nu x)\left(1 + \cos \frac{2 k u x}{F}\right) dx$$

where $s+r/2$ is the lateral shear introduced by the interferometer. Performing the integration and assuming that $r > 1/\nu$

$$\begin{aligned} I(u) = {} & r + r \operatorname{sinc} \frac{k r u}{F} \cos \frac{k u (2 s + r)}{F} \\ & + r/2 \operatorname{sinc} r/2 \left(\frac{2 k u}{F} - \nu\right) \cos \left\{ \frac{k u (2 s + r)}{F} - \phi \right\} \\ & + r/2 \operatorname{sinc} r/2 \left(\frac{2 k u}{F} + \nu\right) \cos \left\{ \frac{k u (2 s + r)}{F} + \phi \right\} \end{aligned} \qquad 6.3.$$

where $\phi = \nu(2 s + r)/2$ and represents a phase shift in the fringes determined by the position of the input frequency relative to the interferometer axis. For the sake of simplicity let $\phi = 0$.

If this interference pattern is recorded on a film of width w such that

$$w = \frac{F \nu}{k}$$

and then presented as the input to the interferometer, the original object will be reconstructed. Thus the intensity distribution in the input plane u' is

$$I(u') = r + \cos \frac{k u' (2 s + r)}{F} \left\{ r \operatorname{sinc} \frac{k r u'}{F} + r/2 \operatorname{sinc} \tfrac{1}{2} \left(\frac{2 k u'}{F} - \nu \right) + r/2 \operatorname{sinc} r/2 \left(\frac{2 k u'}{F} + \nu \right) \right\}$$

where $u' = u + w/2 + v$

and v represents lateral shear introduced by interferometer. Hence the intensity distribution in the output plane becomes

$$I(x') = \int_{v}^{v+s} I(u') \left(1 + \cos \frac{2 k u' x'}{F}\right) d u'$$

The above integral may be simplified by replacing the sinc functions with rect. functions of width w where

$$w' = \frac{\lambda F}{r}$$

Thus the output intensity pattern is

$$\begin{aligned} I(x') = {} & r w + r w \operatorname{sinc} \frac{k x' w}{F} \cos \frac{k x'}{F} (2 v + w) \\ & + r w' \operatorname{sinc} \frac{k w'}{2F} (2 x' + 2 s + r) \cos \left\{ \frac{k x'}{F} (2 v + w) + \phi' \right\} \\ & + r w' \operatorname{sinc} \frac{k w'}{2F} (2 x' - 2 s - r) \cos \left\{ \frac{k x'}{F} (2 v + w) - \phi' \right\} \\ & + r w' \operatorname{sinc} \frac{k w'}{2F} (2 x' + 2 s + r) \cos \left\{ \frac{k w x'}{F} + \phi' \right\} \cos \left\{ \frac{k x'}{F} (2 v + w) + \phi' \right\} \\ & + r w' \operatorname{sinc} \frac{k w'}{2F} (2 x' - 2 s - r) \cos \left\{ \frac{k w x'}{F} - \phi' \right\} \cos \left\{ \frac{k x'}{F} (2 v + w) - \phi' \right\} \end{aligned}$$

6.4.

The first term represents the overall bias offset, while the second is the transform of the bias level in the input plane. The next pair of terms consist of sinc functions positioned at $x' = \pm s + r/2$ and of width $F\lambda/w'$ which on substituting for w' gives a width of r. Thus these terms represent the bias in the original input object, sheared from the axis by an identical amount, although now the object is reconstructed as positive and negative spatial frequencies. This reconstructed bias level appears as an implitude modulation of carrier fringes due to the offset of the zero frequency term in the input from the interferometer axis.

Finally, the last two terms consist of similar sinc functions as in the previous terms, however, there is an added modulating cosine term with a frequency $w/\lambda F$ which on substituting for w gives a frequency of ν. Thus the original object can be perfectly reconstructed. However, the magnitude of this reconstructed object is rw' which is superimposed on a bias level rw. Hence the object to bias ratio is w'/w which becomes $F\lambda/rw$.

Typically, $F\lambda = 5.10^{5}\ \mu m$, while $r = 10^{4}\ \mu m$ and $w = 2.5\ 10^{4} m$, hence

$$\frac{F\lambda}{rw} = 2.10^{-3}$$

Thus the reconstructed object will be submerged in the bias level.

There are two possible methods of overcoming this problem. The first involves the removal of the bias level by electronic filtering, and clearly is only applicable when the transform has been recorded by a television camera. The second technique would be to perform the retransformation optically in a coherent light Fourier transform bench, thus concentrating the bias level into a single point at the centre of the output plane.

In order to verify experimentally, that the interference patterns produced in the previous section were Fourier transforms, the photographic records of the fringes presented in the input plane of a coherent light Fourier transform bench. However, this technique gives rise to an artefact which can be understood on considering the coherent light retransformation of the Fourier transform in the previous example. The input to the coherent processor becomes

$$A(u') = r + r\cos\frac{ku'}{F}(2s + r)\left\{\operatorname{sinc}\frac{kru'}{F} + \tfrac{1}{2}\operatorname{sinc}r/2\left(\frac{2ku'}{F} - \nu\right) + \tfrac{1}{2}\operatorname{sinc}r/2\left(\frac{2ku'}{F} + \nu\right)\right\}$$

Hence the amplitude distribution in the output plane becomes

$$A(x') = \int_{-w/2}^{w/2} A(u') \exp j \frac{k u' x'}{F} \, du'$$

As in the preceeding example the calculation can be simplified by approximating the sinc function in $A(u')$ by a rectangular functions. Hence the intensity in the output plane can be shown to be

$$I(x') = r(w+w') \operatorname{sinc} \frac{k x'}{2F} (w+w')$$

$$+ \frac{r w'}{2} \operatorname{sinc}^2 \frac{k w'}{2F} (2s + r + x') + \frac{r w'}{2} \operatorname{sinc}^2 \frac{k w'}{2F} (2s + r - x')$$

$$+ \frac{r w'}{2} \operatorname{sinc}^2 \frac{k w'}{2F} (2s + r + x') \cos \frac{k w}{F} (2s + r + x')$$

$$+ \frac{r w'}{2} \operatorname{sinc}^2 \frac{k w'}{2F} (2s + r - x') \cos \frac{k w}{F} (2s + r - x')$$

On comparing the above equation with equation 6.4., it can be seen that the overall bias level is contained within a small region about the centre of the output plane. The second point to note is that the object sinc functions, although of the correct width, are positioned at twice the original shear, i.e., at $x' = \pm 2s + r$. This double shear arises because the carrier fringes in the transparency representing the original shear appear as amplitude fringes in the coherent processor and not as the intensity fringes that were initially recorded.

In the first instance, the recorded transforms of a 1mm wide input grating sheared by 1.0mm from the axis was positioned in the input plane of a coherent light Fourier transformer, and the observed output intensity distribution is shown in Figure 6.14. Clearly, twenty-five to thirty vertical bands within a distance of 3mm can be distinguished. The shear from the central spot is approximately 2mm which is, as expected, twice the original shear. The height of the reconstructed images is purely a function of the coherent processor as the transforms produced by the interferometer

was one dimensional and hence all height information was lost. Finally, the star-like nature of the central spot is due to a fourteen element variable aperture stop controlling the extent of the coherent light.

Records of interference patterns produced by wider input object gratings were also reconstructed in a similar manner to that above. Although the reduced light intensity in the larger reconstructed objects made accurate measurements difficult, it became clear that the reconstructed images were smaller than the original input objects. This may result from a non-linear recording of the interference pattern. If the carrier fringes in equation 6.3., were not sinusoidal but more squared, then the input to the coherent processor would be described as the product of the various sinc functions with a comb function at the fringe frequency, all convolved the width of the squared fringe. Thus the output would be as previously described, but now multiplied by the transform of one squared fringe.

If, as suggested, these squared fringes were produced by a non-linear recording process, then their width would be sensibly one half of the ideal sinusiodal fringe period, thus

$$w' = \frac{F}{2\,(2s + r)}$$

The intensity distribution of the transform of such an amplitude function is

$$I(x') = \mathrm{sinc}^2 \frac{k\,x'}{2\,F} \quad \frac{F}{2\,(2s + r)}$$

which gives an effective width of $4s + 2r$. However, the reconstructed images are of width r and sheared from the axis by $2s + r/2$, thus the overall extent of the reconstruction is $4s + 3r$. Hence for r larger than s the extent of the reconstructed objects will be significantly reduced.

CHAPTER 7

Review of the Interferometer

In the opening chapters the intrinsic interest of a wavefront folding interferometer as a real time two dimensional signal processor was unfolded. The following chapters developed the theory and principles of operation of a white light wavefront folding interferometer. Finally, the results of various investigations into the interference pattern produced by the interferometer have been reported. Hence, it is now possible to re-appraise the potential of the spectrally compensated wavefront folding interferometer as a two dimensional real time signal processor.

7.1. Results Achieved

Initially, it was shown how the visibility of the fringe pattern in the absence of any spectral compensation, depended on the spatial and spectral source distributions. When the experiments were repeated with the spectral compensation introduced, it was seen that the fringe visibility was determined solely by the Fourier transform of the input object's spatial extent. This was clearly demonstrated by using a 10 lines per mm binary grating as the input object. In that experiment over four hundred potential fringes were produced using the full spectrum of a tungsten halogen lamp.

Thus it has been demonstrated, albeit in one dimension, that it is possible to produce in real time, an achromatic fringe pattern whose visibility is directly related to the input source's spatial Fourier transform. Hence, this interferometer could be used in any signal processing environment requiring real time presentation of Fourier transform patterns. Typical examples of possible applications might include deconvolutions, especially in processing radar signals or two dimensional feature enhancement or character recognition.

The two dimensional capability of the interferometer could be used to perform some filtering or correlation operation simultaneously on large arrays as for example in automatic speech recognition where a large number of templates must be correlated with the incoming speech waveform.

Finally, the achromatic potential of the interferometer may be used advantageously where monochromatic processing is restrictive. Two immediate examples would be in the various derivatives of the Michelson Stellar Interferometer and secondly in measuring the O. T. F., of optical systems for the complete optical spectrum.

7.2. Practical and Fundamental Difficulties

The largest single difficulty experienced was in aligning the interferometer and ensuring that the path lengths in the two arms were equal. The problems existed mainly because it was difficult to ascertain whether individual components were positioned to the required accuracy until the entire interferometer was built. However, at that stage there were too many degrees of freedom to make any but the smallest improvements. Clearly, this adjustment problem would be much worse for a two dimensional interferometer.

A related problem was that of mechanical stability. The normal vibrations in the building were successfully damped by the massive table and the pneumatic tyre tubes. However, the sensitivity of the individual components to accidental repositioning was very high. This stability problem may be eliminated and the alignment problem reduced to a single occurrence by constructing the interferometer from two glass prisms, in the manner of Breckinridge (1974) and Dainty and Scaddon (1974) . However, such an interferometer, even if made from suitably chosen glasses, may well have different optical path lengths for different wavelengths.

The other set of problems, the very low light levels in the observation plane were the direct result of the interferometer's inefficient use of the source light intensity. The main losses were due to the diffraction gratings and due to the large extent of the interferometer. Clearly, the solid glass interferometer suggested above would reduce the overall size, although it would be as the expense of the input object size. Secondly, more light might be usefully retained by using prisms, instead of gratings, as the dispersive elements. However, these ideas have not been fully developed.

The first theoretical drawback of this interferometer is the restriction of input functions to intensity distributions. Although in the real world this may not prove to be a serious limitation, it is nevertheless a disadvantage when compared with the coherent light Fourier transform processor. A second drawback is the need to record the transform, or the filtered transform, before it can be retransformed by the interferometer, as quite clearly a given point in the transform plane cannot be considered as radiating independently of all the other points. However,

when it is appreciated that many filtering operations can be performed optically and that the interference pattern is an intensity distribution, and as such is readily recorded by electronic techniques, which in themselves have a lot to offer, this disadvantage can become minimal.

7.3. Comparison of the Interferometer With Other Processors

The most common versatile data processor against which any signal processor must be compared is the electronic computer. With the advent of Large Scale Integration techniques, the mini-computer and microprocessor have made computing widely available, resulting in a wealth of algorithms increasing the capabilities of one dimensional digital processors.

A fundamental advantage of electronic processors over their purely optical counterparts, is the ability to amplify a given signal. Clearly, the interferometer, without recourse to any recording process, cannot increase the absolute magnitude of any part of the transform. A second advantage of computers is their versatility and apparent intelligence in making decisions, resulting in the ability to implement a vast range of filtering and processing techniques. However, this advantage is not fundamental but merely the result of intensive research and development during the last three decades.

Against this, the computer is limited to one dimension, hence any two dimensional work is time consuming, whereas the interferometer operates directly on two dimensional signals.

On comparing the interferometer with the coherent light Fourier processor, its fundamental advantage is that it presents the transform as an intensity distribution. Its disadvantages are that it is limited to real input functions and that the transform must be recorded if the processed object is to be re-transformed. However, bearing in mind the comparative ease with which the interferometer's output can be recorded electronically, this disadvantage may well be significantly overcome by the advantages of electronic processing.

Clearly the interferometer by itself, as is true for most optical processors, cannot compete with electronic computers. However, a hybrid system combining the interferometer, with its parallel filtering capabilities, and electronic processing with amplifying power could produce a very powerful signal processor.

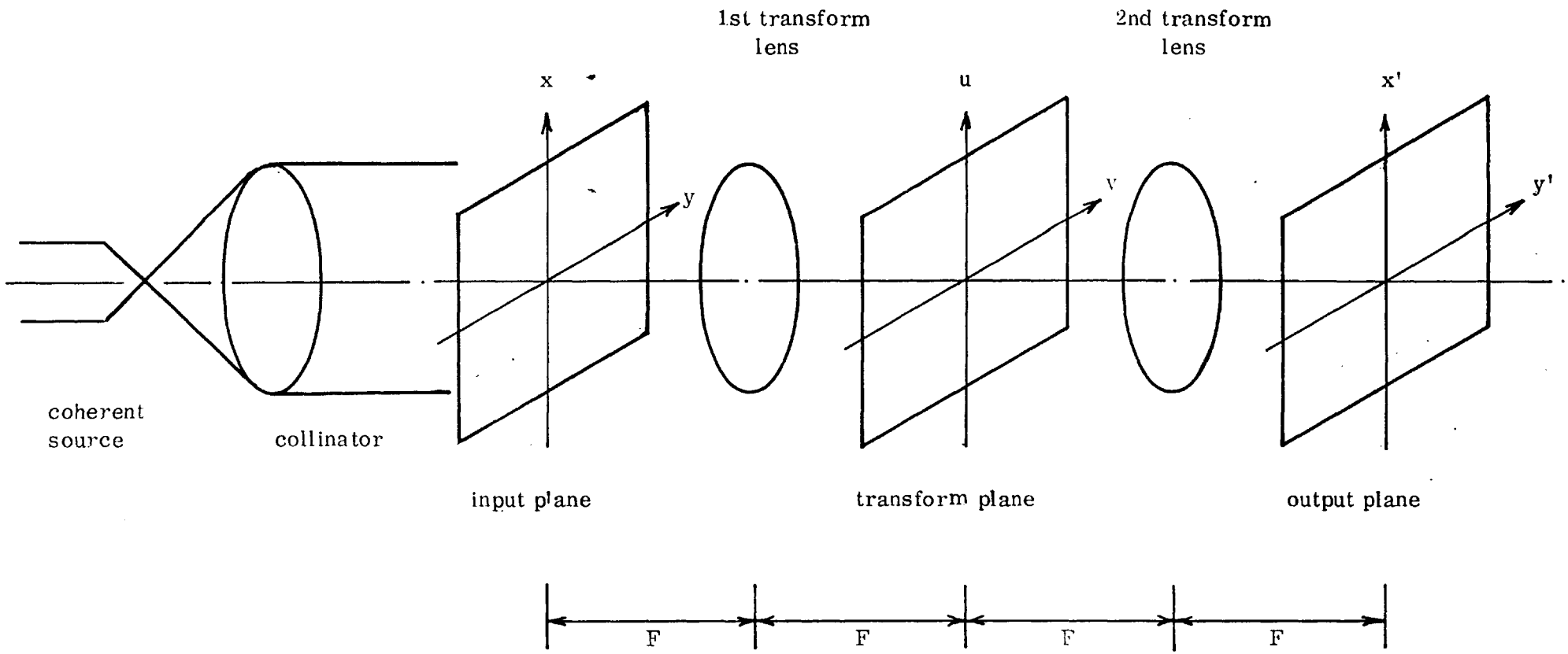


Figure 2.1. Coherent Light Signal Processor

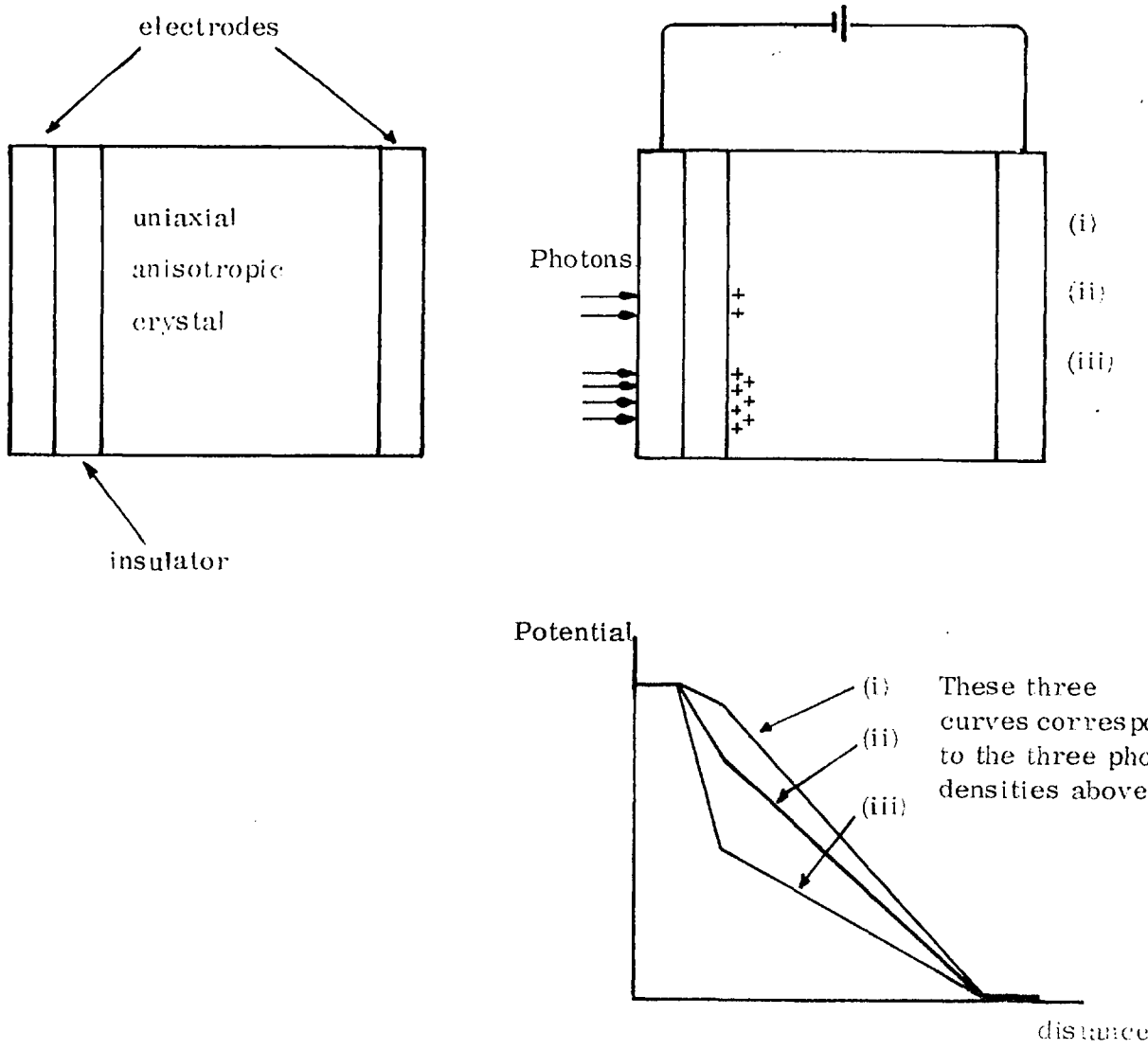


Figure 2.2. Real Time Spatial Modulator

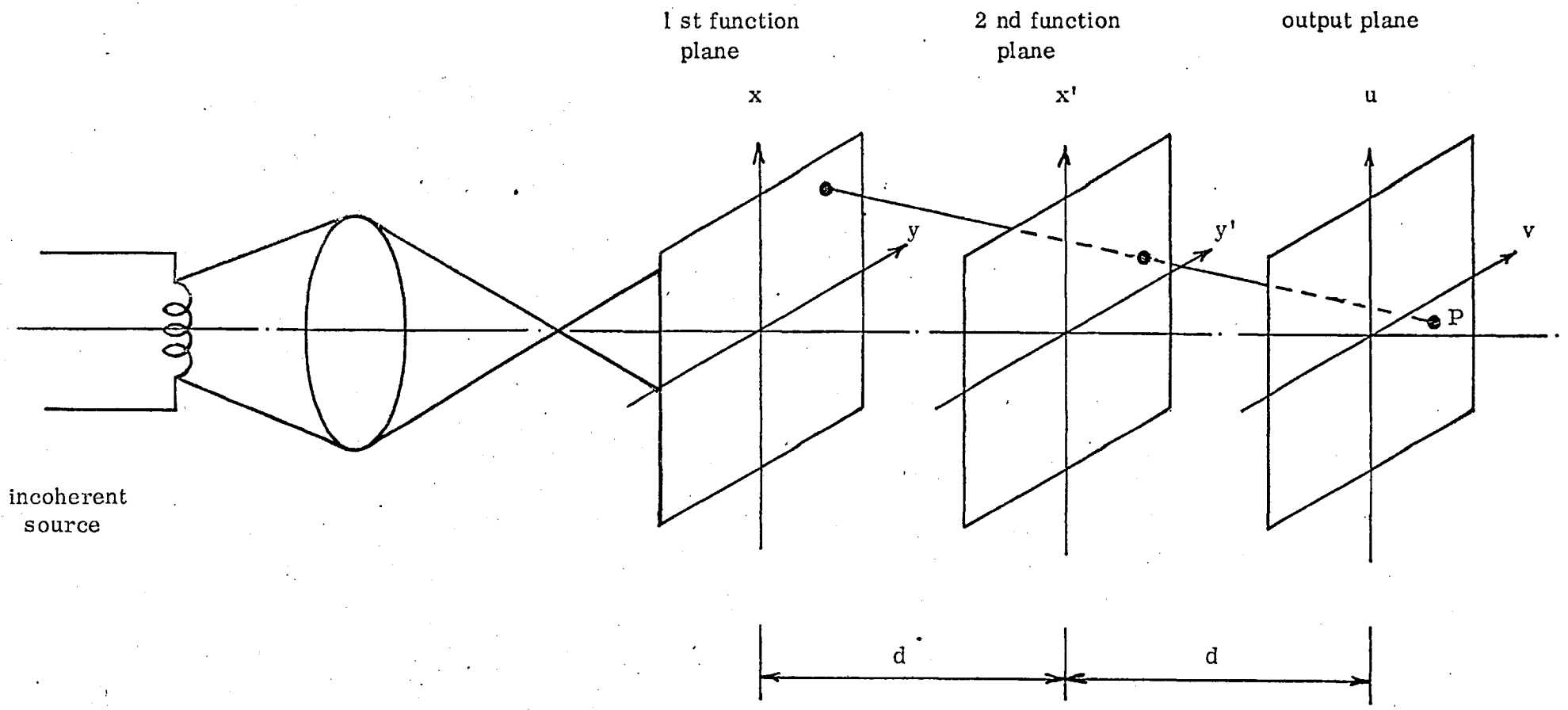


Figure 2. 3. Incoherent Light Signal Correlator

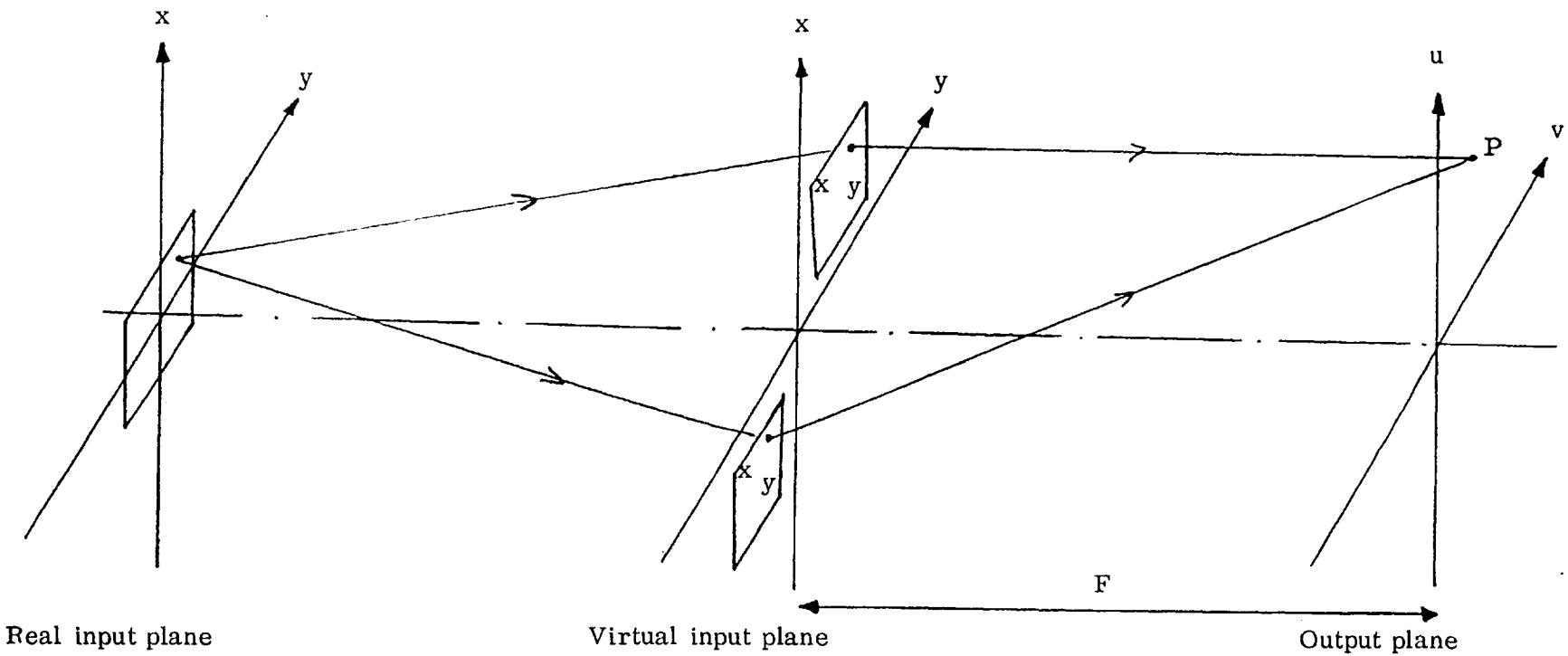


Figure 2.4. Wavefront shearing interferometer.

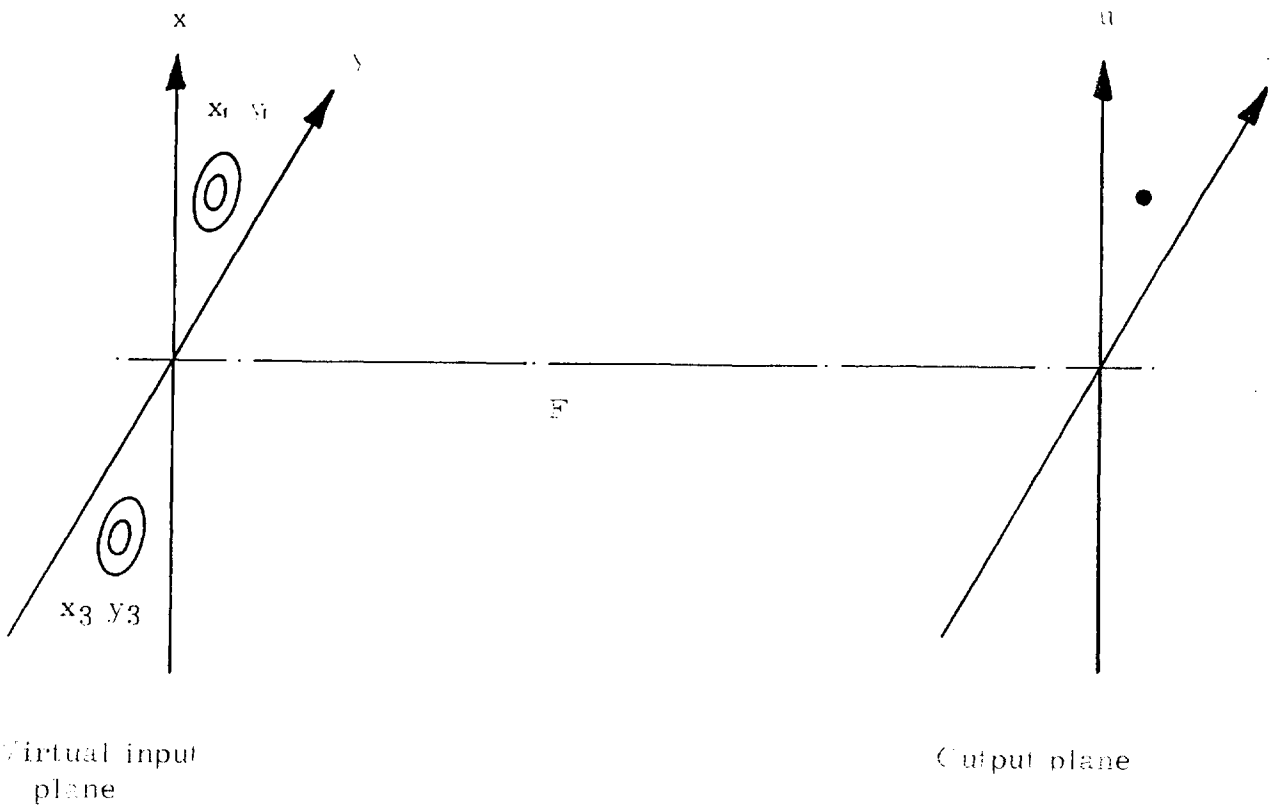


Figure 3.1. Interferometer with spread point input.

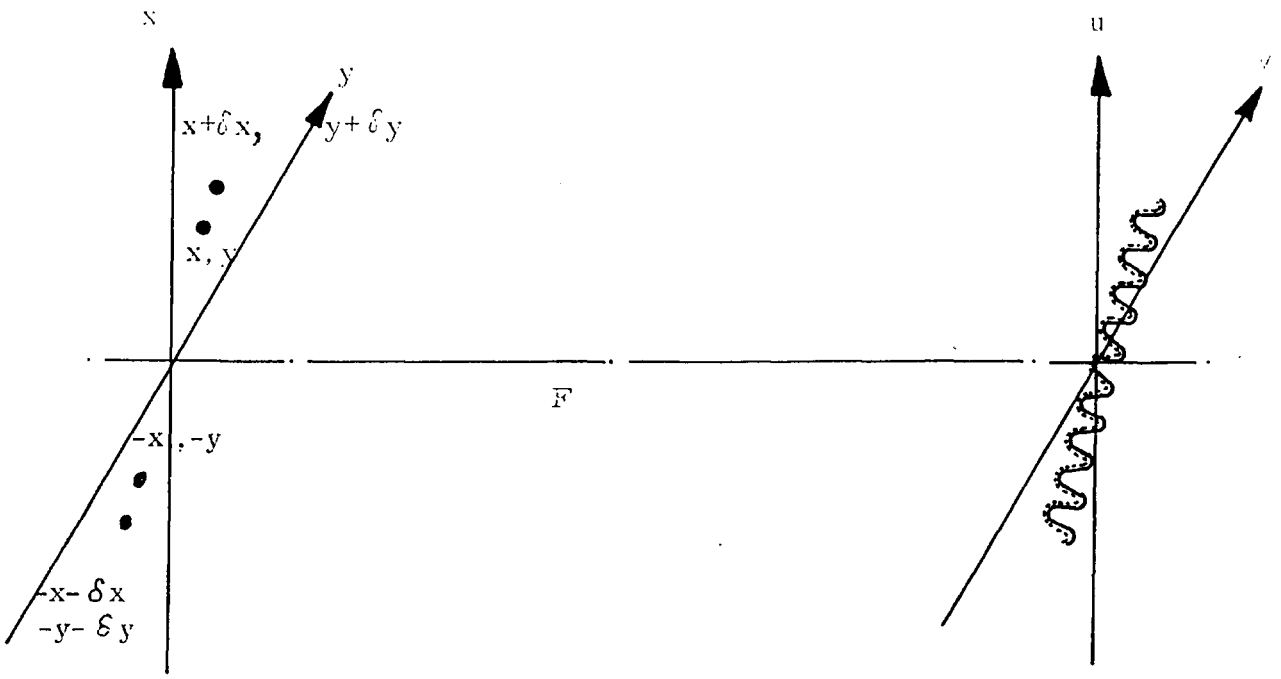


Figure 4.1. Spatial dispersion in the input plane.

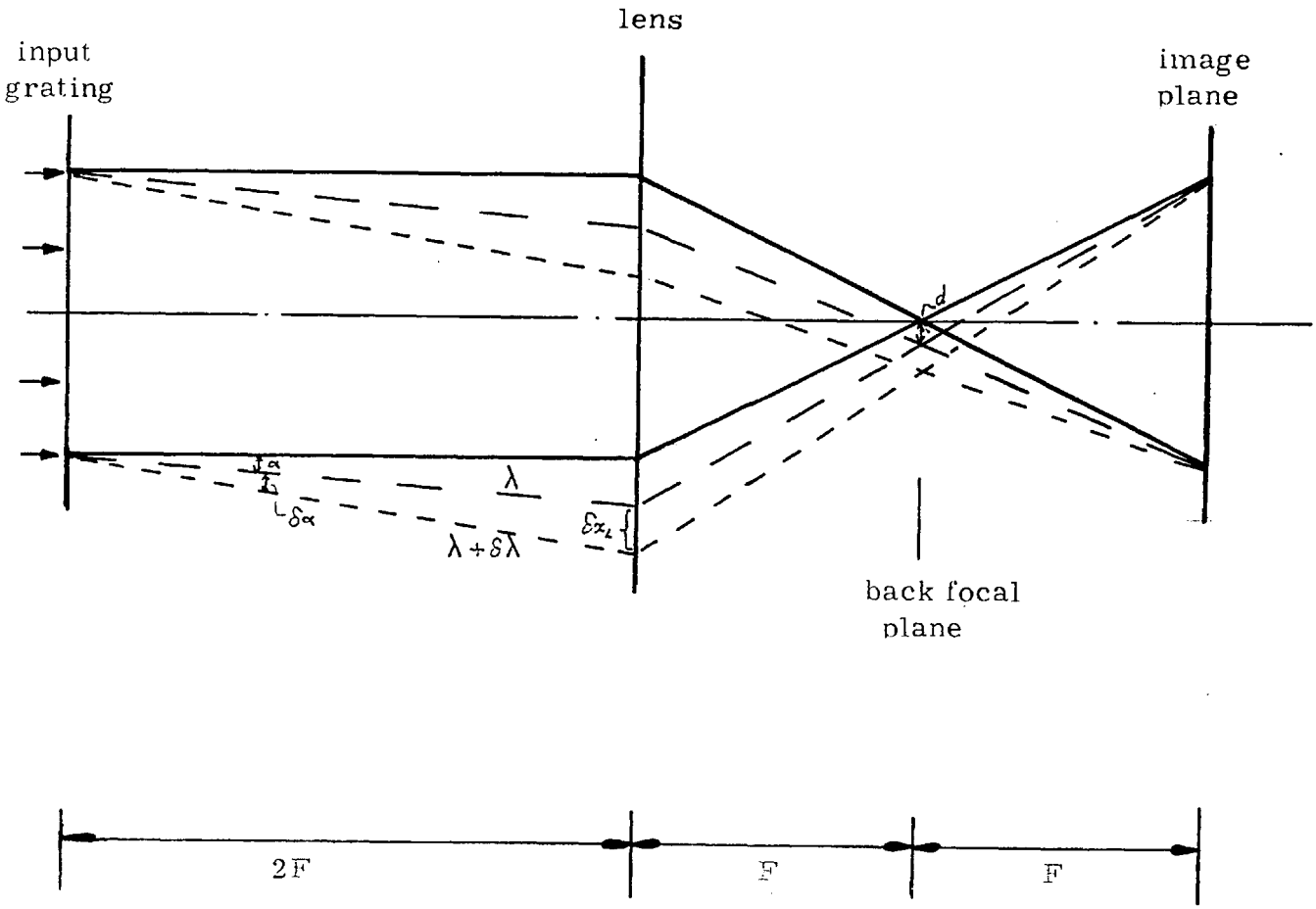


Figure 4. 2. Achromatic Fringes after Leith and Upatnieks

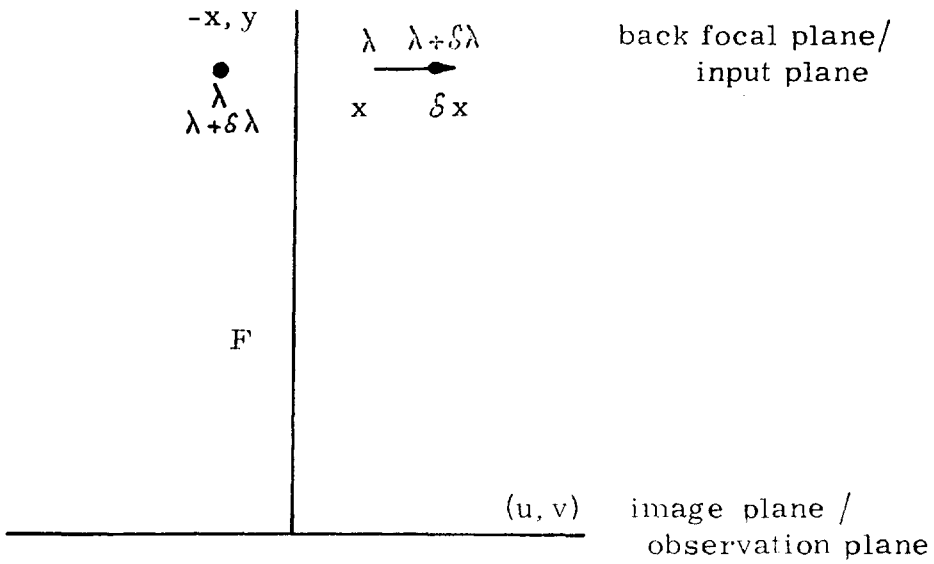


Figure 4.3. Schematic of Leith and Upatnieks system.

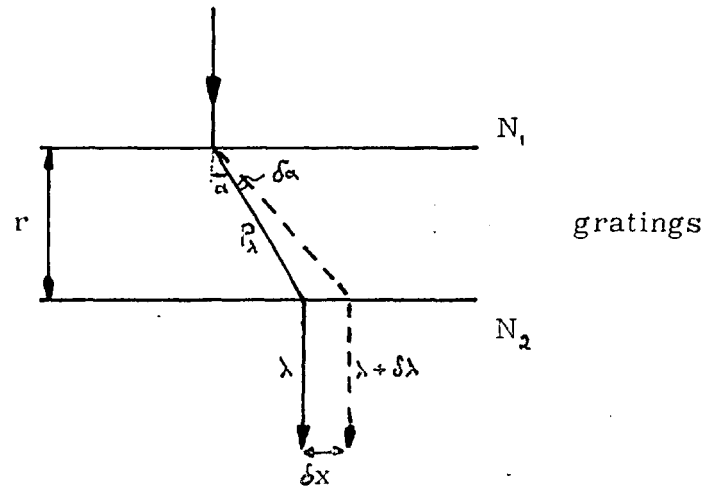


Figure 4.4a. Dispersion by parallel gratings.

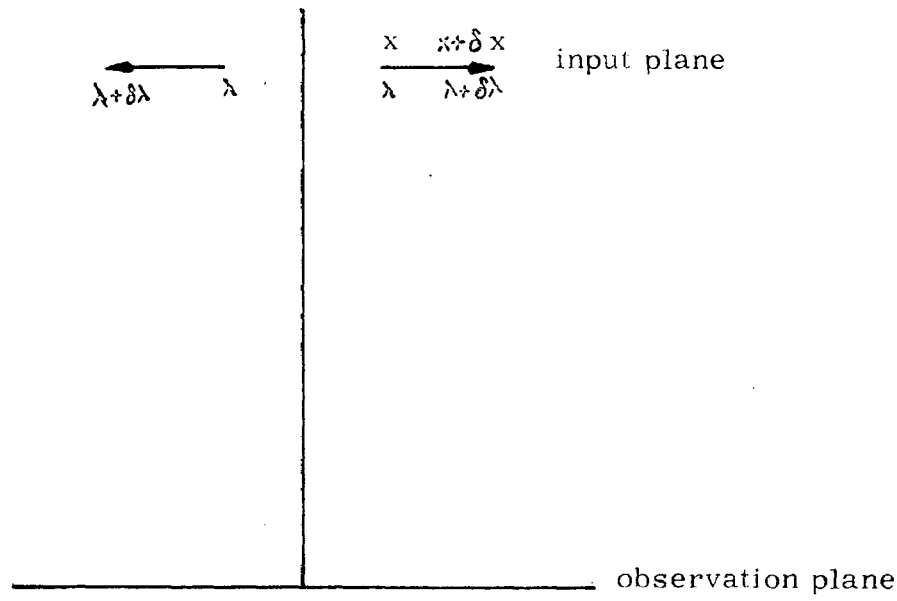


Figure 4.4b. Interferometer Schematic.

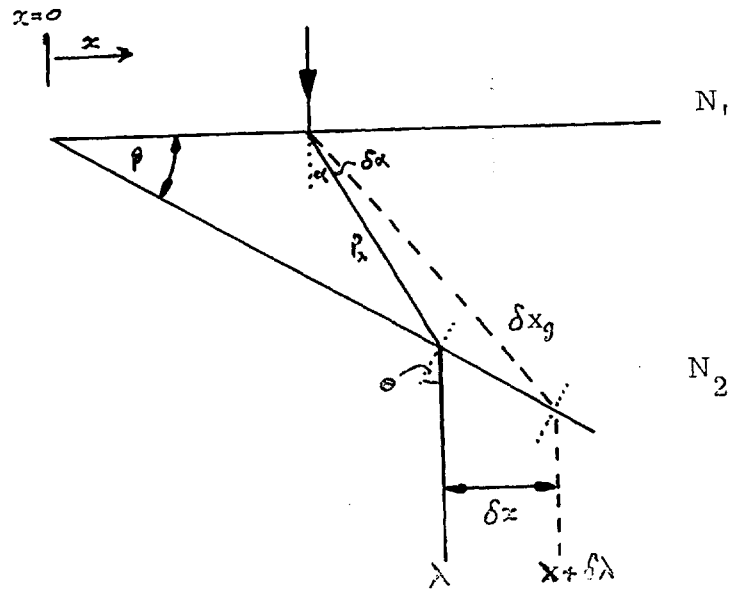


Figure 4.5. Dispersion by inclined gratings.

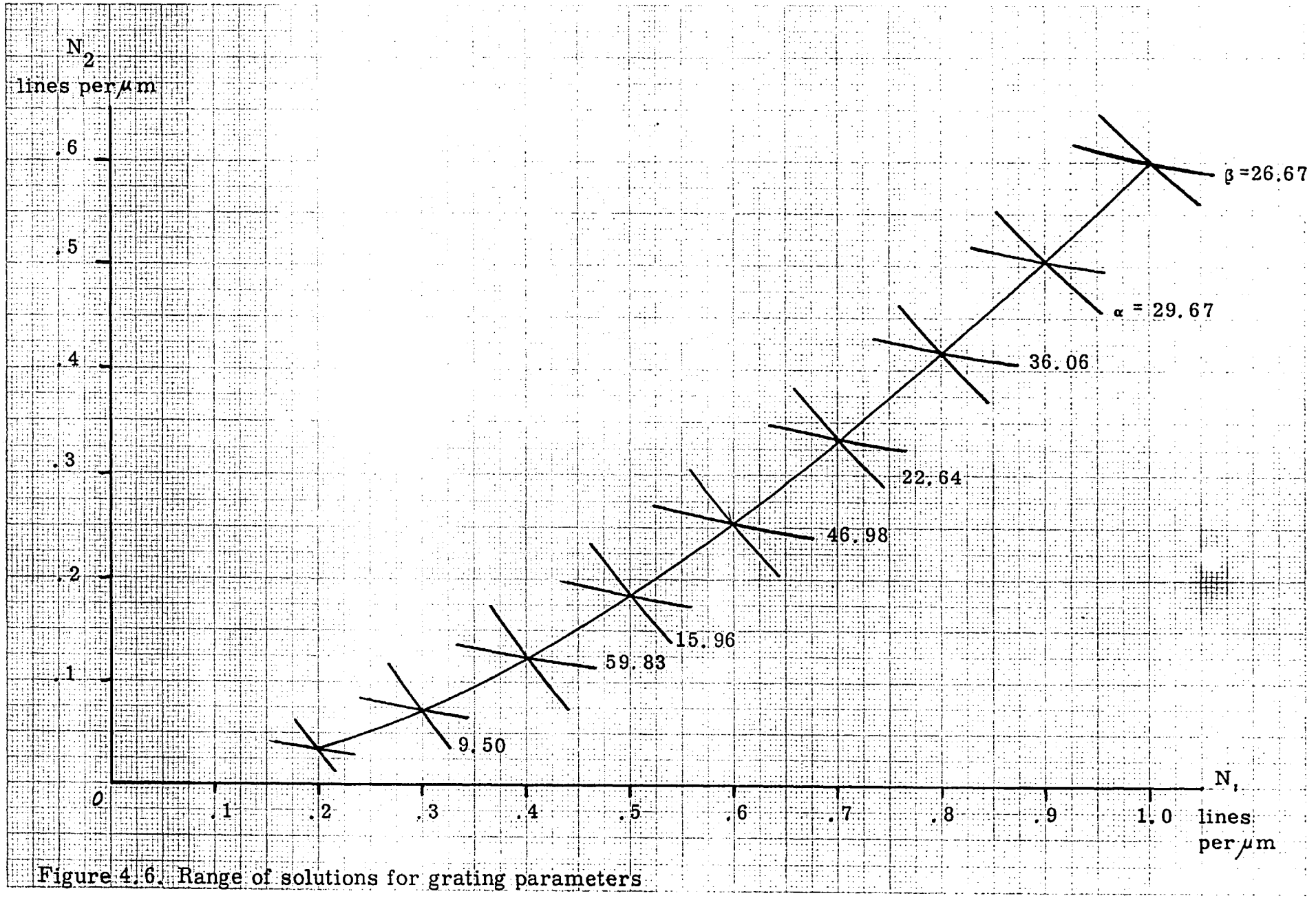


Figure 4.6. Range of solutions for grating parameters

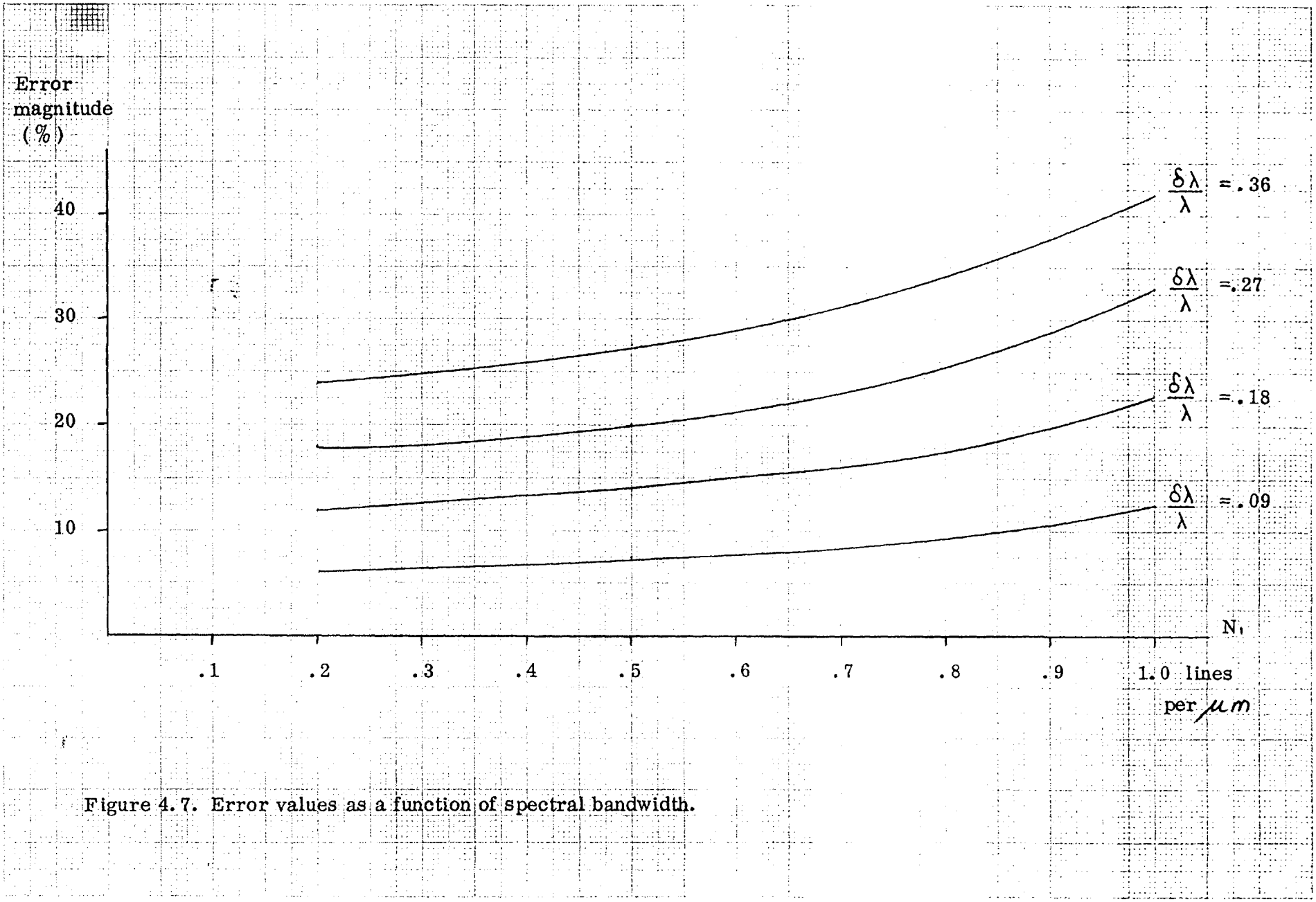


Figure 4.7. Error values as a function of spectral bandwidth.

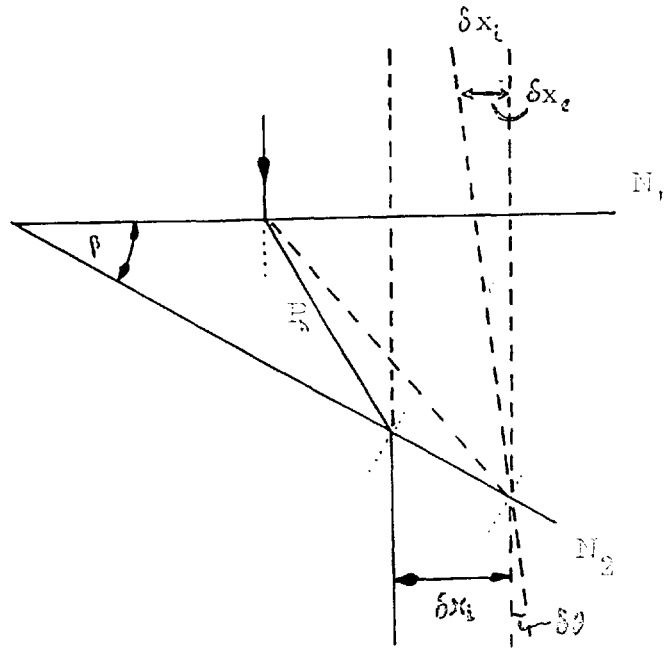


Figure 4.8. Effects of a non-zero $\delta\theta$

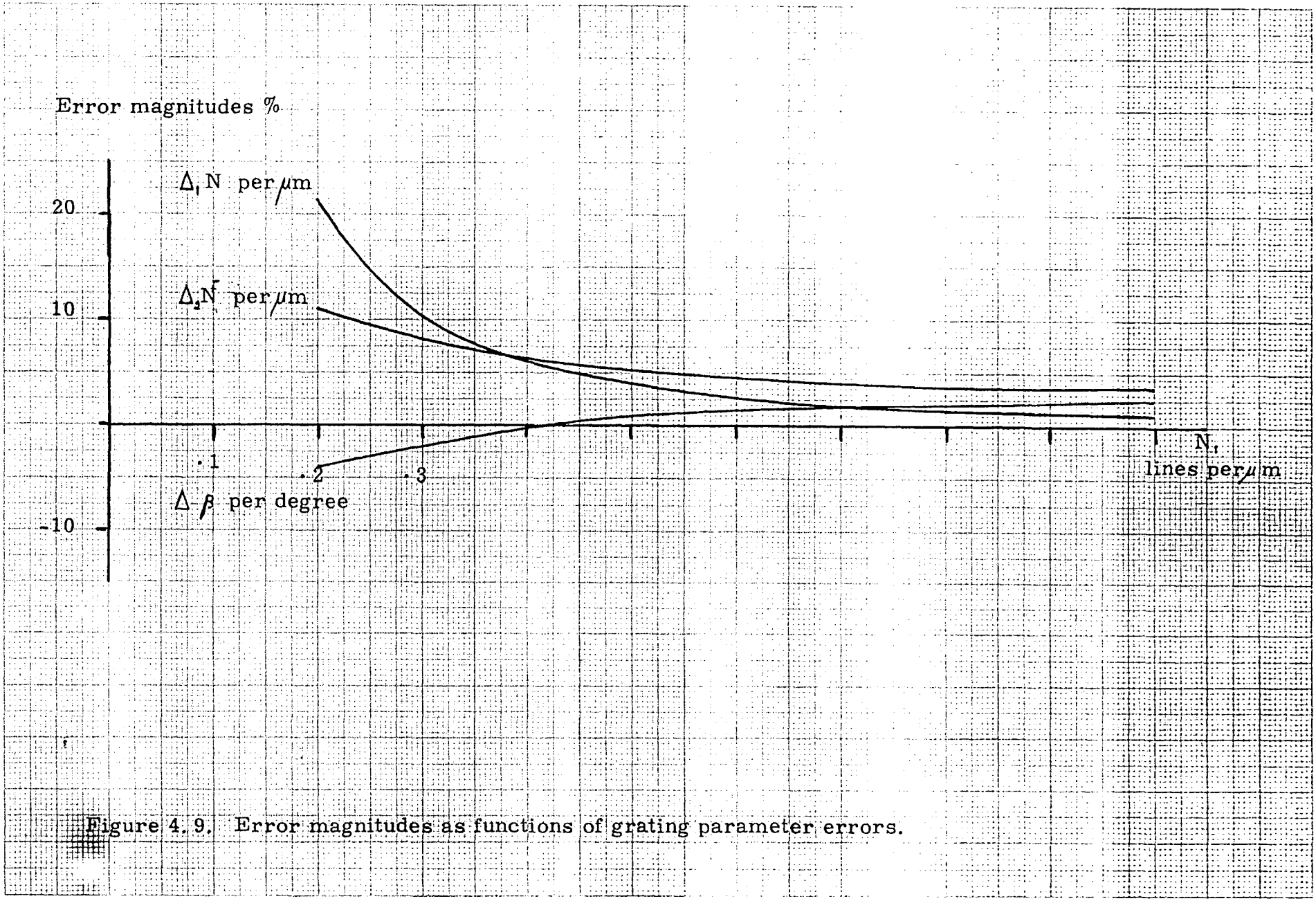


Figure 4.9. Error magnitudes as functions of grating parameter errors.

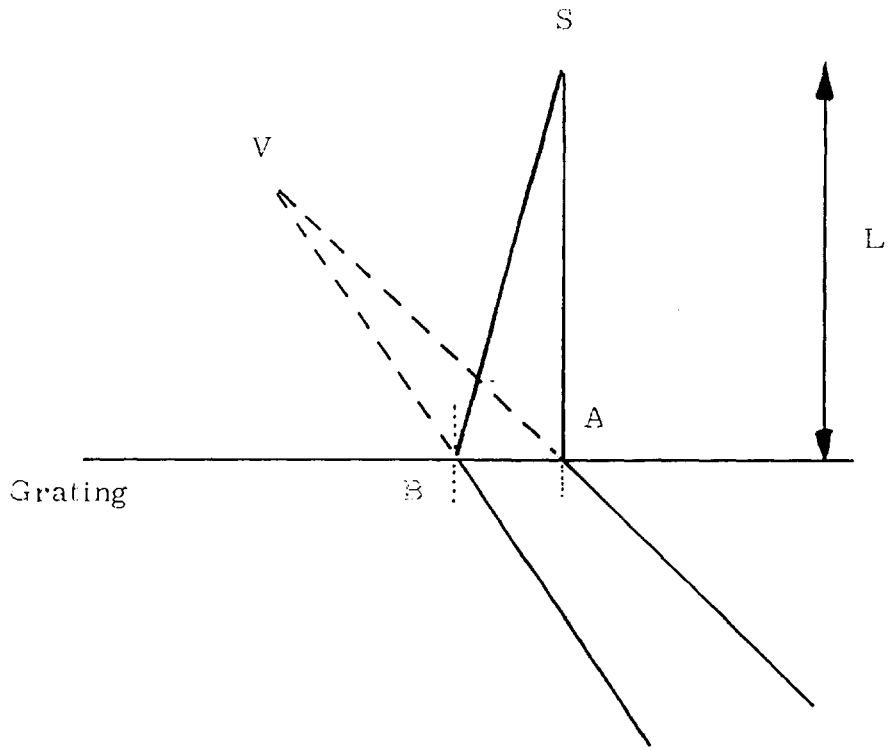


Figure 4.10 Grating Aberrations

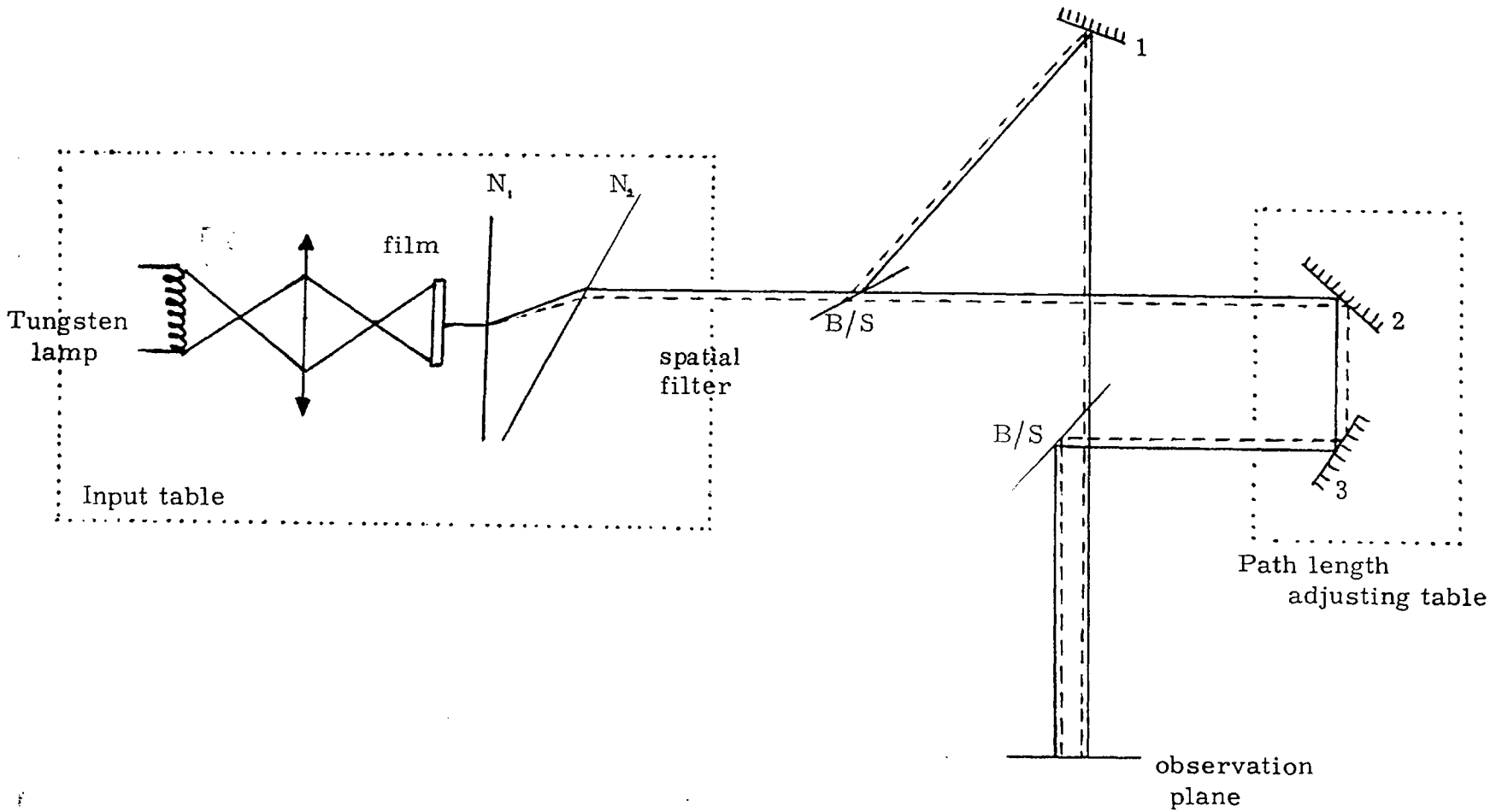


Figure 5.1. Interferometer layout.

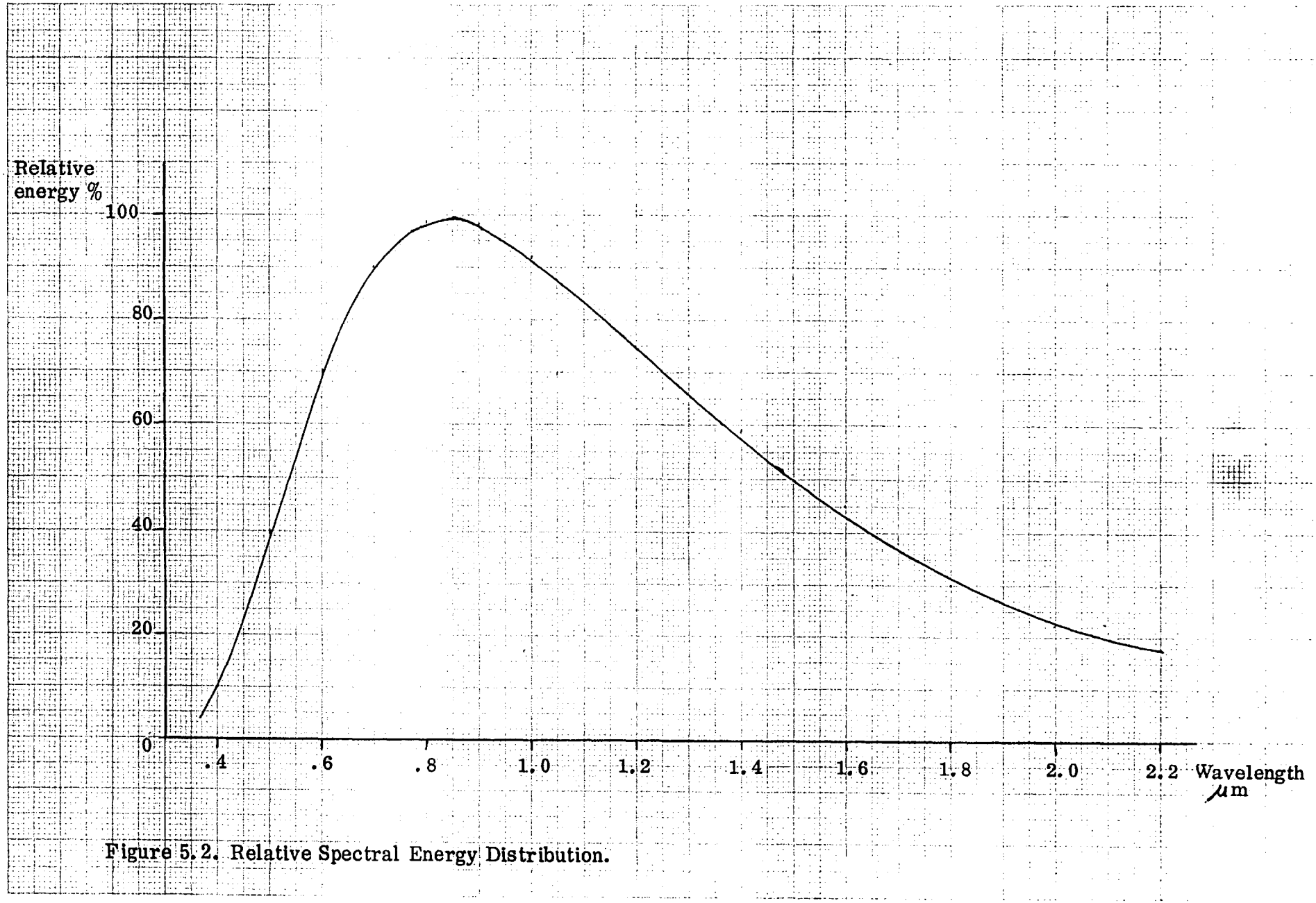


Figure 5.2. Relative Spectral Energy Distribution.

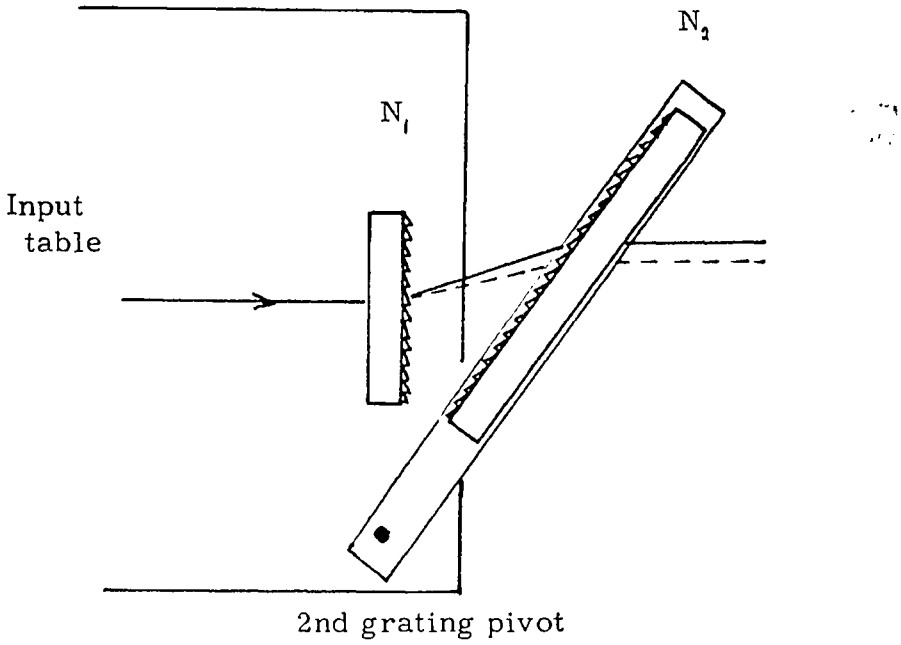
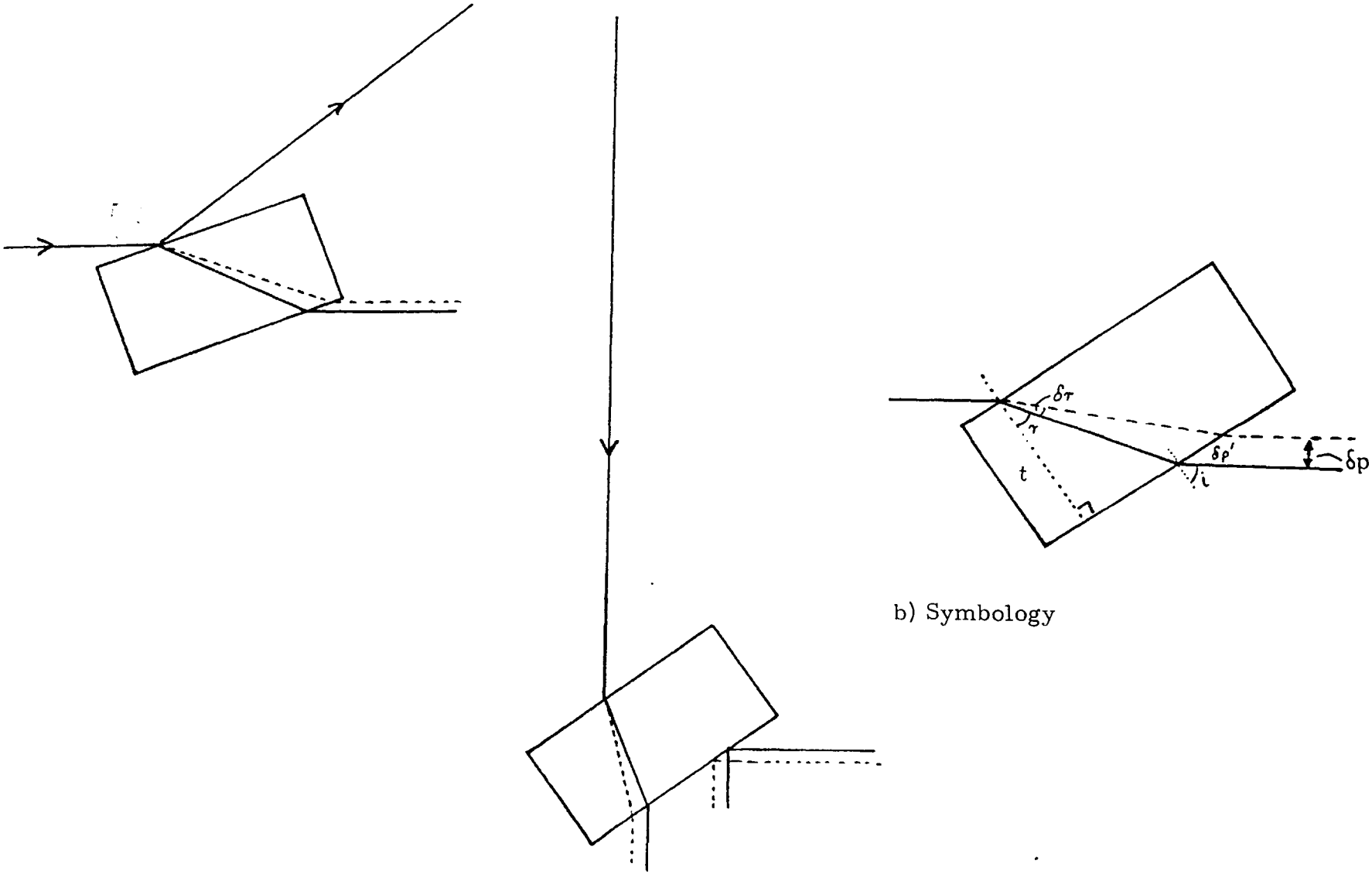


Figure 5.3 Grating mount.



b) Symbology

Figure 5.4. a) Symmetrical dispersion.

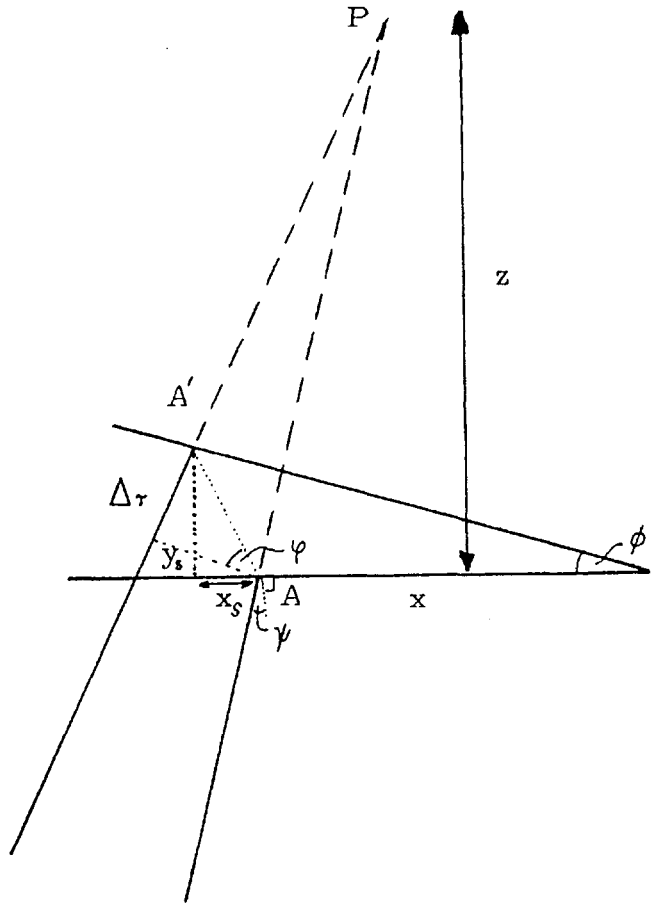


Figure 5.5. Fringe formation by two sources.

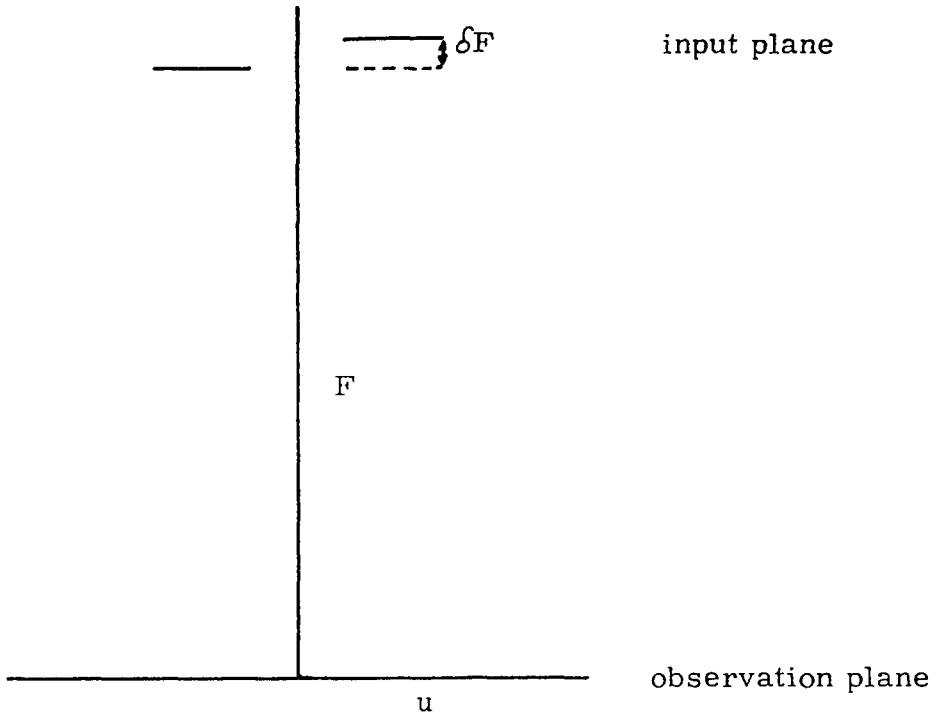


Figure 5 . 6 Schematic of an interferometer with unbalanced arms.

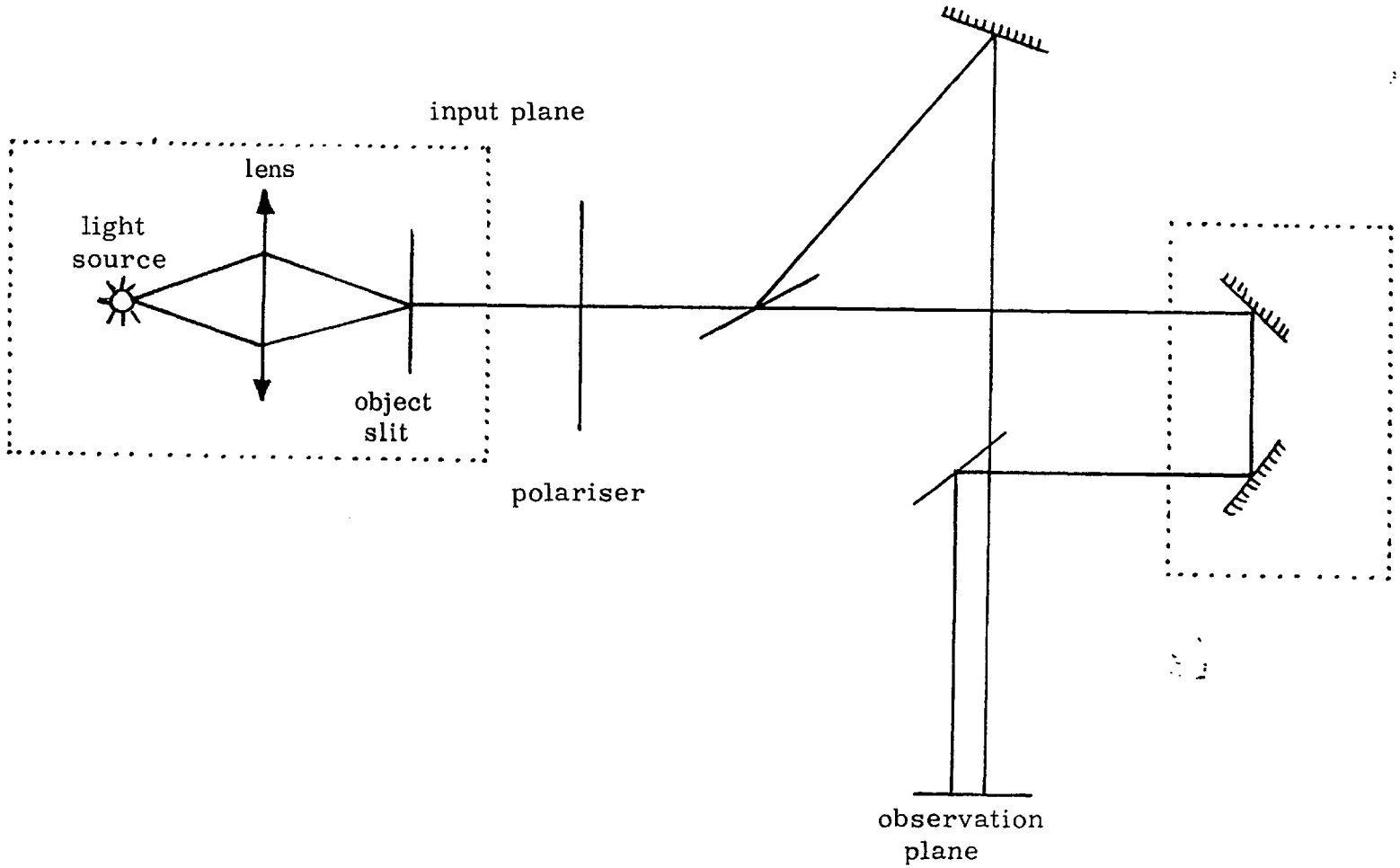


Figure 6.1. Basic Interferometer Layout.

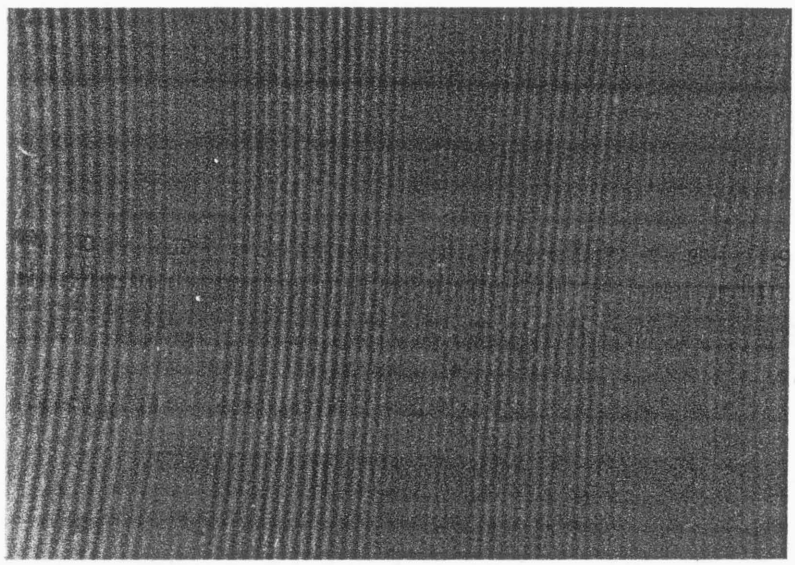


Figure 6.2. Interference fringes due to an uncompensated Hg lamp.

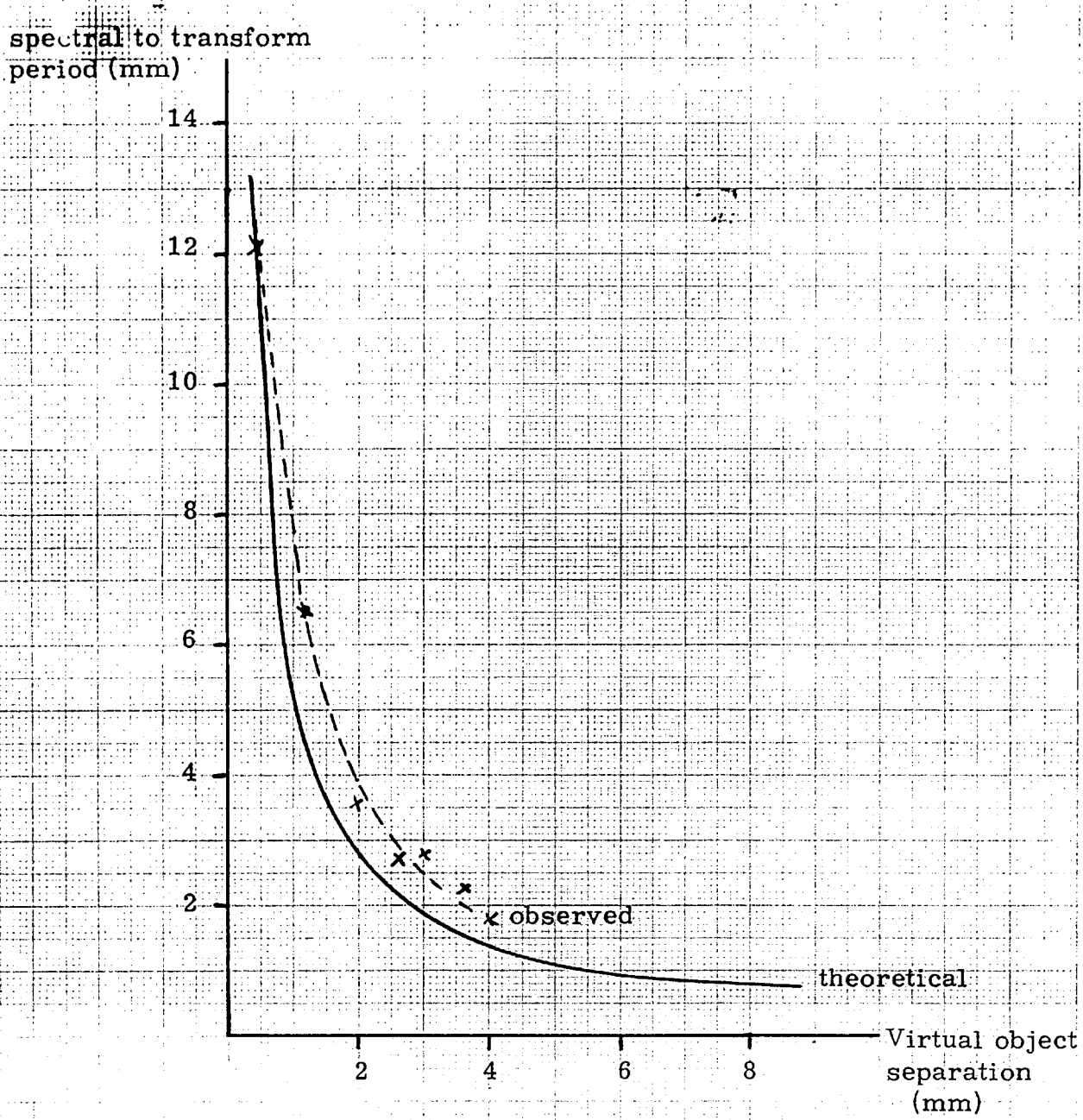


Figure 6.3. Spectral transform period as a function of input object shear.

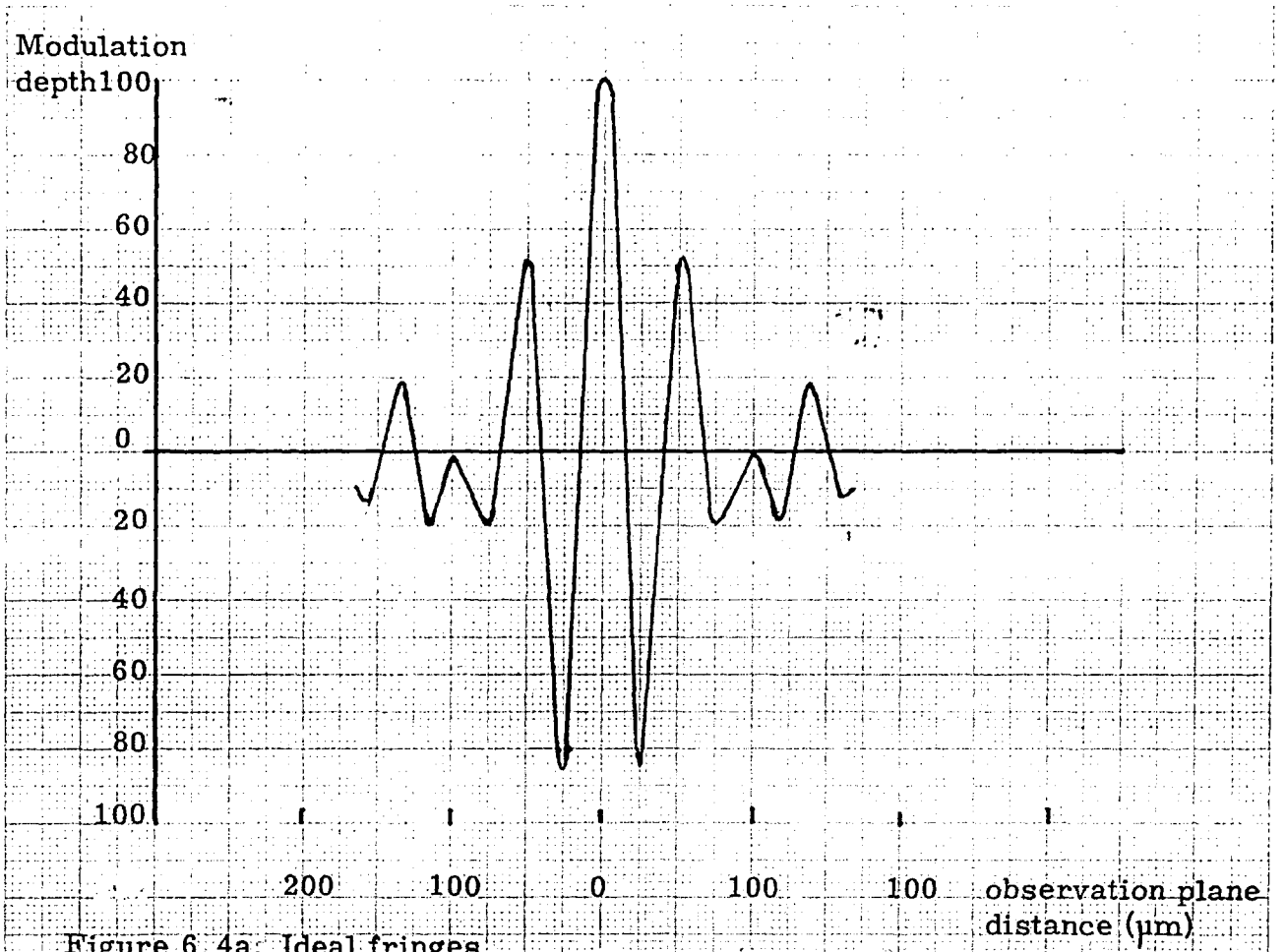


Figure 6.4a. Ideal fringes.

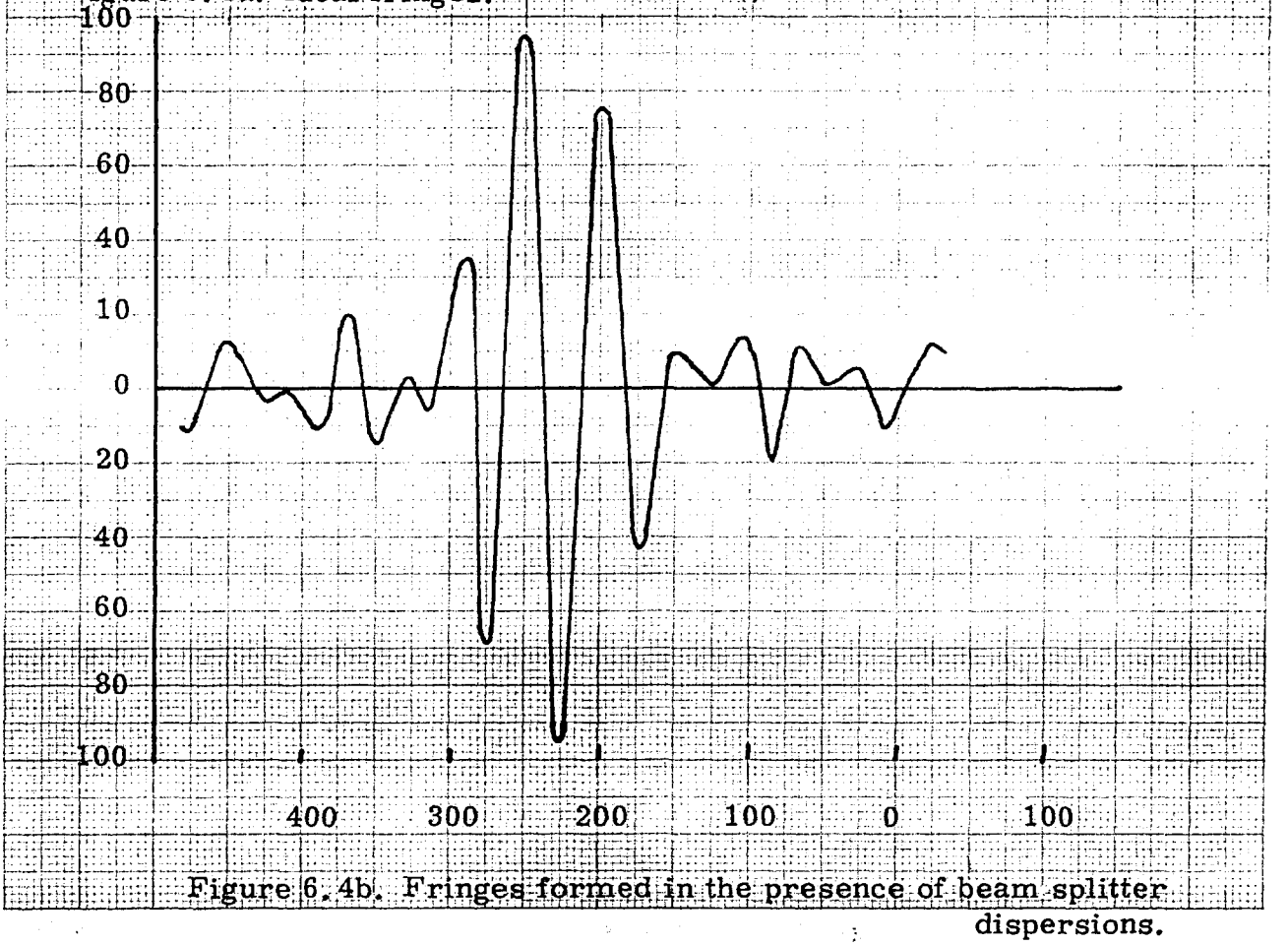


Figure 6.4b. Fringes formed in the presence of beam splitter dispersions.

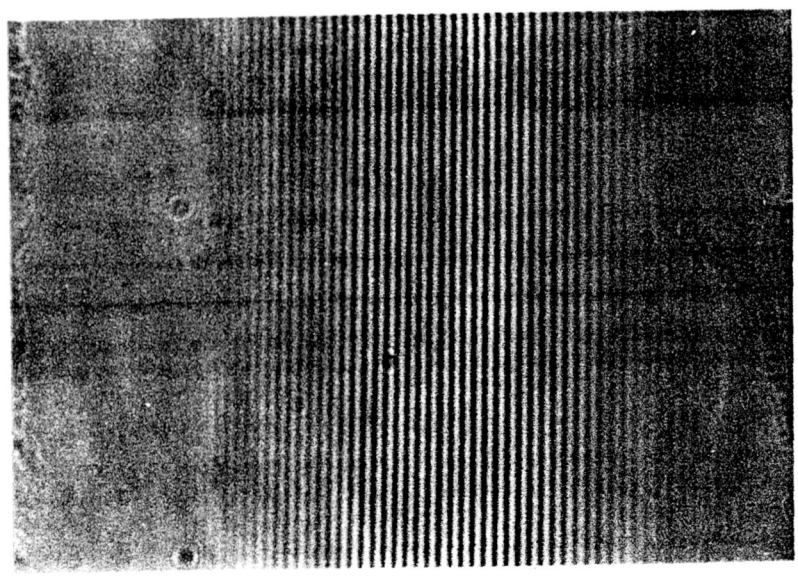


Figure 6. 5. Interference fringes due to uncompensated Tungsten-Halogen lamp.

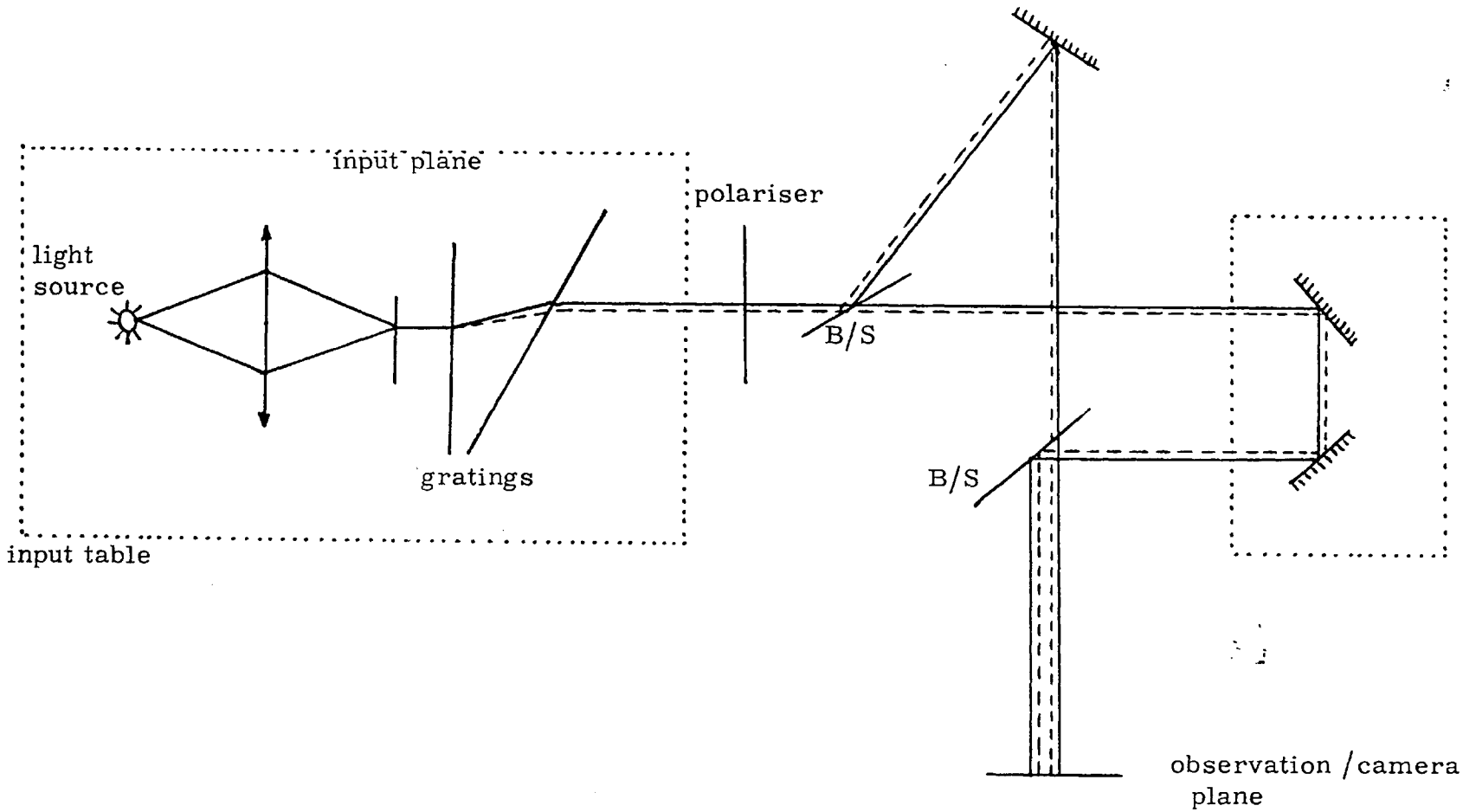
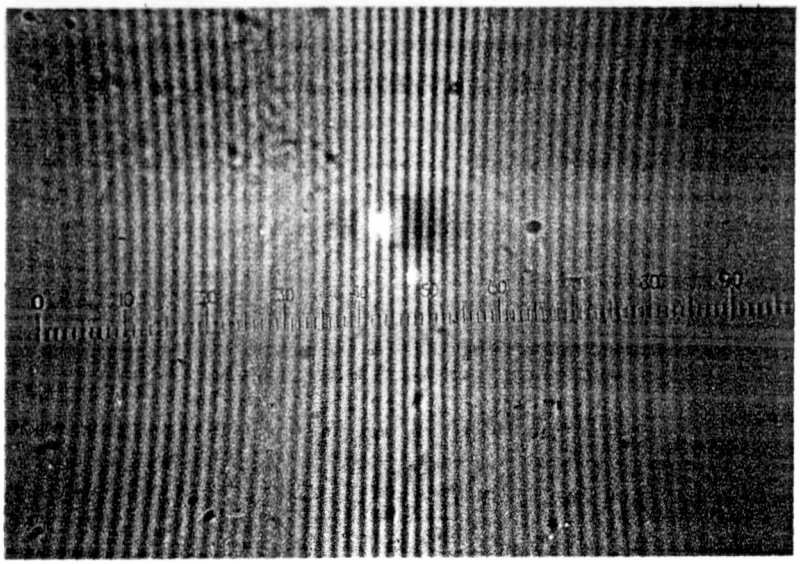
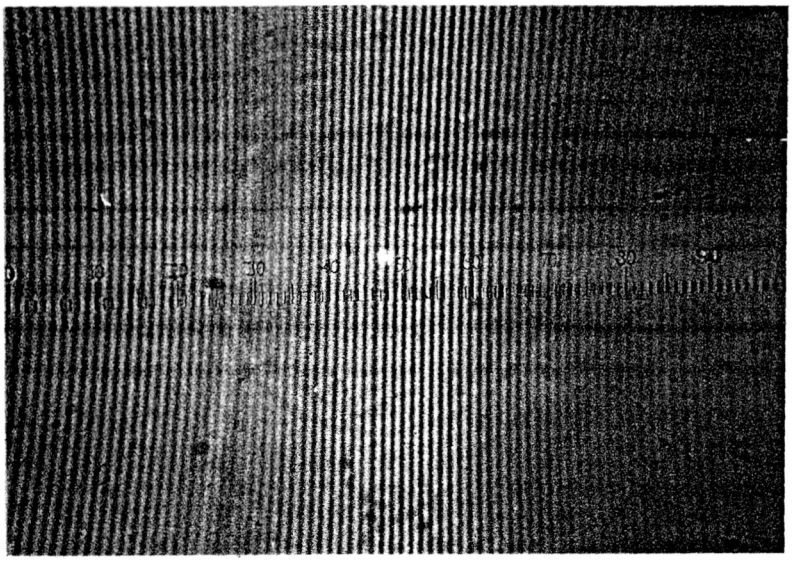


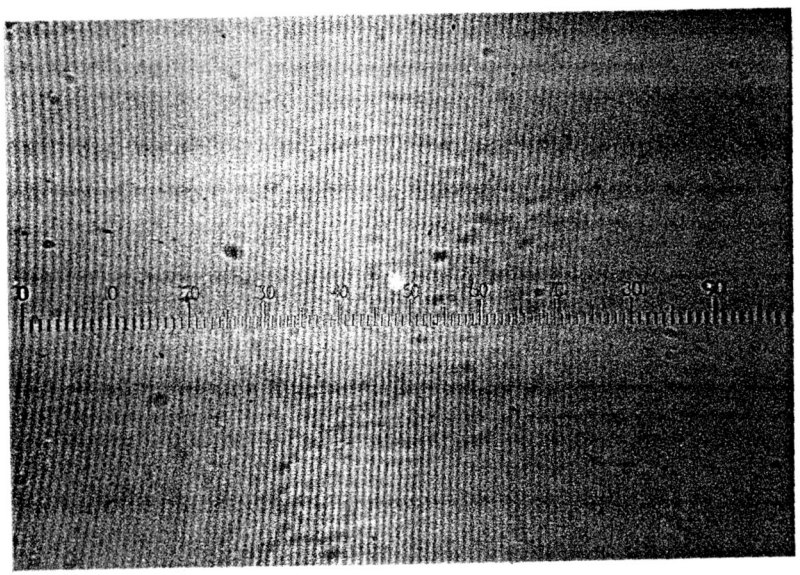
Figure 6.6. Fully compensated interferometer.



igure 6.7a



gure 6.7b



gure 6.7c

Figure 6.7. Increasing fringe numbers for different shears.

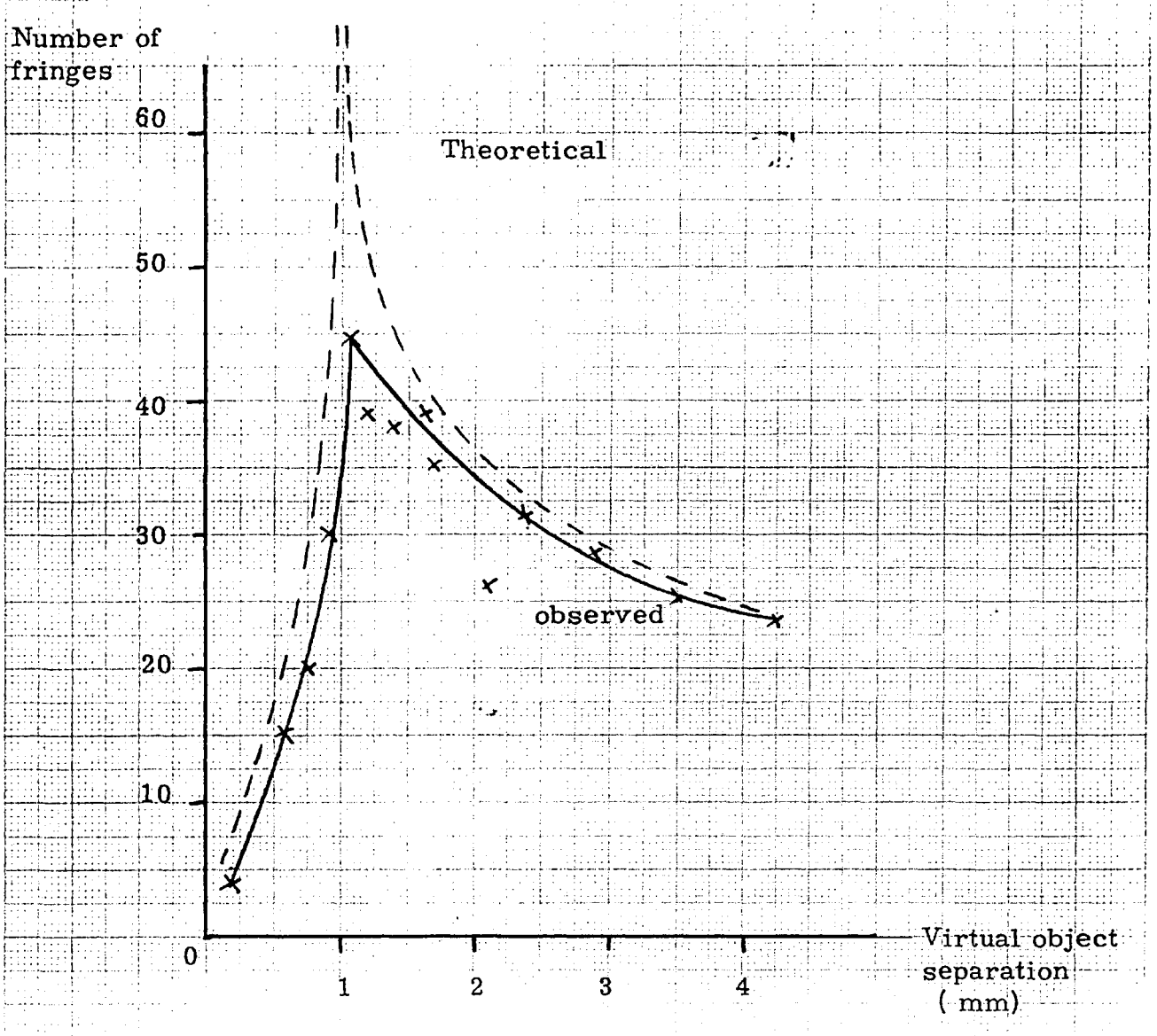


Figure 6. 8. Number of compensated Hg lamp fringes as a function of virtual object separation.

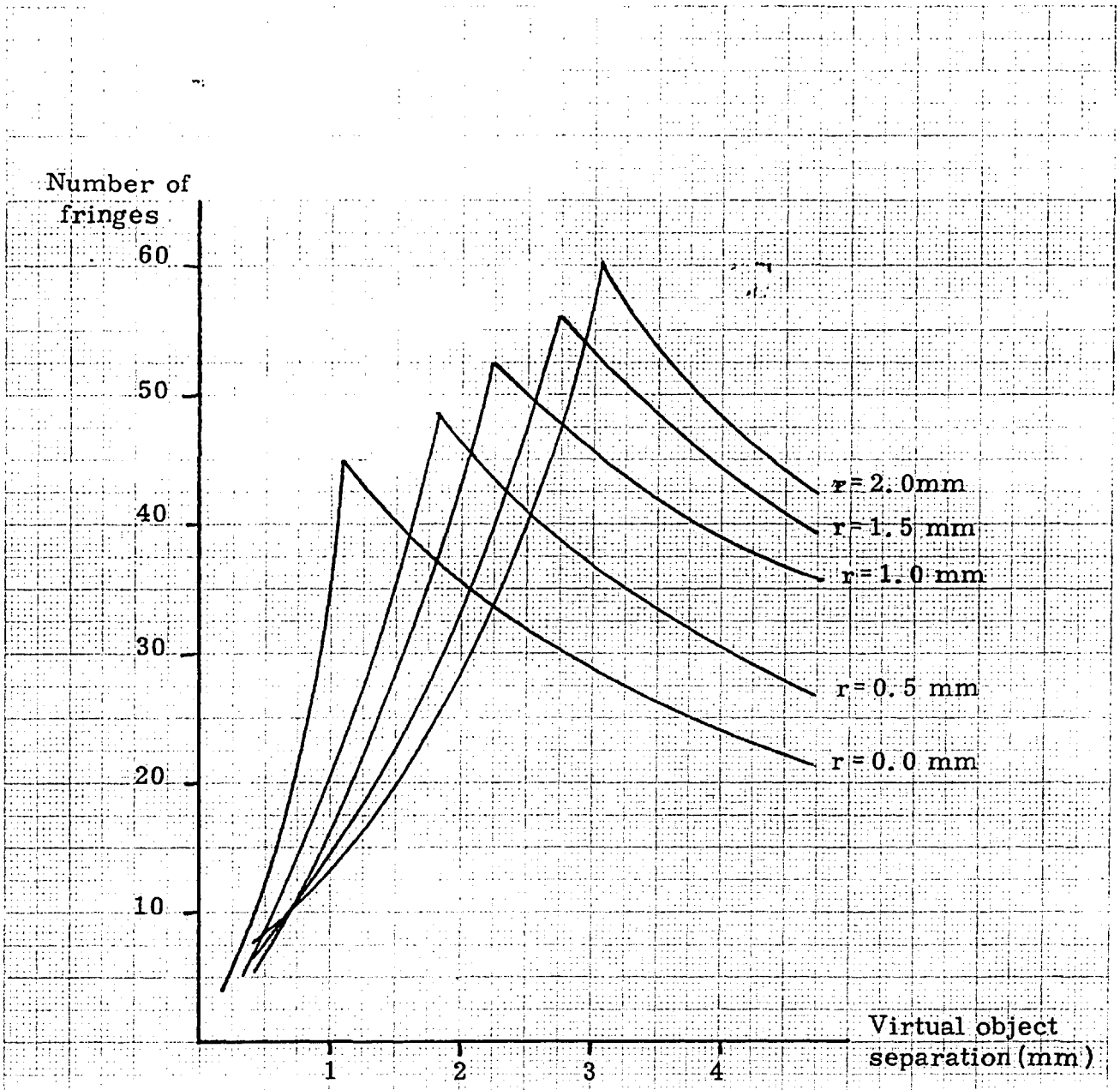
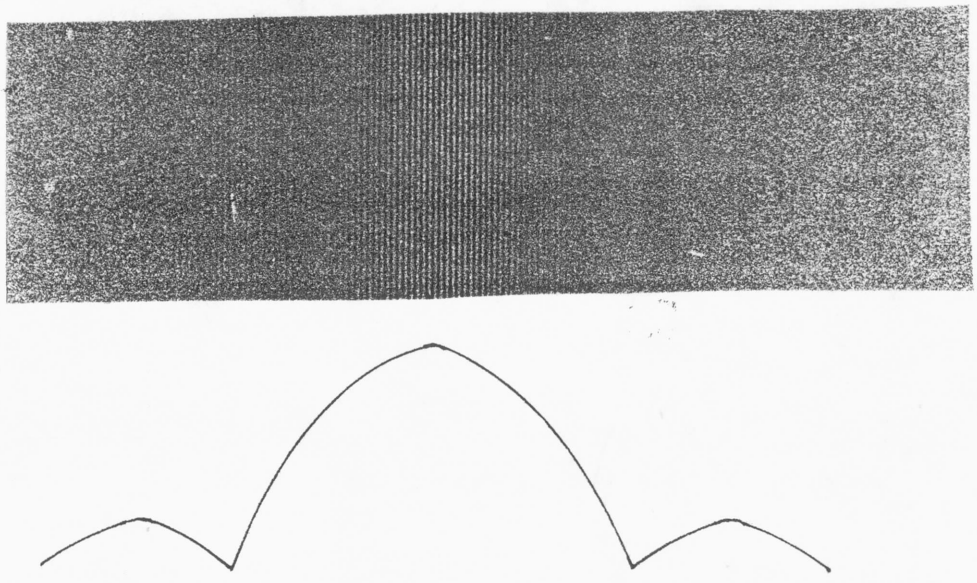
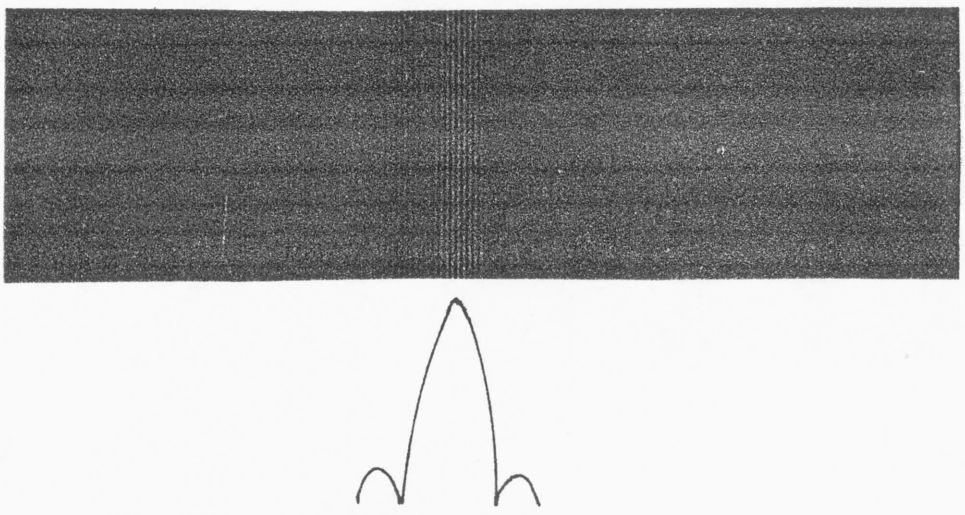


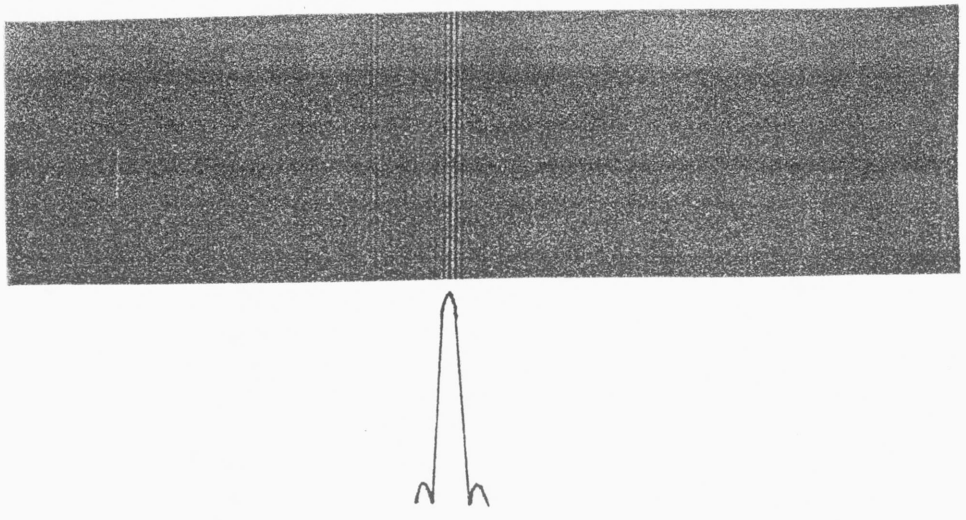
Figure 6.9. Fringe Number distributions for various input positions.



igure 6.10a



igure 6.10b



igure 6.10c

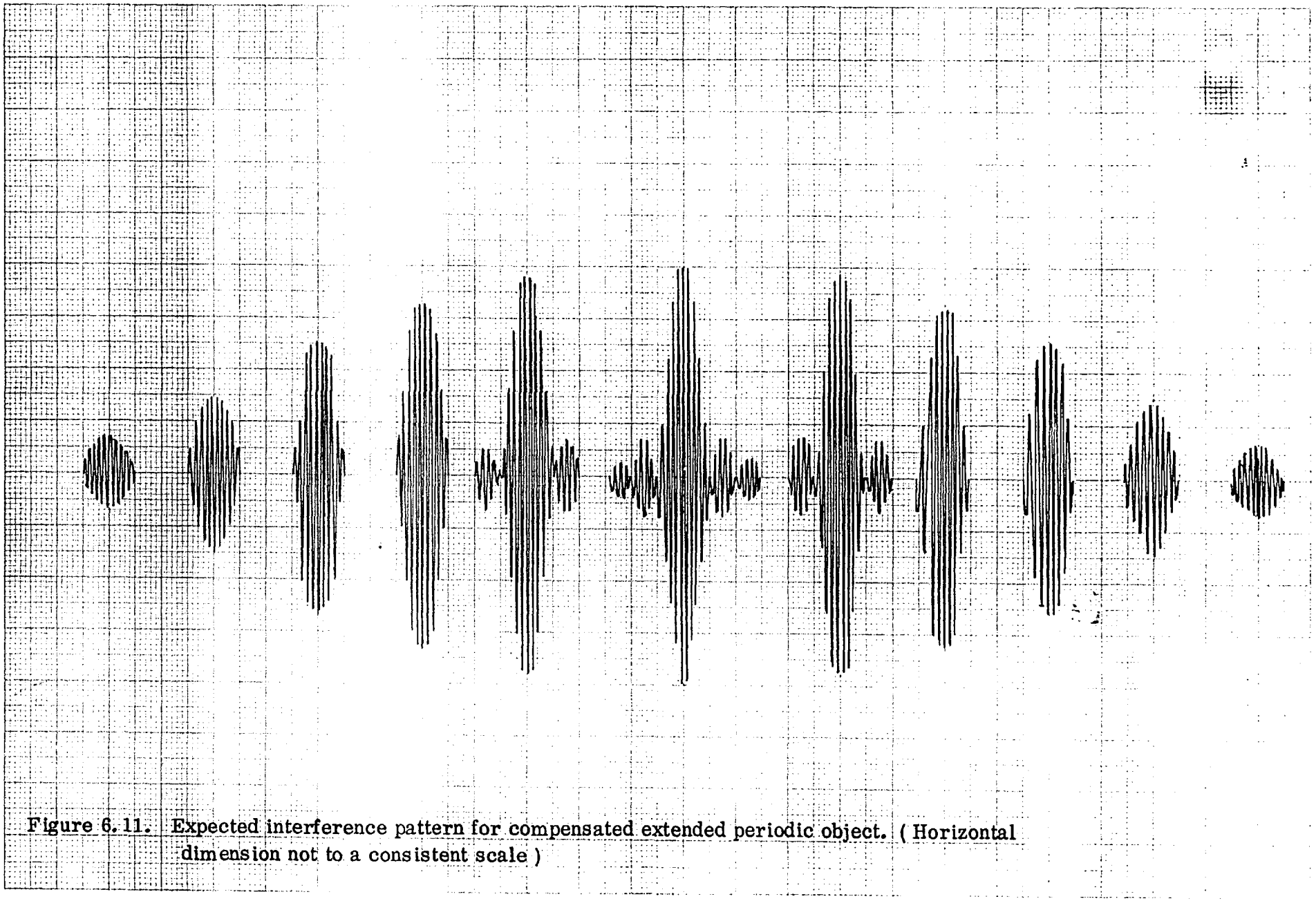
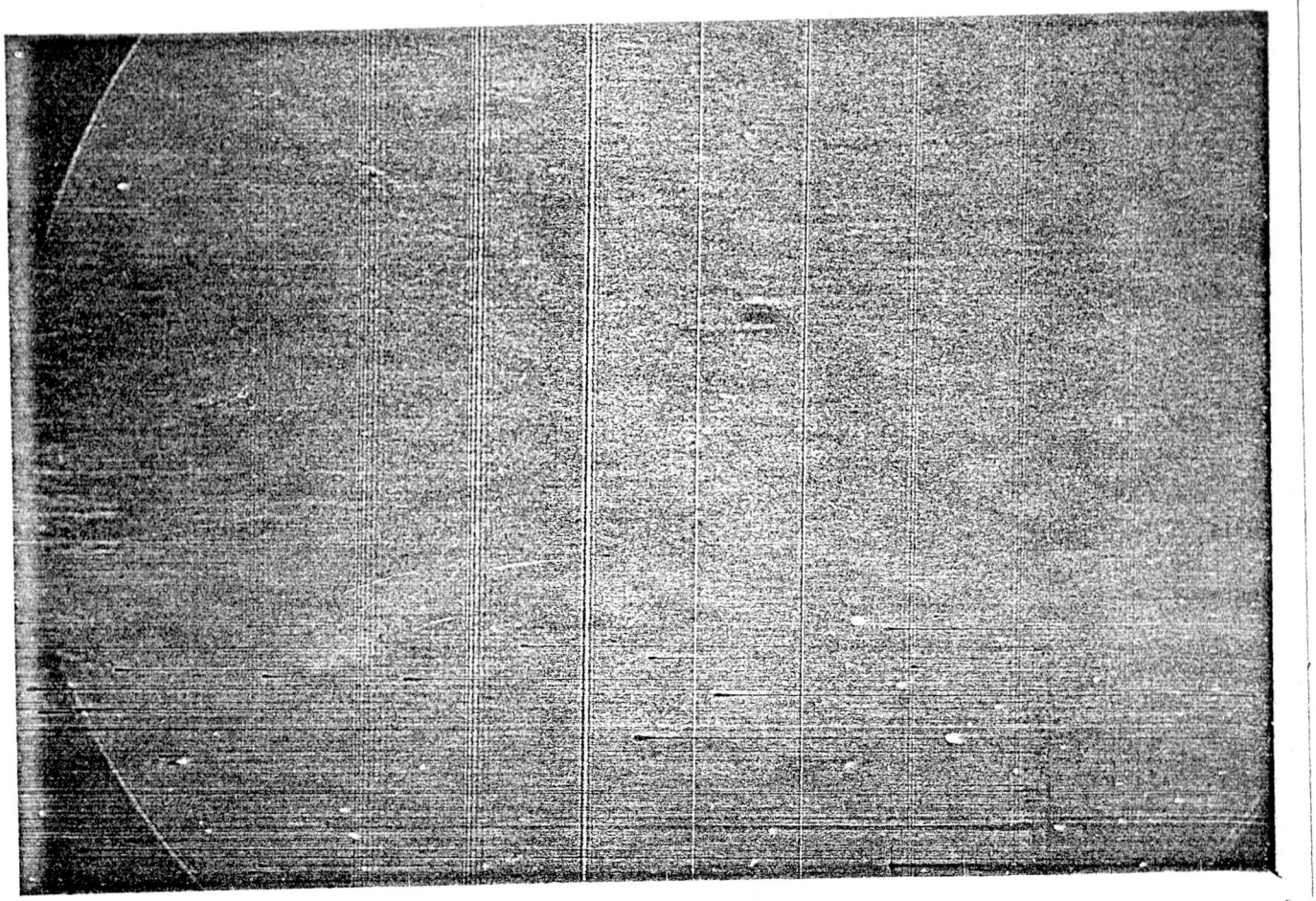


Figure 6.11. Expected interference pattern for compensated extended periodic object. (Horizontal dimension not to a consistent scale)



igure 6.12. Observed interference pattern for Compensated extended periodic object.

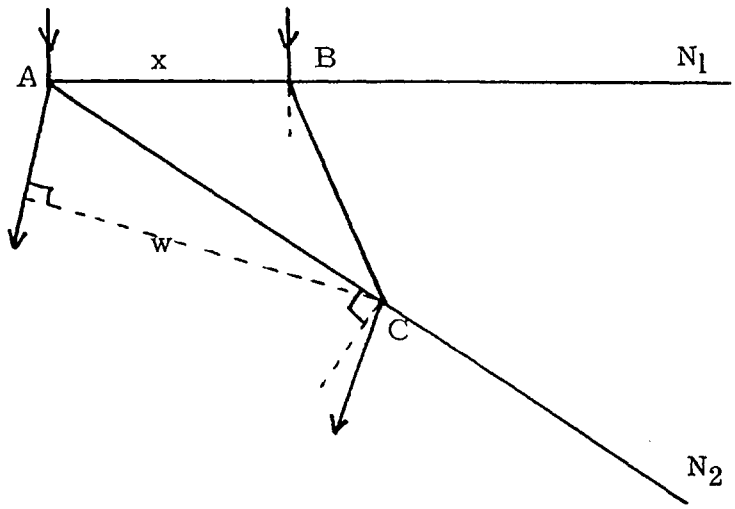
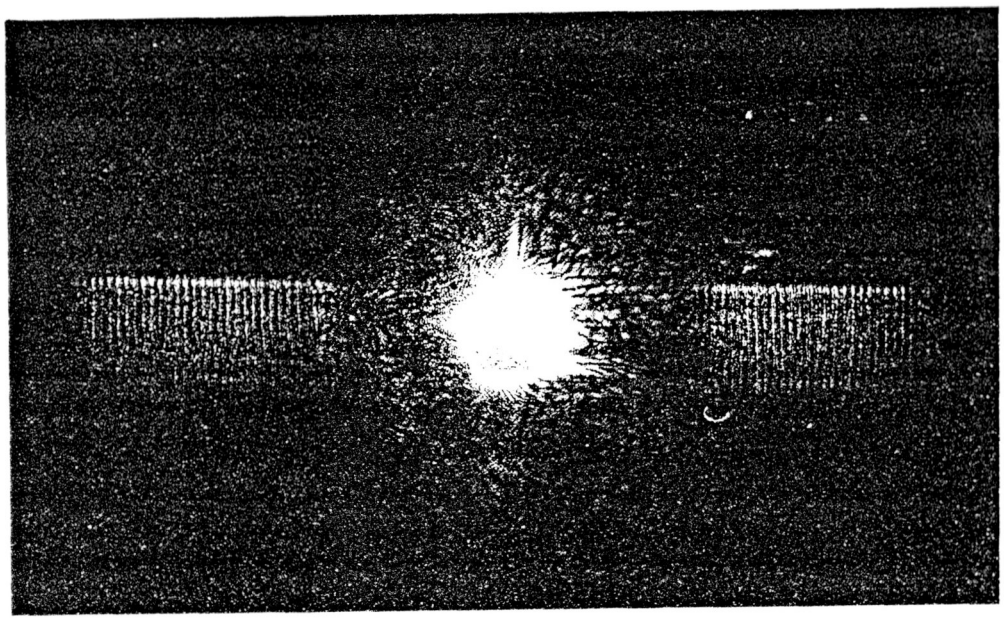


Figure 6.13. Magnification due to gratings.



igure 6.14. Reconstructed object.

REFERENCES

Allen J., Proc. IEEE, 63, 624, (1975)

Barret H.H. Wilson D.T. De Meester G.D., Opt.Comm., 5, 398, (1972)

Barret H.H. Harrington H.A., App. Opt., 12, 2686, (1973)

Bertolotti M. Muzii L. Setle D., J.Opt.Soc.Amer., 60, 1603, (1970)

Beutler H.G., J.Opt.Soc.Amer., 35, 311, (1945)

Born M. Wolf E., "Principles of Optics", Pergamon Press 1970 4th. ed.

Bose S.M., PhD Thesis, Imperial College, to be published.

Breckinridge J.B., App.Opt., 11, 2996, (1972)

Breckinridge J.B., App.Opt., 13, 2760, (1974)

Bryngdahl O. Lohmann A., J.Opt.Soc.Amer., 60, 281, (1970)

Chang B.J., Opt.Comm., 9, 357, (1973)

Crochiere R.E. Oppenheim A.V., Proc.IEEE, 63, 581, (1975)

Cutrona L.J. et al., IRE Trans.Info. Theory, IT-6, 386,(1960)

Cutter B.W. Lohmann A., Opt.Comm., 12, 220, (1974)

Dainty J.C. Scaddon R.J., Mon.Nat.R.Astr.Soc., 167, 69, (1974)

Feinlab J. Oliver D.S., App.Opt., 11, 2753, (1972)

Felstead E.B., IEEE Trans.Aero.Elec.Sys., AES-3, 907, (1967)

Freeney S.L., Proc. IEEE, 63,633, (1975)

Gillieson G.H.C.P., J.Sci.Inst., 26, 335, (1949)

Goodman J.W., "Introduction to Fourier Optics", McGraw-Hill 1968

Haag G., Nature (London), 153, 81, (1944)

Hou S.L. Oliver D.S., App.Phys.Lett., 18,325, (1971)

Kato M. Suzuki T., J.Opt.Soc.Amer., 59, 303, (1960)

Kovasznay L.S.G. Arman T., Rev.Sci., 28, 795, (1957)

Lacourt A. Froully C. Neiras J., Opt.Acta, 19, 485, (1972)

Leifer I. Rogers G.L. Stephens N.W.F., Opt.Acta, 16, 535, (1969)

Leith E.N. Upatnieks J., J.Opt.Soc.Amer., 57, 975, (1967)

Leith E.N. Chang B.J., App.Opt., 12, 1957, (1973)

Leverenz H.W."Luminescence of solids",Chapman & hall 1950 p72

Lipson S.G. Nissenson P., App.Opt., 13, 2052, (1974)
Lohmann A., Opt.Acta, 9, 1, (1962)
Lohmann A. , App.Opt., 7, 561, (1968)
Maloney W.T., App.Opt., 10, 2127, (1971)a
Maloney W.T., App.Opt., 10, 2554, (1971)b
McLachlen D., J.Opt.Soc.Amer., 52, 454, (1962)
Mertz L. Young N.O., Proc.ICO Conf.Opt.Inst. London (1961)
Murty M.V.R.K., J.Opt.Soc.Amer., 54, 1187, (1964)
Nissenson P. Iwasa S., App.Opt., 11, 2760, (1972)
Oliver D.S. et al., App.Phys.Lett., 17, 416, (1970)
Oliver D.S. Buchan W.R., IEEE Trans.Elec.Dev., ED-18, 769, (1971)
Philips W.R. McLachlen D., Rev.Sci.Inst., 25, 123, (1954)
Rhodes J., Am.J.Phys., 21, 337, (1953)
Robertson J.M., Phil.Mag., 13, 413, (1932)
Rogers G.L., Proc.Roy.Soc. (Edinburgh), A63, 193, (1952)
Rogers G.L., Proc.Inst.Opt.Comp.Conf. Zurich (1974)
Schneider W. Fink W., Opt.Acta, 22, 879, (1975)
Silva R. Rogers G.L., J.Opt.Soc.Amer., 65, 1448, (1975)
Sommerfeld A., "Optics" (Lectures on Theorectical Physics Vol. 4), Academic Press, New York 1954, p.247
Sondhi M.M., Proc. IEEE, 60, 842, (1972)
Stochham T.G. et al., Proc. IEEE, 63, 678, (1975)
Stoke G.W. Restrick R.C., App.Phys.Lett., 7, 229, (1965)
Vander Lugt A., IEEE Trans.Info. Theory, IT-10, 139, (1964)
Vander Lugt A., App.Opt., 6, 1221, (1967)
Vohl P. Nissenson P. Oliver D.S., IEEE Trans.Elec.Dev., ED-20, 1032, (1973)
Wessely H.W. Bolstad J.O., J.Opt.Soc.Amer., 60,678, (1970)
Wilson D.T. et al., Opt.Comm., 8, 384, (1973)
Wood L.C. Treitel S., Proc.IEEE, 63, 649, (1975)
Worthington H.R., J.Opt.Soc.Amer., 56, 1397, (1966)

AN ACHROMATIC WAVEFRONT FOLDING INTERFEROMETER

G.R. WLOCH and J.R. COZENS
Electrical Engineering Department, Imperial College, London S.W. 7, UK

Received 8 March 1976, revised version received 14 April 1976

A wavelength compensated wavefront folding interferometer is described which produces over 200 achromatic fringes from extended polychromatic objects.

In a wavefront folding interferometer, light from a two-dimensional, quasi-monochromatic incoherent light distribution is amplitude divided, and one of the beams subjected to a shear and to a two-fold rotation, before interference fringes are formed. The intensity distribution in the fringe pattern represents directly the two-dimensional cosine Fourier transform of the initial intensity distribution, with the point of folding as the co-ordinate centre [1,2]. For a polychromatic input the number of fringes visible is limited, thus restricting the range of spatial frequencies observable in the transform [3,4].

Consider the operation of a typical interferometer with a polychromatic input and, for simplicity, a single fold. By taking a single fold, the fringe pattern represents only a one-dimensional transform of the input distribution, and can usefully be applied only to line inputs or to inputs with no variations in the orthogonal direction. A point source, spectral width $\pm\delta\lambda$ about λ, would appear as two sources, at say $\pm x$ from the axis, as shown in fig. 1. These virtual polychromatic point sources would produce approximately N fringes of periodicity $\lambda F/2x$, where $N = \lambda/2\delta\lambda$. For white light, $N \simeq 3$.

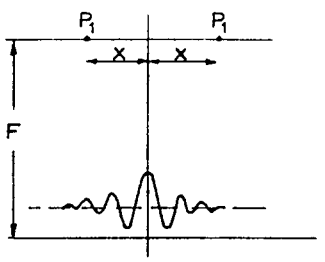


Fig. 1. The fringe pattern for a polychromatic point source P_1 with no dispersed displacement.

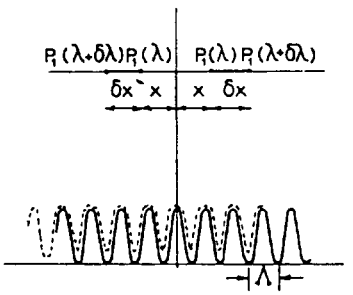


Fig. 2. The fringe pattern for a polychromatic point source P_1, with correct displacements. The radiation at $(\lambda + \delta\lambda)$ appears to be shifted to $(x + \delta x)$ so that the periodicity of all the fringes, Λ, is constant i.e. $\Lambda = \frac{1}{2}\lambda F/x = \frac{1}{2}(\lambda+\delta\lambda)F/(x+\delta x)$.

If a critical amount of dispersed displacement can be introduced in the interferometer, so that radiation from the point source at $(\lambda + \delta\lambda)$ appears to originate at $\pm(x + \delta x)$, (fig. 2) then it can easily be shown that a large number of white fringes will be observed if

$$\delta x = x\,\delta\lambda/\lambda. \qquad (1)$$

These fringes are produced by the superposition of many separately monochromatic fringes of the same periodicity.

Thus the possibility of compensating for the wide optical spectrum exists. A number of authors have de-

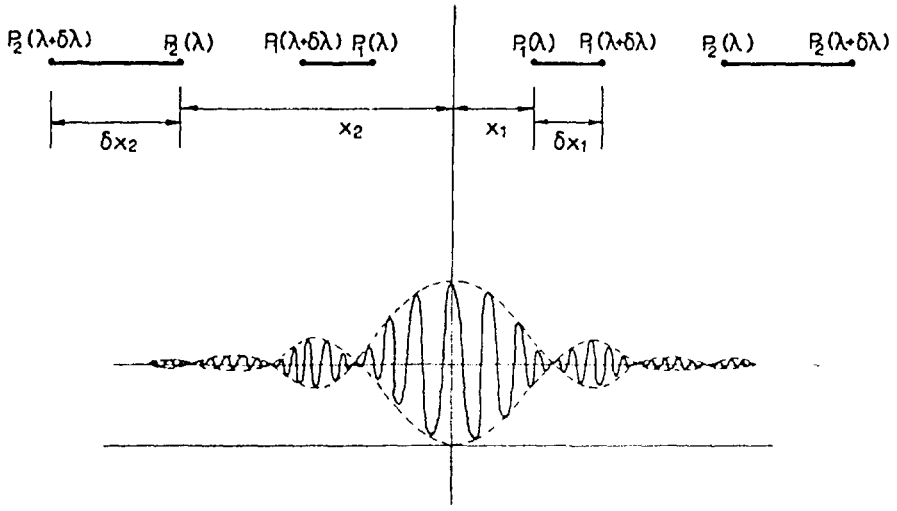


Fig. 3. The point sources P_1 and P_2 are at the extremities of a line source with displacements such that visibility is determined by the Fourier transform of the source intensity, and not its spectral composition.

scribed techniques, usually involving a pair of parallel gratings, or their equivalents, for obtaining achromatic fringe patterns, which are essentially similar to this [5–8].

An obvious limit to the usefulness of this type of compensation, for the applications considered here, is that the required displacements are functions of source position x, as well as $\delta\lambda$ [eq. (1)]. Hence this form of compensation, in which δx is proportional to $\delta\lambda$, is ideally valid only for a point object situated in a unique position. If the displacements can themselves be made to vary linearly with source position, then many achromatic fringes can be realized with objects of any dimension (fig. 3).

One way in which this can be approached is by means of two transmission gratings with different numbers of lines, mutually inclined at a suitable angle. Such an arrangement has been investigated in a single fold interferometer, as shown in fig. 4. The gratings had 590 and 250 lines per mm respectively and were mounted with an angle of 48° between them. These gratings do not form an ideal combination, but were the nearest readily available. The object used was a 10 lines per mm binary transmission grating, with a mark to space ratio of 4:1. A tungsten halide projection lamp was focussed onto this object grating.

The light distribution in the output plane was recorded on a photographic emulsion, which was subsequently analysed on a microdensitometer. A typical result is shown in fig. 5. The fringe visibility is an expected for the cosine Fourier transform of the grating.

The most significant feature of the result is that white fringes have been formed over a wide range of the output plane; the equivalent number of visible fringes being in excess of 200. The size of each patch of fringes is determined by the overall object width, which in this case was 1.2 mm. If this is increased, the width of each path is correspondingly reduced, but the range over which patches are observed, and

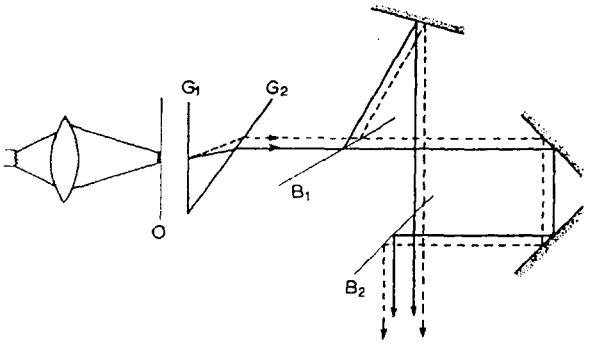


Fig. 4. Schematic view of the interferometer, showing folding and dispersion. o – object plane; G_1,G_2 – transmission gratings (G_1 = 590 lines/mm, G_2 = 250 lines/mm) B_1, B_2 – beam splitters. —— path followed by light, wavelength λ; – – – path followed by light, wavelength $\lambda + \delta\lambda$.

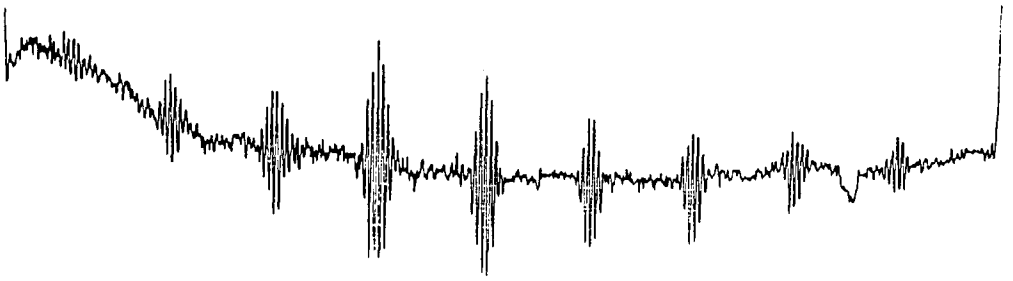


Fig. 5. Microdensitometer record of a photograph of white fringes formed with the interferometer, for a binary grating object.

hence the equivalent number of fringes, is not reduced. Object dimensions up to 1 cm have been used without a decrease in spatial resolution. This limit is imposed by the size of the lamp filament.

It is interesting to note that to achieve this number of fringes without dispersion, a filter of width $\simeq 30$ Å would be required. This system can thus be regarded as offering an effective power gain of $\simeq 85$ by making the full spectral width of the source available.

One possible area of application for this technique, extended to two-dimensions, is in the real time processing of information presented as a self-liminous distribution, e.g. a C.R.T. screen or an astronomical source. The transform could either be recorded, or processed with a two-dimensional filter in the interference plane and read-out with a camera tube. Alternatively, the camera video signal could be filtered conventionally with the re-transformation being subsequently carried in a second interferometer.

References

[1] L. Mertz, Transforms in Optics (Wiley, New York, 1965), p. 111.
[2] G.W. Stroke and R.C. Restrick, App. Phys. Lett. 7 (1965) 229.
[3] J.C. Dainty and R.J. Scaddon, Mon. Nat. R. Astr. Soc. 167 (1974) 69.
[4] J.B. Breckinridge, App. Opt. 13 (1974) 2760.
[5] E.N. Leith and J. Upatnieks, J. Opt. Soc. Am. 57 (1967) 975.
[6] O. Bryngdahl and A. Lohmann, J. Opt. Soc. Am. 60 (1970) 281.
[7] F. Bridow and J.J. Clair, Opt. Act. 19 (1972) 865.
[8] D.W. Cutter and A.W. Lohmann, Opt. Comm. 12 (1974) 220.

Reprinted from

Applications of

Holography and Optical Data Processing

Proceedings of the International Conference

Jerusalem, August 23-26, 1976.

Editors Marom, Friesem and Wiener-Auneav

Pengamon Press 1976.

ACHROMATIC FRINGES FROM EXTENDED SOURCES

G. R. Wloch, S. M. Bose and J. R. Cozens

Electrical Engineering Department, Imperial College, London

ABSTRACT

A wavelength compensated wavefront folding interferometer is described which produces over 400 achromatic fringes from extended polychromatic objects.

INTRODUCTION

The inherently two-dimensional nature of optical wavefronts leads to the expectation that the potential of parallel optical signal processing would prove considerable. The technique most widely developed so far involves coherent light, its power deriving from the remarkable Fourier transform property of lenses. While there are many applications for which this system is admirable, there are many others for which the requirement that the information has to be impressed onto a coherent carrier is very restrictive. Furthermore, coherent processing relates the transforms of complex amplitudes, so that the Fourier transform of a signal cannot be recorded directly with a detector.

If we consider the form in which two-dimensional signals are likely to occur in practice, among the most obvious examples might be photographic transparencies, optical images (i.e. intensity distributions) and cathode ray tube displays. Only the first of these is readily imposed onto a coherent beam.

One useful development might therefore be towards parallel processing systems that can use direct signals as inputs. We have chosen to aim for a system that could use, for example, a C.R.T. screen display directly as an input signal. The system must therefore be able to operate with spatially incoherent, polychromatic signals in real time.

The system we are investigating is based on the wavefront folding interferometer (Refs. 1-4). In this instrument, the light from a two-dimensional incoherent source is amplitude divided, with one beam then being subjected to a relative two-fold rotation and shear. Thus the intensity g (x,y) at a point in the original source will become $g(x-X_o, y-Y_o)$ and $g(-x-X_o, -y-Y_o)$ in the two virtual sources for shears of $\pm (X_o, Y_o)$, respectively. Subsequent interference between the beams is effected in the plane of a screen or detector some distance from the source, Fig. 1. If we represent the source as an array of mutually independent point sources, then only corresponding points in (a b c d) and (a'b'c'd') can interfere in the output (ξ ,η) plane. The complementary sources at (x,y) and (-x,-y) will produce a sinusoidal intensity distribution in the output plane. These fringe patterns for each pair of points will add in intensity, leading to the intensity distribution in the

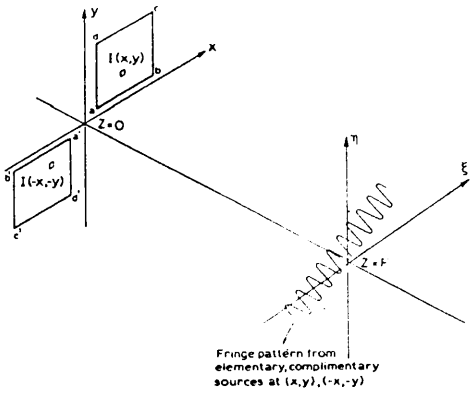


Fig. 1 Intensity distributions in the input and output planes of a wavefront folding interferometer

output plane (for large F)

$$I^1(\xi,\eta) = \frac{2}{(\lambda F)^2} \int_0^\infty I(x,y) \left\{ 1 + \cos \frac{2k}{F} \left[\xi x + \eta y \right] \right\} dxdy \qquad (1)$$

where $I(x,y) = g(x-X_o, y-Y_o)$. Thus the output contains a spatial representation of the cosine Fourier transform of the source intensity distribution which can be filtered, and also directly recorded.

We can obtain further insight into the nature of the output if eqn.(1) is written in the form:

$$I^1(\xi,\eta) = \frac{2}{(\lambda F)^2} \left[G_e\left(\frac{2\xi}{\lambda F}, \frac{2\eta}{\lambda F}\right) \cos \left\{ 2\pi \left(\frac{2\xi}{\lambda F} X_o + \frac{2\eta}{\lambda F} Y_o \right) \right\} \right.$$

$$\left. - j\, G_o\left(\frac{2\xi}{\lambda F}, \frac{2\eta}{\lambda F}\right) \sin \left\{ 2\pi \left(\frac{2\xi}{\lambda F} X_o + \frac{2\eta}{\lambda F} Y_o \right) \right\} \right] + C_1$$

where G_e and G_o are the Fourier transforms of the even (g_e) and odd (g_o) parts respectively of the function $g(x,y)$ and C_1 is a constant. Thus it is evident that the output consists of three parts:

(a) a cosine fringe pattern modulated by the transform of the even part of the input signal,

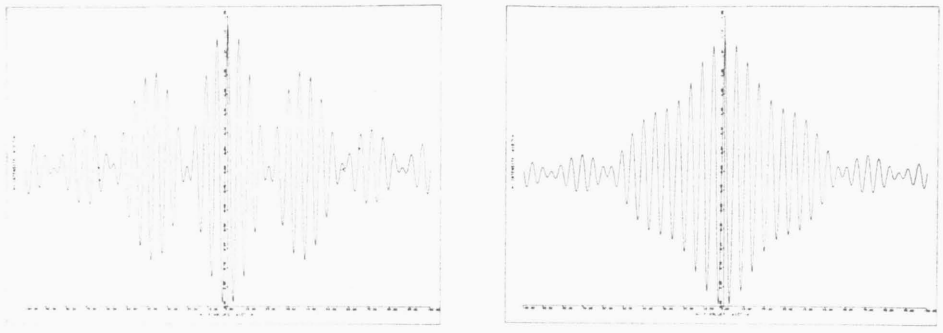
(b) a sine fringe pattern modulated by the transform of the odd part of the input,

(c) a constant intensity which allows the sinusoidal terms to be directly observed.

Since the transform of g_e is real and the transform of g_o is imaginary, complete information about the complex transform is available in the form of a real signal in the output plane. This is illustrated, Fig. 2, by means of the calculated one dimensional outputs for input functions.

$$g(x) = \text{rect}(\tfrac{x}{b}).\ (1 + \cos \tfrac{x}{d}) \qquad \text{(purely even)}$$

$$g(x) = \text{rect}(\tfrac{x}{b}).\ (1 + \sin \tfrac{x}{d}) \qquad \text{(odd + even)}$$



(a) (b)

Fig. 2 Plots of output intensity for
(a) $g(x) = \text{rect}(\frac{x}{b})\ (1 + \cos \frac{x}{d})$
(b) $g(x) = \text{rect}(\frac{x}{b})\ (1 + \sin \frac{x}{d})$
both for b = d.

The ability of the interferometer to distinguish and display both transforms is clear.

A calculated two-dimensional fringe pattern for a rectangular slot is also shown, Fig. 3, where the phase information in the transform is apparent in the fringe displacements.

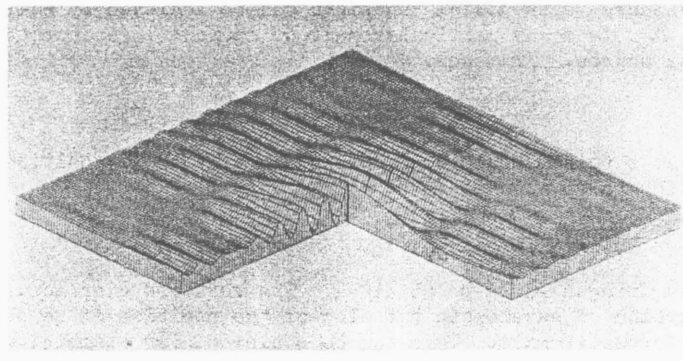


Fig. 3 2-dimensional fringe pattern for a rectangular slot input

It should be noted that the input to the interferometer is in the form of an intensity distribution and hence g(x,y) can only be real and positive. It is therefore not possible to have a truly odd input function and the output intensity distribution always has its maximum at the centre.

Bandwidth Restriction

One important restriction limits useful application of the instrument in the form so far described. As we move from the centre of the output plane, the visibility of the fringes from each pair of point sources will fall, due to their finite temporal coherence. Since regions further from the centre display the higher spatial frequencies of the source, some trade-off must exist between the optical bandwidth, $\lambda/\Delta\lambda$, and the maximum detectable spatial frequency, ω_s. From eqn.(1) we can simply show that the displacement from the origin $R(\omega_s)$ representing ω_s is

$$R(\omega_s) = \frac{F}{2k} \cdot \omega_s \qquad (2)$$

whereas the maximum range, R(max), for fringe formation by a pair of sources at $\pm(x,y)$ is

$$R_{max} = \frac{\pi F}{k(x^2+y^2)^{1/2}} \cdot \frac{\lambda}{\Delta\lambda} \qquad (3)$$

Combining eqn.(2) and (3),

$$\omega_s = 2\pi \frac{1}{(x^2+y^2)^{1/2}} \cdot \frac{\lambda}{\Delta\lambda} .$$

For fairly wide bandwidths, say $\lambda/\Delta\lambda \approx 10$, and spatial frequencies up to, say, 10 lines/m.m., source dimensions of only 1 m.m. can be accommodated. Alternatively we could use a large source and increase $\lambda/\Delta\lambda$ with a filter. Thus a T.V. screen, dimension say D_o, would have spatial frequencies up to about $2\pi 300/D$, which for $\lambda = 6000$ A, would require $\Delta\lambda \approx 20$ A. These restrictions would be clearly unacceptable in many cases.

The Achromatic Interferometer

A considerable increase in useful bandwidth can be achieved by making the fringe period, Λ, from any particular pair of sources independent for a one dimensional source. From eqn.(1)

$$\Lambda = \frac{F\lambda}{2x} ,$$

thus if all source points could be displaced so that their positions linearly shifted according to wavelength as

$$\delta x = x \cdot \frac{\delta\lambda}{\lambda} ,$$

then achromatic fringes would be obtained. The required displacement is thus not only a function of wavelength but also of the position of an object point from the interferometer axis. This can be achieved, for example, with a pair of mutually inclined transmission gratings, as shown in Fig. 4. A suitable choice of N_1, N_2 and θ leads to the formation of a large number of fringes.

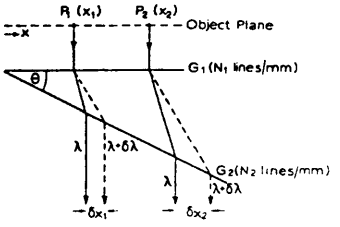


Fig. 4 The passage of two rays through a pair of inclined gratings

The ray representation indicated in Fig. 4 allows a simple illustrative analysis to be performed, but contains an uncertain level of approximation. A more rigorous treatment is outlined below.

A point source in the input plane is represented by a uniform spectrum of plane waves. The passage of each component wave through the gratings can then be described accurately by the grating equation. In Fig. 5, the component waves from a source $S(x_1, y_1, 0)$ propagate through gratings 1 and 2; after diffraction these components are modified in amplitude and phase and are expressed in the co-ordinate system (x_3, y_3, z_3). This system is then rotated by θ_0 and translated by 'a' and 'b' to the final system (x_5, y_5, z_5), centre O_5. The z_5-axis is the interferometer axis and the plane $z_5=0$ is the folding plane.

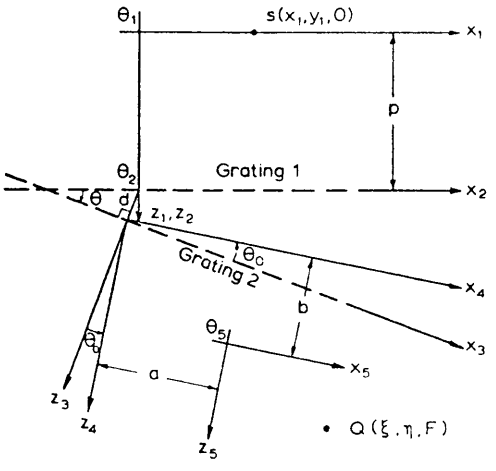


Fig. 5 Co-ordinate systems for plane wave analysis

The contribution to the field at a distant point $Q(\xi, \eta, F)$, in the output plane, from any component wave can be written as

$$dE = dA \exp - jk\left[(\alpha_5 \xi + \beta_5 \eta + \gamma_5 F) - \phi_5\right]$$

where dA depends upon the source amplitude, grating dimensions and other standard factors; α_5, β_5, γ_5 are the direction cosines of the particular component and ϕ_5 is a phase term which depends on the system configuration and grating parameters. For small α_5, β_5, we have

$$\phi_5 \approx \left[\alpha_5 x_5 + \beta_5 y_5 + (W_o - W_1 \alpha_5^2 - W_2 \beta_5^2) + \text{higher order terms in } \alpha_5, \beta_5\right],$$

where x_5, y_5, W_o, W_1, W_2 are dependent on the grating and system parameters N_1, N_2, d, θ_o θ and λ. The dimensions x_5, y_5 can be interpreted as effective displacements from the interferometer axis.

By a suitable choice of parameters, we can make $W_1 \approx W_2 \left[= z_5/2.\right]$ so that

$$\phi_5 \approx \left[\alpha_5 x_5 + \beta_5 y_5 + z_5\left(1 - \frac{\alpha_5^2}{2} - \frac{\beta_5^2}{2}\right) + W + \text{higher terms}\right]$$

where W is a constant independent of α_5, β_5. This phase term implies that the elementary plane wave can be interpreted as originating from a virtual source at x_5, y_5, z_5 and propagating with an intrinsic phase determined by W and the higher order terms. For regions far away from the source, the stationary phase method (Ref. 3) can be used to determine the amplitude at Q. Denoting the stationary point by α_{o5}, β_{o5} and restricting our argument to one dimensional folding, we obtain

$$\alpha_{o5} = f(\xi, \eta, F) \; ; \; \beta_{o5} = 0; \quad \gamma_{o5} = \sqrt{1 - \alpha_{o5}^2}$$

Thus the stationary phase solution leads to

$$E_Q = A \exp - jk\left[(\alpha_{o5} \xi + \gamma_{o5} F) - (\alpha_{o5} x_5 + \gamma_{o5} z_5) - W - \text{higher terms}\right]$$

where A is the complex amplitude. Interference between the wave from x_5, y_5, z_5 and its folded companion from $-x_5$, $-y_5$, z_5 occurs at Q. As Q varies, a fringe pattern is generated.

On substituting typical experimental values, the contribution of the higher order terms to the intensity at Q is found to be at least 3 orders of magnitude smaller than the dominant (sinusoidal) term. Thus it is evident that the aberrations introduced by the gratings are small, at least in the one-dimensional case.

The expression for x_5 is of the form

$$x_5 = \left[\left\{A + F_1(\lambda)\right\} + \left\{B + F_2(\lambda)\right\}\right] x_1$$

and describes the mapping of x_5 into the virtual object plane. The constants A, B and the functions F_1, F_2 depend on the system parameters and can be chosen to yield the maximum number of polychromatic fringes. The optimum condition

is determined by a numerical analysis.

EXPERIMENTAL OBSERVATIONS

The achromatic system was tested in a one-dimensional version for simplicity. The configuration of the interferometer is shown in Fig. 6. The source used was a photographic transparency with a tungsten halide projection lamp focussed onto it. The lamp, lens and transparency were mounted on the same plate, along with the two gratings. This plate could be rotated and displaced linearly with respect to the rest of the system.

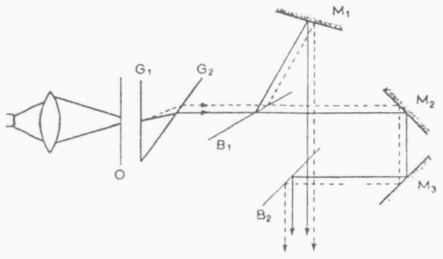


Fig. 6 The experimental interferometer arrangement

The arrangement of beam splitters and mirrors achieved the necessary 1-fold rotation by introducing an even number of reflections on one path and an odd number in the other. The careful positioning of the second beam splitter for superposition of the two virtual objects is useful in aligning the interferometer. Mirrors 2 and 3 were mounted on a common plate which could be moved with a micrometer drive so that the path lengths of the two beams were equalized to within 1 μm. The required shear, X_o, was introduced by driving the complete object plate perpendicular to the interferometer axis. This ensured that both virtual objects were displaced by the same distance from the axis but in opposite directions, thereby maintaining overall alignment.

When an object transparency, in the form of a one-dimensional binary grating, 10 lines/m.m., mark-space ratio 4:1, is illuminated in the interferometer, the output pattern is as shown in Fig. 7.

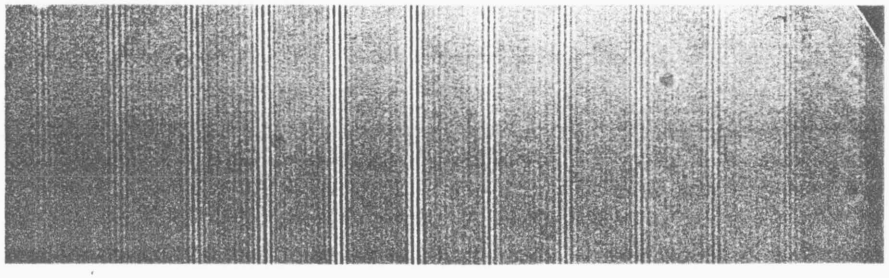


Fig. 7 Output fringe pattern for an input binary grating of 10 lines/m.m., mark-space ratio 4:1

As expected, the transform of the input varies the visibility of a set of carrier fringes. The positions of the bands of fringes are determined by the grating periodicity, while their width is determined by the overall aperture. The maximum fringe visibility in each band is governed by the mark-space ratio. The extent of the transform is not in fact limited by loss of achromaticity, but by the exit aperture of the final beam splitter. The range over which fringes can be observed in the present system shows that approximately 250 black and white fringes can be formed from a white source. In other examples, up to 450 have been obtained.

For the purposes of demonstration, the object width was limited to 1.5 m.m., so that several fringes could be clearly seen in each patch. However, if the object size is increased to 8 m.m. a comparable number of fringes can be obtained, showing that achromaticity is maintained for extended objects. The limit of 8 m.m. was imposed purely by the size of the lamp filament.

It should be possible to use the output transparency subsequently as the input to the interferometer, hence reforming the original input directly. However, the normal output contains a large uniform background component, which, on a second transit through the interferometer, would swamp the useful signal. This difficulty could be overcome using electronic techniques in conjunction with a T.V. camera tube as the detecting element.

CONCLUSION

We have shown that a one-dimensional wavefront folding interferometer can be made essentially achromatic for extended objects. The extension of the technique to two-dimensions should not present any further difficulty. Since the interferometer output is a display of the cosine Fourier transform of the object as an intensity distribution, it can be recorded with a camera tube, the video signal then being available for electronic processing. This feature makes the system potentially attractive as a hybrid optical-electronic real-time processor.

REFERENCES

1. Mertz, L., Transformations in Optics, Wiley, New York, p.111, 1965.

2. Stroke, G.W., Restrick, R.C., App. Phys. Lett., 7, 229, (1965).

3. Dainty, J.C., Scaddon, R.J., Mon. Nat. R. Astr. Soc., 167, 69, (1974).

4. Breckinridge, J.B., App. Opt., 13, 2760, (1974).

5. Jones, D.S., Kline, M., J. Maths. Phys., 37, 1, (1958).

Reprinted from

Applications of

Holography and Optical Data Processing

Proceedings of the International Conference

Jerusalem, August 23-26, 1976.

Editors Marom, Friesem and Wiener-Auneav

Pengamon Press 1976.

ACHROMATIC FRINGES FROM EXTENDED SOURCES

G. R. Wloch, S. M. Bose and J. R. Cozens

Electrical Engineering Department, Imperial College, London

ABSTRACT

A wavelength compensated wavefront folding interferometer is described which produces over 400 achromatic fringes from extended polychromatic objects.

INTRODUCTION

The inherently two-dimensional nature of optical wavefronts leads to the expectation that the potential of parallel optical signal processing would prove considerable. The technique most widely developed so far involves coherent light, its power deriving from the remarkable Fourier transform property of lenses. While there are many applications for which this system is admirable, there are many others for which the requirement that the information has to be impressed onto a coherent carrier is very restrictive. Furthermore, coherent processing relates the transforms of complex amplitudes, so that the Fourier transform of a signal cannot be recorded directly with a detector.

If we consider the form in which two-dimensional signals are likely to occur in practice, among the most obvious examples might be photographic transparencies, optical images (i.e. intensity distributions) and cathode ray tube displays. Only the first of these is readily imposed onto a coherent beam.

One useful development might therefore be towards parallel processing systems that can use direct signals as inputs. We have chosen to aim for a system that could use, for example, a C.R.T. screen display directly as an input signal. The system must therefore be able to operate with spatially incoherent, polychromatic signals in real time.

The system we are investigating is based on the wavefront folding interferometer (Refs. 1-4). In this instrument, the light from a two-dimensional incoherent source is amplitude divided, with one beam then being subjected to a relative two-fold rotation and shear. Thus the intensity g (x,y) at a point in the original source will become $g(x-X_o, y-Y_o)$ and $g(-x-X_o, -y-Y_o)$ in the two virtual sources for shears of $\pm (X_o, Y_o)$, respectively. Subsequent interference between the beams is effected in the plane of a screen or detector some distance from the source, Fig. 1. If we represent the source as an array of mutually independent point sources, then only corresponding points in (a b c d) and (a'b'c'd') can interfere in the output (ξ, η) plane. The complementary sources at (x,y) and (-x,-y) will produce a sinusoidal intensity distribution in the output plane. These fringe patterns for each pair of points will add in intensity, leading to the intensity distribution in the

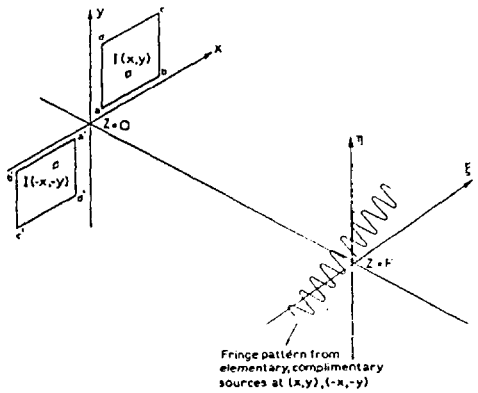


Fig. 1 Intensity distributions in the input and output planes of a wavefront folding interferometer

output plane (for large F)

$$I^1(\xi,\eta) = \frac{2}{(\lambda F)^2} \int_0^{\infty} I(x,y) \left\{1 + \cos \frac{2k}{F} \left[\xi x + \eta y\right]\right\} dxdy \qquad (1)$$

where $I(x,y) = g(x-X_o, y-Y_o)$. Thus the output contains a spatial representation of the cosine Fourier transform of the source intensity distribution which can be filtered, and also directly recorded.

We can obtain further insight into the nature of the output if eqn.(1) is written in the form:

$$I^1(\xi,\eta) = \frac{2}{(\lambda F)^2} \left[G_e\left(\frac{2\xi}{\lambda F}, \frac{2\eta}{\lambda F}\right) \cos \left\{2\pi \left(\frac{2\xi}{\lambda F} X_o + \frac{2\eta}{\lambda F} Y_o\right)\right.\right.$$

$$\left.\left. - j\, G_o\left(\frac{2\xi}{\lambda F}, \frac{2\eta}{\lambda F}\right) \sin \left\{2\pi \left(\frac{2\xi}{\lambda F} X_o + \frac{2\eta}{\lambda F} Y_o\right)\right\}\right] + C_1$$

where G_e and G_o are the Fourier transforms of the even (g_e) and odd (g_o) parts respectively of the function $g(x,y)$ and C_1 is a constant. Thus it is evident that the output consists of three parts:

(a) a cosine fringe pattern modulated by the transform of the even part of the input signal,

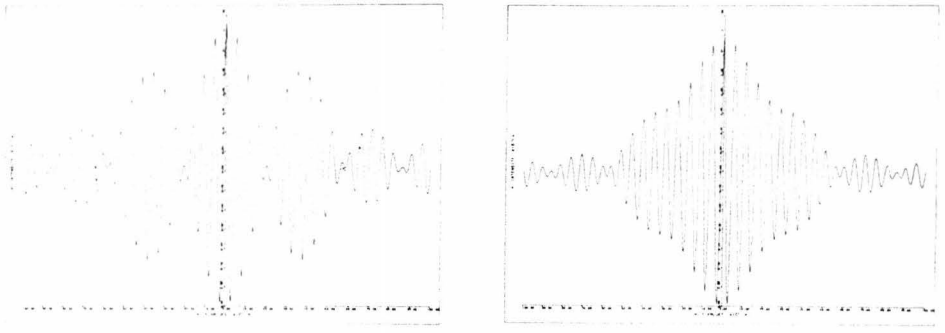
(b) a sine fringe pattern modulated by the transform of the odd part of the input,

(c) a constant intensity which allows the sinusoidal terms to be directly observed.

Since the transform of g_e is real and the transform of g_o is imaginary, complete information about the complex transform is available in the form of a real signal in the output plane. This is illustrated, Fig. 2, by means of the calculated one dimensional outputs for input functions.

$$g(x) = \text{rect}(\tfrac{x}{b}).\ (1 + \cos \tfrac{x}{d}) \qquad \text{(purely even)}$$

$$g(x) = \text{rect}(\tfrac{x}{b}).\ (1 + \sin \tfrac{x}{d}) \qquad \text{(odd + even)}$$



(a) (b)

Fig. 2 Plots of output intensity for
(a) $g(x) = \text{rect}(\frac{x}{b})\ (1 + \cos \frac{x}{d})$
(b) $g(x) = \text{rect}(\frac{x}{b})\ (1 + \sin \frac{x}{d})$
both for b = d.

The ability of the interferometer to distinguish and display both transforms is clear.

A calculated two-dimensional fringe pattern for a rectangular slot is also shown, Fig. 3, where the phase information in the transform is apparent in the fringe displacements.

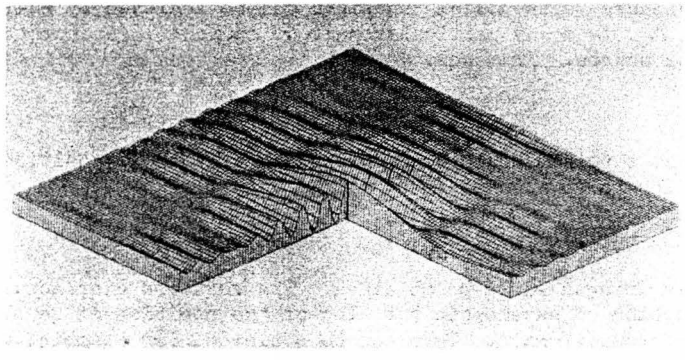


Fig. 3 2-dimensional fringe pattern for a rectangular slot input

It should be noted that the input to the interferometer is in the form of an intensity distribution and hence g(x,y) can only be real and positive. It is therefore not possible to have a truly odd input function and the output intensity distribution always has its maximum at the centre.

Bandwidth Restriction

One important restriction limits useful application of the instrument in the form so far described. As we move from the centre of the output plane, the visibility of the fringes from each pair of point sources will fall, due to their finite temporal coherence. Since regions further from the centre display the higher spatial frequencies of the source, some trade-off must exist between the optical bandwidth, $\lambda/\Delta\lambda$, and the maximum detectable spatial frequency, ω_s. From eqn.(1) we can simply show that the displacement from the origin $R(\omega_s)$ representing ω_s is

$$R(\omega_s) = \frac{F}{2k} \cdot \omega_s \qquad (2)$$

whereas the maximum range, R(max), for fringe formation by a pair of sources at $\pm(x,y)$ is

$$R_{max} = \frac{\pi F}{k(x^2+y^2)^{1/2}} \quad \frac{\lambda}{\Delta\lambda} \qquad (3)$$

Combining eqn.(2) and (3),

$$\omega_s = 2\pi \frac{1}{(x^2+y^2)^{1/2}} \frac{\lambda}{\Delta\lambda} .$$

For fairly wide bandwidths, say $\lambda/\Delta\lambda \approx 10$, and spatial frequencies up to, say, 10 lines/m.m., source dimensions of only 1 m.m. can be accommodated. Alternatively we could use a large source and increase $\lambda/\Delta\lambda$ with a filter. Thus a T.V. screen, dimension say D_o, would have spatial frequencies up to about $2\pi\,300/D$, which for $\lambda = 6000$ A, would require $\Delta\lambda \approx 20$ A. These restrictions would be clearly unacceptable in many cases.

The Achromatic Interferometer

A considerable increase in useful bandwidth can be achieved by making the fringe period, Λ, from any particular pair of sources independent for a one dimensional source. From eqn.(1)

$$\Lambda = \frac{F\lambda}{2x} ,$$

thus if all source points could be displaced so that their positions linearly shifted according to wavelength as

$$\delta x = x \cdot \frac{\delta\lambda}{\lambda} ,$$

then achromatic fringes would be obtained. The required displacement is thus not only a function of wavelength but also of the position of an object point from the interferometer axis. This can be achieved, for example, with a pair of mutually inclined transmission gratings, as shown in Fig. 4. A suitable choice of N_1, N_2 and θ leads to the formation of a large number of fringes.

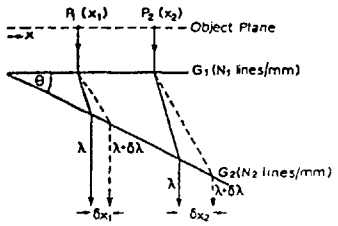


Fig. 4 The passage of two rays through a pair of inclined gratings

The ray representation indicated in Fig. 4 allows a simple illustrative analysis to be performed, but contains an uncertain level of approximation. A more rigorous treatment is outlined below.

A point source in the input plane is represented by a uniform spectrum of plane waves. The passage of each component wave through the gratings can then be described accurately by the grating equation. In Fig. 5, the component waves from a source $S(x_1, y_1, 0)$ propagate through gratings 1 and 2; after diffraction these components are modified in amplitude and phase and are expressed in the co-ordinate system (x_3, y_3, z_3). This system is then rotated by θ_0 and translated by 'a' and 'b' to the final system (x_5, y_5, z_5), centre O_5. The z_5-axis is the interferometer axis and the plane $z_5=0$ is the folding plane.

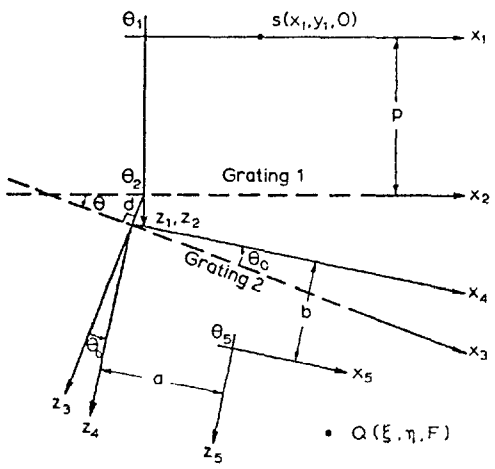


Fig. 5 Co-ordinate systems for plane wave analysis

The contribution to the field at a distant point $Q(\xi, \eta, F)$, in the output plane, from any component wave can be written as

$$dE = dA \exp - jk\left[(\alpha_5 \xi + \beta_5 \eta + \gamma_5 F) - \phi_5\right]$$

where dA depends upon the source amplitude, grating dimensions and other standard factors; α_5, β_5, γ_5 are the direction cosines of the particular component and ϕ_5 is a phase term which depends on the system configuration and grating parameters. For small α_5, β_5, we have

$$\phi_5 \approx \left[\alpha_5 x_5 + \beta_5 y_5 + (W_o - W_1 \alpha_5^2 - W_2 \beta_5^2) + \text{higher order terms in } \alpha_5, \beta_5\right],$$

where x_5, y_5, W_o, W_1, W_2 are dependent on the grating and system parameters N_1, N_2, d, θ_o θ and λ. The dimensions x_5, y_5 can be interpreted as effective displacements from the interferometer axis.

By a suitable choice of parameters, we can make $W_1 \approx W_2 \left[= z_5/2,\right]$ so that

$$\phi_5 \approx \left[\alpha_5 x_5 + \beta_5 y_5 + z_5\left(1 - \frac{\alpha_5^2}{2} - \frac{\beta_5^2}{2}\right) + W + \text{higher terms}\right]$$

where W is a constant independent of α_5, β_5. This phase term implies that the elementary plane wave can be interpreted as originating from a virtual source at x_5, y_5, z_5 and propagating with an intrinsic phase determined by W and the higher order terms. For regions far away from the source, the stationary phase method (Ref. 3) can be used to determine the amplitude at Q. Denoting the stationary point by α_{o5}, β_{o5} and restricting our argument to one dimensional folding, we obtain

$$\alpha_{o5} = f(\xi, \eta, F) \; ; \; \beta_{o5} = 0; \quad \gamma_{o5} = \sqrt{1 - \alpha_{o5}^2}$$

Thus the stationary phase solution leads to

$$E_Q = A \exp - jk\left[(\alpha_{o5} \xi + \gamma_{o5} F) - (\alpha_{o5} x_5 + \gamma_{o5} z_5) - W - \text{higher terms}\right]$$

where A is the complex amplitude. Interference between the wave from x_5, y_5, z_5 and its folded companion from $-x_5$, $-y_5$, z_5 occurs at Q. As Q varies, a fringe pattern is generated.

On substituting typical experimental values, the contribution of the higher order terms to the intensity at Q is found to be at least 3 orders of magnitude smaller than the dominant (sinusoidal) term. Thus it is evident that the aberrations introduced by the gratings are small, at least in the one-dimensional case.

The expression for x_5 is of the form

$$x_5 = \left[\left\{A + F_1(\lambda)\right\} + \left\{B + F_2(\lambda)\right\}\right] x_1$$

and describes the mapping of x_5 into the virtual object plane. The constants A, B and the functions F_1, F_2 depend on the system parameters and can be chosen to yield the maximum number of polychromatic fringes. The optimum condition

is determined by a numerical analysis.

EXPERIMENTAL OBSERVATIONS

The achromatic system was tested in a one-dimensional version for simplicity. The configuration of the interferometer is shown in Fig. 6. The source used was a photographic transparency with a tungsten halide projection lamp focussed onto it. The lamp, lens and transparency were mounted on the same plate, along with the two gratings. This plate could be rotated and displaced linearly with respect to the rest of the system.

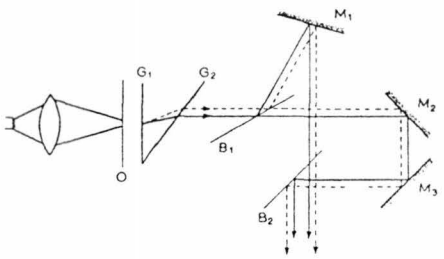


Fig. 6 The experimental interferometer arrangement

The arrangement of beam splitters and mirrors achieved the necessary 1-fold rotation by introducing an even number of reflections on one path and an odd number in the other. The careful positioning of the second beam splitter for superposition of the two virtual objects is useful in aligning the interferometer. Mirrors 2 and 3 were mounted on a common plate which could be moved with a micrometer drive so that the path lengths of the two beams were equalized to within 1 μm. The required shear, X_o, was introduced by driving the complete object plate perpendicular to the interferometer axis. This ensured that both virtual objects were displaced by the same distance from the axis but in opposite directions, thereby maintaining overall alignment.

When an object transparency, in the form of a one-dimensional binary grating, 10 lines/m.m., mark-space ratio 4:1, is illuminated in the interferometer, the output pattern is as shown in Fig. 7.

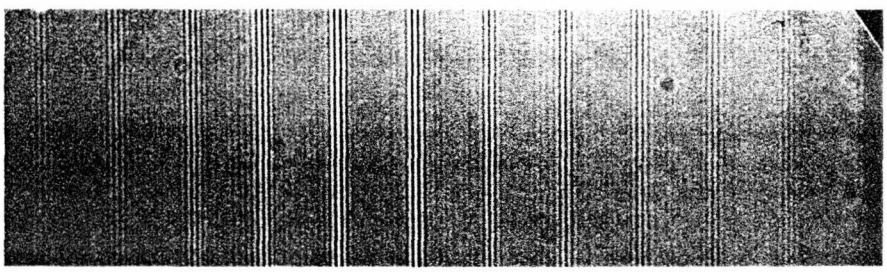


Fig. 7 Output fringe pattern for an input binary grating of 10 lines/m.m., mark-space ratio 4:1

As expected, the transform of the input varies the visibility of a set of carrier fringes. The positions of the bands of fringes are determined by the grating periodicity, while their width is determined by the overall aperture. The maximum fringe visibility in each band is governed by the mark-space ratio. The extent of the transform is not in fact limited by loss of achromaticity, but by the exit aperture of the final beam splitter. The range over which fringes can be observed in the present system shows that approximately 250 black and white fringes can be formed from a white source. In other examples, up to 450 have been obtained.

For the purposes of demonstration, the object width was limited to 1.5 m.m., so that several fringes could be clearly seen in each patch. However, if the object size is increased to 8 m.m. a comparable number of fringes can be obtained, showing that achromaticity is maintained for extended objects. The limit of 8 m.m. was imposed purely by the size of the lamp filament.

It should be possible to use the output transparency subsequently as the input to the interferometer, hence reforming the original input directly. However, the normal output contains a large uniform background component, which, on a second transit through the interferometer, would swamp the useful signal. This difficulty could be overcome using electronic techniques in conjunction with a T.V. camera tube as the detecting element.

CONCLUSION

We have shown that a one-dimensional wavefront folding interferometer can be made essentially achromatic for extended objects. The extension of the technique to two-dimensions should not present any further difficulty. Since the interferometer output is a display of the cosine Fourier transform of the object as an intensity distribution, it can be recorded with a camera tube, the video signal then being available for electronic processing. This feature makes the system potentially attractive as a hybrid optical-electronic real-time processor.

REFERENCES

1. Mertz, L., Transformations in Optics, Wiley, New York, p.111, 1965.

2. Stroke, G.W., Restrick, R.C., App. Phys. Lett., 7, 229, (1965).

3. Dainty, J.C., Scaddon, R.J., Mon. Nat. R. Astr. Soc., 167, 69, (1974).

4. Breckinridge, J.B., App. Opt., 13, 2760, (1974).

5. Jones, D.S., Kline, M., J. Maths. Phys., 37, 1, (1958).