University of London

Applied Optics Section
Department of Physics
Imperial College of Science and Technology

"Some extrinsic factors affecting the performance of an image-forming optical system."

A thesis submitted for the
degree of Doctor of Philosophy
by
Roland Vincent Shack

May 1965

TABLE OF CONTENTS

ABSTRACT

In Part I a general two-dimensional expression for the transfer function of the photographic exposure image is derived, including as factors the optical transfer function, the film transfer function, the shutter function, and the motion of the image. The effect of shutter operation is discussed and it is shown that a focal-plane shutter acts as an independent agent with its own effective transfer function, whereas a between-the-lens shutter is not independent but acts as an apodizing agent. The effect of image motion is then discussed. The concept of an equivalent spread function for image motion is developed, and uniform linear image motion and simple harmonic image motion are discussed in detail. An approximation for small degradations is then obtained in which all possible combinations of uniform linear motion and simple harmonic motion are contained.

Part II deals with the classical properties of an

image formed by an optical system looking through a turbulent medium, namely scintillation, agitation, and blur. First a simple but adequate statistical model is obtained describing the pupil function in the presence of random wavefront disturbances in both amplitude and phase. A measure of scintillation is then obtained in the form of a normalized, weighted integral in frequency space containing all the pertinent properties of the optical system. A similar measure of agitation is then developed, and finally a description of the blur is obtained in the form of an equivalent transfer function for the turbulent medium in which the contribution from amplitude variations and the contribution from phase variations are independent.

GENERAL INTRODUCTION

The relatively recent realization that an optical image-forming system can be treated as a communication channel has led to a remarkably fruitful application of many of the ideas developed in the field of communication engineering.[1] Particular emphasis has been placed on the application of Fourier analysis, although modern statistical theory and information theory have also begun to be applied.[2,3] The use of these tools has extended and deepened our understanding of the image-forming process, and is making the conception, design, and evaluation of image-forming systems less of an art and more of a science.

The classical communication channel[4] consists of three basic operations: origination, transmission, and reception. In the origination stage a carrier (transmissible form of energy) is modified in one way or another so as to conform to the message to be sent. This modification may be analogous in form to the message

---

(1) A superscript in parentheses denotes a reference in the Bibliography.

(e.g. amplitude variations of a radio frequency carrier which follow the form of the sound pressure variations sensed by a microphone), or may be in coded form, for example, the dots and dashes of Morse code. The carrier must be transmitted by an appropriate medium in the course of which the signal may be altered or degraded, and spurious (unwanted) signals and noise may be picked up. The transmitted signal is received by a mechanism capable of translating the modification of the carrier into a form which can be comprehended as the received message.

An image-forming optical system can be treated in an analogous fashion. In the origination stage the carrier is the illumination incident on the object. This carrier is modified by the geometric structure and the reflectance (or transmittance) variations of the object. If the object is self-luminous it can be assumed to be both source and modifier. The transmission channel is the optical path between the object and the detector, this path containing an optical system which forms a real image of the object at the detector. In a visual system the

retina of the eye is the detector. In a photoelectric system a photosensitive element behind an aperture of some sort is the detector. In a photographic system the photographic emulsion is the detector which converts the received illumination into a developable latent image, and subsequent development converts this latent image into a permanent silver image.

We are still in the process of understanding the various phenomena which are involved in each of these systems from this new viewpoint. Unfortunately a large part of both the visual and the photographic process is non-linear, and the most powerful tools which are available, in particular Fourier analysis, depend for their validity on the linearity of the system. However, it is possible to consider the linear and non-linear parts of the system separately*, and at least discover as much as

---

* The non-linearity can to some degree be circumvented by restricting the range of the variables sufficiently so that the system is adequately close to being linear within this small range.

we can about the linear parts.

The performance of any image-forming system is determined primarily by the performance of the optical system which it contains, and the performance of the detector. These are the intrinsic factors. There are, however, for some systems other factors which can have a significant effect on the performance. Since these other factors may or may not be significant, we shall call them extrinsic factors. The aim of this thesis is the investigation of several of these extrinsic factors.

In part I we shall investigate two important time-dependent factors affecting the performance of a photographic system. These are the effects of the shutter in obscuring part of the pupil of the optical system during the finite opening and closing time it requires, and the effects of motion of the image relative to the film during the exposure time. In part II we shall investigate the effects of a turbulent medium between the object and the optical system on the properties of point images. This part covers the classical phenomena of scintillation,

agitation (image motion), and blur which have been observed for many years by astronomers.

PART I. INTRODUCTION

The spatial distribution of the effective exposure obtained in creating a photographic image is not, in general, identical to that of the instantaneous optical image being recorded. Apart from the blurring resulting from the diffusion of light in the emulsion, effectively treated elsewhere[5], there are, in general, time-dependent variations in the optical image taking place during the exposure time. Two classes of phenomena which are considered in this thesis are: (1) the modification of the image structure resulting from the action of the shutter, which causes the shape and size of the pupil to vary as it opens and closes, and (2) motion of the image.

For the purposes of this thesis we do not wish to concern ourselves with the photographic process as a whole but only with the formation of the latent image and not even with all of that. The latent image consists of a distribution of discrete grains of silver halide made

developable by the action of light. This discrete structure can be related to the practically continuous structure of the optical image by postulating a continuous statistical latent image, the value of which at any point represents the probability that a developable silver halide grain will occur at that point after exposure. This probability will depend on a variety of factors, for example, the distribution of grain sensitivities and the population density of the grains, but the major factor which determines the spatial structure of the image is the amount of light received and integrated over the exposure time.

Thus the statistical latent image is related to the optical image producing it by the process of time integration. The time-integrated optical image is itself a kind of hypothetical image, which is in units of energy distribution rather than power distribution over the image surface. It is commonly called the exposure image.

The exposure image is an intermediate stage between the optical image and the statistical latent image. The

statistical latent image can be obtained from the exposure image through a nonlinear transfer, taking into account the grain distribution factors, etc. However, if reciprocity can be assumed to hold, the exposure image can be obtained from the optical image by a linear process of integration. This means that the formation of the exposure image is the logical termination of the linear input stage of the photographic process, after which nonlinear procedures must be employed. Throughout the linear stage, Fourier methods can be applied and the image-forming capability of that part of the photographic system can be characterized by its transfer function, or, alternatively, its spread function.

Throughout the thesis the transfer functions and spread functions discussed are functions of two spatial dimensions. To avoid unnecessary complication in the symbology, the same spatial coordinates u,v are used to represent object and image space. It is understood that the necessary, and simple, geometrical transformation has taken place. Also, to avoid using a great

many different symbols, the author has chosen to use the tilde over those functions and variables which are in the spatial frequency domain, but maintaining the symbol which is used in the spatial domain. Thus u and v represent spatial coordinates whereas $\tilde{u}$ and $\tilde{v}$ represent the corresponding spatial frequency coordinates, and $\tilde{F}$ is the Fourier transform of F. The tilde, which is seldom used in optical symbology, suggests a sine wave.

To simplify the appearance of several of the mathematical expressions, $\mathcal{S}(x)$ will be used to represent the frequently occurring $(\sin x)/x$.

In chapter I-1 a general expression for the transfer function of the exposure image is derived. Chapter I-2 discusses the possible effects of a focal-plane and a between-the-lens shutter on the transfer function in the absence of image motion. Chapter I-3 discusses the effect of image motion. The concept of the image-motion spread function is developed and uniform linear and simple harmonic motion are discussed in detail, the latter especially

for exposure times which are of the same order of magnitude as the period of vibration. An approximation appropriate for small degradations is obtained, and combined motions are considered.

CHAPTER I-1. THE TRANSFER FUNCTION OF THE EXPOSURE IMAGE

In this chapter we shall derive a general expression for the transfer function of the exposure image. In order to do this we must examine the process of formation of the exposure image, starting with the object.

The Fourier Description of the Object

Let $N(u,v)$ represent the radiance distribution over the object plane, limited to a finite field of area A. It will be advantageous to factor out the mean value: thus

$$N(u,v) = \bar{N}\, n(u,v) \qquad \text{I-1.1}$$

where

$$\bar{N} = A^{-1} \iint_A N(u,v)\, du\, dv = A^{-1} \int_{-\infty}^{\infty} \int N(u,v)\, du\, dv. \qquad \text{I-1.2}$$

The total radiant intensity of the object is given by $A\bar{N}$. The purpose of the factoring is the separation of the object function into two factors, one without spatial structure but carrying the dimensional units, and the other providing the structural information in a dimensionless form.

The spatial Fourier transform of the object, which we shall call its structure spectrum, is given by

$$\tilde{N}(\tilde{u},\tilde{v}) = \iint_{-\infty}^{\infty} N(u,v)\exp[-2\pi i(\tilde{u}u+\tilde{v}v)]\,du\,dv, \qquad I-1,3$$

Here, too, we wish to separate the function into a dimensional factor and a structural factor. In order to do this, we define the normalized structure spectrum by

$$\tilde{n}(\tilde{u},\tilde{v}) \equiv \frac{\tilde{N}(\tilde{u},\tilde{v})}{\tilde{N}(0,0)} = \frac{\bar{N}\iint_{-\infty}^{\infty} n(u,v)\exp[-2\pi i(\tilde{u}u+\tilde{v}v)]\,du\,dv}{\bar{N}\iint_{-\infty}^{\infty} n(u,v)\,du\,dv}$$

$$= A^{-1}\iint_{-\infty}^{\infty} n(u,v)\exp[-2\pi i(\tilde{u}u+\tilde{v}v)]\,du\,dv, \qquad I-1,4$$

which is seen to be the normalized Fourier transform of the structural factor of the object distribution. The dimensional structure spectrum of Eq. I-1.3 is then given by

$$\tilde{N}(\tilde{u},\tilde{v}) = A\bar{N}\,\tilde{n}(\tilde{u},\tilde{v}), \qquad \text{I-1.5}$$

from which it is seen that the dimensional factor of the structure spectrum is simply the total radiant intensity of the object.

The Fourier Description of the Optical Image

In the same fashion, assuming the effective size of the spread function is much smaller than A, the irradiance in the image plane is given by

$$H(u,v) = \bar{H}\, h(u,v), \qquad \text{I-1.6}$$

and the image structure spectrum by

$$\tilde{H}(\tilde{u},\tilde{v}) = A\bar{H}\,\tilde{h}(\tilde{u},\tilde{v}). \qquad \text{I-1.7}$$

The average irradiance in the image is related to the average radiance of the object by

$$\bar{H} = K\bar{N}, \qquad \text{I-1.8}$$

where K represents the usual photometric factor, and the normalized image structure spectrum is related to the corresponding object spectrum by

$$\tilde{h}(\tilde{u},\tilde{v}) = \tilde{\Phi}_o(\tilde{u},\tilde{v})\,\tilde{n}(\tilde{u},\tilde{v}), \qquad \text{I-1.9}$$

where $\tilde{\Phi}_o$ is the transfer function of the optical system.

The Fourier Description of the Exposure Image

The exposure image is obtained by a time integration of the optical image in the emulsion. Thus,

$$E(u,v) = \int_{-\infty}^{\infty} S(t)\, H(u,v)\, dt$$

$$= \int_{-\infty}^{\infty} S(t) \left\{ \iint_{-\infty}^{\infty} \tilde{H}\, \exp\left[2\pi i(\tilde{u}u + \tilde{v}v)\right] d\tilde{u}\, d\tilde{v} \right\} dt, \qquad \text{I-1.10}$$

where $S(t)$ is the shutter function in the usual sense, and H is retained under the integral sign because it may also be time-dependent. If we assume that the image is moving,

$$H(u,v) = H(u,v;t) = H\left[u - \underline{u}(t),\ v - \underline{v}(t)\right], \qquad \text{I-1.11}$$

where $\underline{u}(t)$ and $\underline{v}(t)$ are parametric expressions describing the path of motion. Then, on interchanging the order of integration, Eq. I-1.10 becomes

$$E(u,v) = \iint_{-\infty}^{\infty} \tilde{E}(\tilde{u},\tilde{v}) \exp\left[2\pi i(\tilde{u}u + \tilde{v}v)\right] d\tilde{u}\, d\tilde{v}, \qquad \text{I-1.12}$$

where the structure spectrum of the exposure image is

$$\tilde{E}(\tilde{u},\tilde{v}) = \int_{-\infty}^{\infty} S\,\tilde{H}\,\exp[-2\pi i(\tilde{u}\underline{u} + \tilde{v}\underline{v})]\,dt. \qquad \text{I-1.13}$$

Substituting for $\tilde{H}$ from eq. I-1.7 and using eq. I-1.8 we obtain

$$\tilde{E}(\tilde{u},\tilde{v}) = AK\bar{N}t_e\left\{t_e^{-1}\int_{-\infty}^{\infty} S\,\tilde{\Phi}_o \exp[-2\pi i(\tilde{u}\underline{u} + \tilde{v}\underline{v})]\,dt\right\}\tilde{n}(\tilde{u},\tilde{v}), \qquad \text{I-1.14}$$

where $t_e = \int_{-\infty}^{\infty} S\,dt$ is the effective exposure time. Comparing Eq. I-1.14 with Eqs. I-1.5 and I-1.7 shows that $\tilde{E}$ can be factored in the same way as $\tilde{N}$ and $\tilde{H}$; that is,

$$\tilde{E}(\tilde{u},\tilde{v}) = A\bar{E}\,\tilde{e}(\tilde{u},\tilde{v}), \qquad \text{I-1.15}$$

where

$$\bar{E} = K\bar{N}t_e. \qquad \text{I-1.16}$$

which is the usual sensitometric expression, and

$$\tilde{e}(\tilde{u},\tilde{v}) = \tilde{\Phi}_E(\tilde{u},\tilde{v})\,\tilde{n}(\tilde{u},\tilde{v}) \qquad \text{I-1.17}$$

where

$$\tilde{\Phi}_E(\tilde{u},\tilde{v}) = t_e^{-1}\int_{-\infty}^{\infty} S\,\tilde{\Phi}_o \exp[-2\pi i(\tilde{u}\underline{u} + \tilde{v}\underline{v})]\,dt \qquad \text{I-1.18}$$

is the effective transfer function for the exposure image.

The optical transfer function $\tilde{\tilde{\Phi}}_o$ in the integrand of Eq. I-1.18 is the optical transfer function in the emulsion, i.e., it includes the effect of light scatter in the emulsion. Thus

$$\tilde{\tilde{\Phi}}_o = \tilde{\tilde{\Phi}}_L \tilde{\tilde{\Phi}}_F \qquad \text{I-1.19}$$

where $\tilde{\tilde{\Phi}}_L$ is the transfer function for the optical system per se, and $\tilde{\tilde{\Phi}}_F$ is the transfer function for the scattering in the emulsion. The latter is independent of time and can be taken outside the integral. The former, as we shall see in the next chapter, is not, in general, independent of time and must be kept inside. The exposure-image transfer function then becomes

$$\tilde{\tilde{\Phi}}_E = \tilde{\tilde{\Phi}}_F t_e^{-1} \int_{-\infty}^{\infty} S \tilde{\tilde{\Phi}}_L \exp[-2\pi i(\tilde{u}\underline{u} + \tilde{v}\underline{v})]\, dt \qquad \text{I-1.20}$$

CHAPTER I-2. THE EFFECT OF THE SHUTTER

In this section we shall assume the absence of image motion. The transfer function for the exposure image then becomes

$$\tilde{\Phi}_E = \tilde{\Phi}_F \, t_e^{-1} \int_{-\infty}^{\infty} S \, \tilde{\Phi}_L \, dt \,, \qquad \text{I-2.1}$$

If we can also assume that $\tilde{\Phi}_L$ is independent of time, then

$$\tilde{\Phi}_E = \tilde{\Phi}_L \tilde{\Phi}_F \, t_e^{-1} \int_{-\infty}^{\infty} S \, dt = \tilde{\Phi}_L \tilde{\mathcal{J}}_F \,, \qquad \text{I-2.2}$$

and the shutter has no effect.

Strictly speaking, this condition holds only when the opening and closing times of the shutter are negligibly small with respect to the effective exposure time, or when the shutter action is obtained by varying the transparency of the pupil uniformly, as with an ideal polarizing shutter. In practice, however, most shutters are mechanical devices which vary the shape of the pupil during their opening and closing times, thereby

altering the pupil function and the optical transfer function $\tilde{\mathcal{F}}_L$ which depends on it.

The two most common types of shutter are the focal-plane and the between-the-lens shutter. We shall consider both in our investigation.

The Focal-Plane Shutter

The influence of a narrow focal-plane shutter on the image has been investigated by Bechtel[6] and by Asakura[7] by calculating the effect on the spread function of diffraction by the slit. Approaching the problem from the point of view of transfer function theory, however, leads to a very simple and elegant solution, as follows.

For mechanical reasons the so-called focal-plane shutter is never actually in the focal plane, but is a considerable distance forward of it. We shall assume for our purposes that the shutter is sufficiently removed from the focal plane for its effect to be

virtually identical to that of an equivalent shutter in the pupil. Then the shutter function can be incorporated as a time-dependent modifier of the pupil function from which the optical transfer function can be calculated. Thus

$$\tilde{\Phi}_L = \tilde{\Phi}_L(\tilde{u}, \tilde{v}; t), \quad \tilde{\Phi}_L(0, 0; t) = S(t),$$

and

$$\tilde{\Phi}_E = \tilde{\Phi}_F \, t_e^{-1} \int_{-\infty}^{\infty} \tilde{\Phi}_L \, dt. \qquad \text{I - 2, 3}$$

The modified optical transfer function can be expressed as the autocorrelation integral of the modified pupil function.

Let us represent the unmodified pupil function by $f(x, y)$ and the pupil shutter function by $s(x + \dot{x}t)$, where the shutter velocity $\dot{x}$ is constant and the shutter is moving in the negative direction of x. Then the modified pupil function is given by fs and the transfer function for the exposure image is given by

$$\tilde{\Phi}_E = \tilde{\Phi}_F\, t_e^{-1} \int_{-\infty}^{\infty} \left\{ A^{-1} \iint_{-\infty}^{\infty} \left[\mathcal{A}(x + \dot{x}t + \tfrac{1}{2}\tilde{u})\, f(x + \tfrac{1}{2}\tilde{u},\, y + \tfrac{1}{2}\tilde{v}) \right] \right.$$

$$\left. \left[\mathcal{A}^*(x + \dot{x}t - \tfrac{1}{2}\tilde{u})\, f^*(x - \tfrac{1}{2}\tilde{u},\, y - \tfrac{1}{2}\tilde{v}) \right] dx\, dy \right\} dt$$

$$= \tilde{\Phi}_F A^{-1} \iint_{-\infty}^{\infty} \left[t_e^{-1} \int_{-\infty}^{\infty} \mathcal{A}(x + \dot{x}t + \tfrac{1}{2}\tilde{u})\, \mathcal{A}^*(x + \dot{x}t - \tfrac{1}{2}\tilde{u})\, dt \right]$$

$$\left[f(x + \tfrac{1}{2}\tilde{u},\, y + \tfrac{1}{2}\tilde{v})\, f^*(x - \tfrac{1}{2}\tilde{u},\, y - \tfrac{1}{2}\tilde{v}) \right] dx\, dy, \qquad \text{I-2.4}$$

where A represents the area of the pupil and the asterisks indicate complex conjugates. The time integral in the brackets is in the form of an autocorrelation integral in which both components have been shifted by an amount X but equal shifts in the same direction do not alter the value of the integral which is therefore independent of X. The integral is also independent of y and therefore may be taken outside the larger integral:

$$\tilde{\Phi}_E = \tilde{\Phi}_F \left[t_e^{-1} \int_{-\infty}^{\infty} \mathcal{A}(\dot{x}t + \tfrac{1}{2}\tilde{u})\, \mathcal{A}(\dot{x}t - \tfrac{1}{2}\tilde{u})\, dt \right]$$

$$\left[A^{-1} \iint_{-\infty}^{\infty} f(x + \tfrac{1}{2}\tilde{u},\, y + \tfrac{1}{2}\tilde{v})\, f^*(x - \tfrac{1}{2}\tilde{u},\, y - \tfrac{1}{2}\tilde{v})\, dx\, dy \right]. \qquad \text{I-2.5}$$

The expression within the second square bracket is simply the transfer function for the optical system with the shutter fully open, so the expression within the first square bracket must describe the equivalent transfer function for the focal plane shutter, and

$$\tilde{\Phi}_E = \tilde{\Phi}_L \tilde{\Phi}_F \tilde{\Phi}_S . \qquad \text{I - 2.6}$$

By changing the variable of integration, letting $x' = x' t$ the form of the integral for the shutter transfer function is given by

$$\tilde{\Phi}_S = \frac{\int_{-\infty}^{\infty} A(x' + \frac{1}{2}\tilde{u}) A^*(x' - \frac{1}{2}\tilde{u}) dx'}{\int_{-\infty}^{\infty} A(x') A^*(x') dx'} , \qquad \text{I - 2.7}$$

which is the form of a normalized spatial autocorrelation function strictly analogous to the expression for the optical transfer function. It is, however, a function of $\tilde{u}$ alone, being independent of $\tilde{v}$.

It is interesting to note the rather unexpected result that the equivalent shutter transfer function for a focal plane shutter is quite independent of the shape and size of the pupil and the aberrations of the optical system. Unfortunately, this independence only holds in the absence of image motion.

The derivation assumes the most general complex form for the pupil shutter function, subject to the restrictions that it be independent of y, convergent in x, and constant in its velocity. However, the most common form of shutter is a simple slit. If we let w represent the ratio of the shutter width projected into the pupil to the radius of the pupil aperture, then, for a simple slit shutter

$$\tilde{\Phi}_s = 1 - \frac{|\tilde{u}|}{w}, \quad 0 < |\tilde{u}| < w$$
$$= 0, \quad |\tilde{u}| > w. \qquad \text{I-2.8}$$

This is a roof-shaped function of triangular cross section, as shown in Fig. I-1.

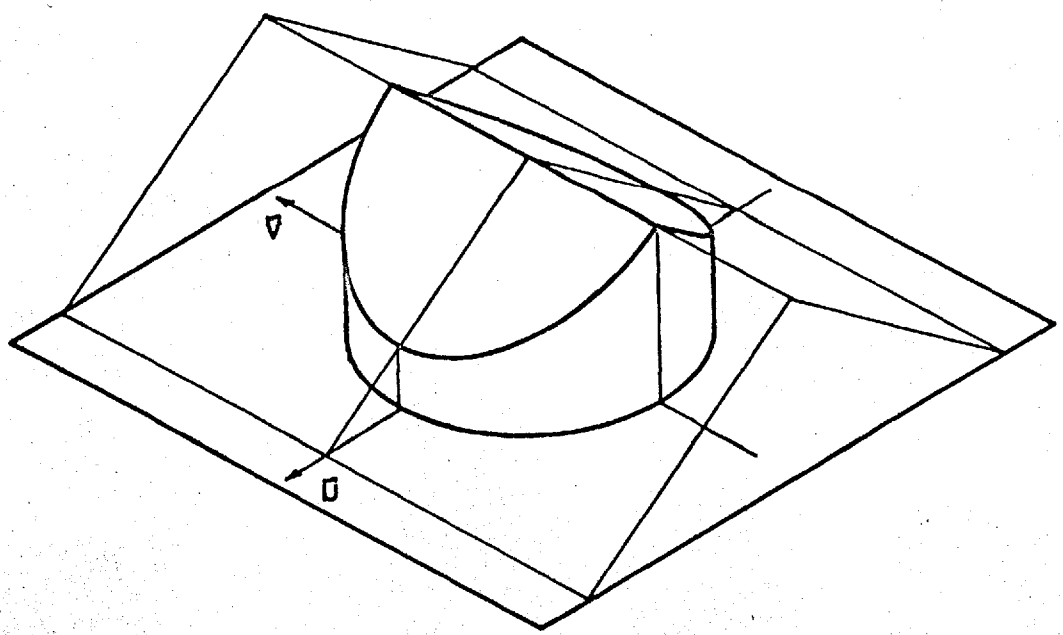


Figure I-1. Transfer Function for Focal-Plane Shutter. The cylinder in the figure represents the boundary of the frequency domain limited by optical diffraction.

The Between-the-Lens Shutter

The between-the-lens shutter was also investigated by Bechtel[6] by calculating the spread function. The following illustrates the transfer function approach.

Unfortunately, the between-the-lens shutter which operates approximately radially in the pupil does not lend itself to the elegant reduction of the previous section. It is not possible in general to separate the shutter effect from the aberrations or the diaphragm setting, and the problem is usually complicated by the shape of the shutter aperture, which usually resembles a bent star or pinwheel.

In general Eq. I-2.1 must be used where the variation of $\tilde{\Phi}_L$ with time must be determined beforehand. The shutter function S is given by the ratio of the area of the clear aperture at time t to the area of the aperture when the shutter is fully open. Both $\tilde{\Phi}_L$ and S will depend on the diaphragm setting as well.

In order to get a general idea of the effect of a between-the-lens shutter, a simplified model consisting of a circular aperture, the radius of which varies with time, will be examined. It will be assumed that its area increases linearly during its opening and that it closes linearly at the same rate.

Figure I-2 is a graphical interpretation of Eq. I-2.1 for this simplified model in the absence of aberrations. Note that as the spatial frequency of a component increases there is a corresponding delay before a contribution to that component can pass through the spatial filter, and a corresponding reduction in the effective transfer function. This occurs because the smaller aperture of the partly open shutter results in a lower maximum spatial frequency which can be transferred.

Figure I-3 is a corresponding illustration for the same model with a defect of focus of $\frac{1}{2}\lambda$. Here the same spatial frequency limit applies as in Fig. I-2 but the amplitudes are reduced. The amount of reduc-

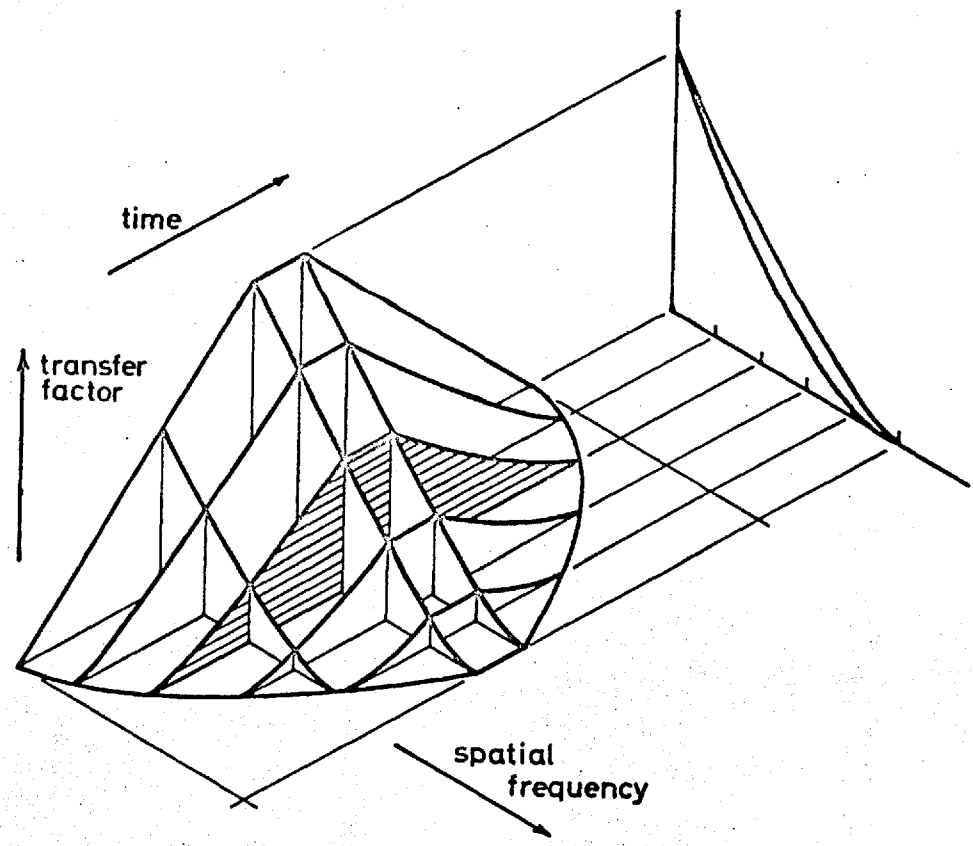


Figure I-2. Influence of a Between-the-Lens Shutter on the Transfer Function. This figure is a graphical interpretation of Equation (I-2.1). The optical system is assumed to be free of aberrations.

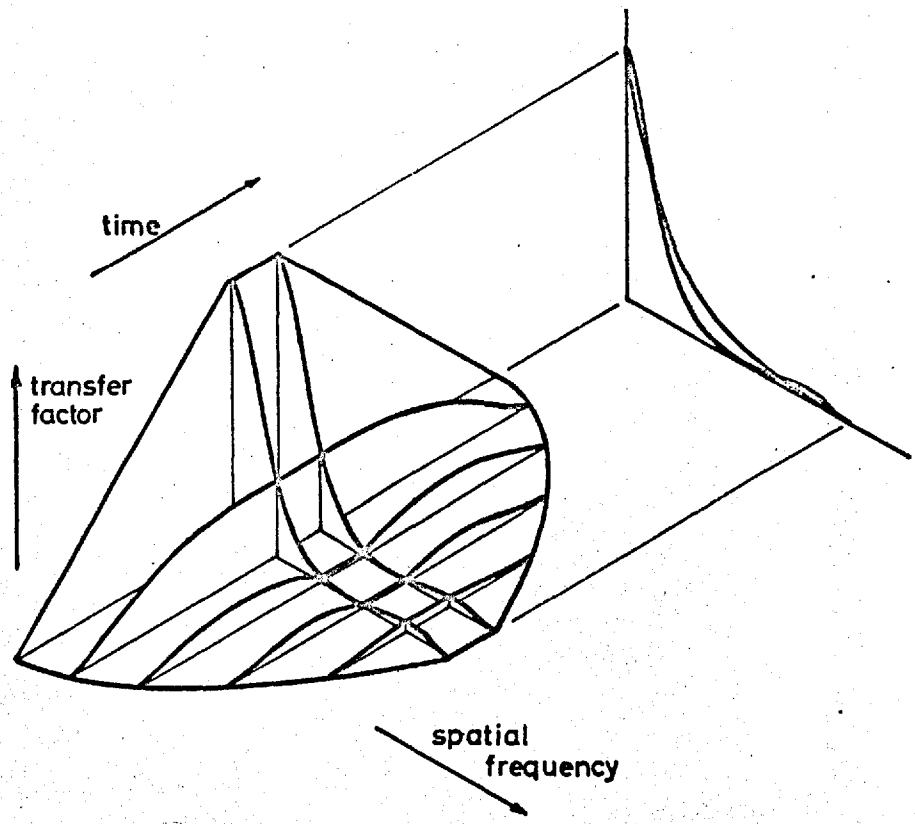


Figure I-3. Same as Figure I-2, except that the optical system is defocussed by $\lambda/2$.

tion, however, increases rapidly with the size of the shutter aperture, so that the amplitude with the shutter partly open may exceed that obtained with the shutter fully open, and the resulting transfer function value may be greater than that obtained with a "perfect" shutter which opens and closes instantaneously.

Figure I-4 shows for the same model the difference between the two extreme shutter functions, the rectangular "perfect" shutter function and the triangular "50% efficient" shutter function. In Fig. I-4a we see the difference in shape of the transfer functions and in Fig. I-4b the manner in which the transfer function values vary with focus for a particular frequency.

Taking the defect of focus as an aberration, it is clear that, although a finite shutter effect degrades an aberration-free system, it can improve an aberrated system. The improvement should be even more striking for higher order aberrations.

The character of the curves in Fig. I-4 implies that the ripples of the corresponding spread functions must

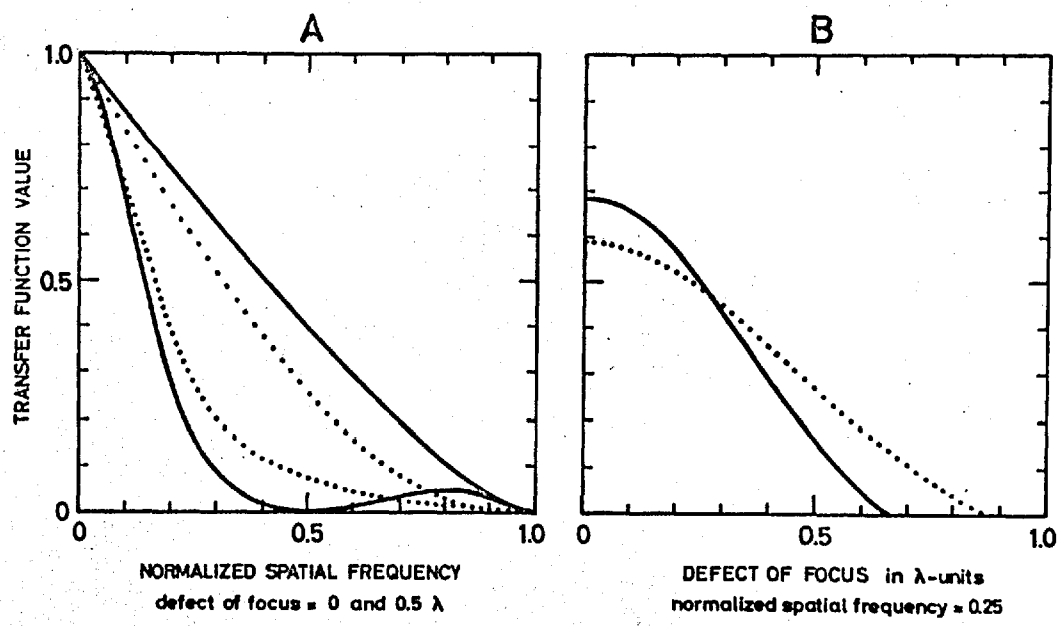


Figure I-4. Influence of a Between-the-Lens Shutter on the Transfer Function. The solid curves are for the rectangular shutter function and the dotted curves for the triangular shutter function.

be considerably reduced for the triangular shutter function relative to those of the rectangular shutter function; that is, a kind of apodization has taken place. In fact, the simplified model of shutter we employed does result in noncoherent apodization for the exposure image, and the effect of a between-the-lens shutter can be thought of in terms of such an apodization. The manner and extent of the apodization will depend heavily on the shape of the actual shutter aperture, however, and the concept of apodization in the usual sense only applies in the absence of image motion.

CHAPTER I-3. IMAGE MOTION

Earlier investigations of the effect of image motion on the quality of the image have mostly been concerned with predicting the reduction in resolving power resulting from image motion. Scott[8] reported expressions for simple forms of one-dimensional image motion. Rosenau[9] reported an expression for parabolic image motion. Chang[10] has investigated the effect of simple harmonic motion on resolving power when the exposure time is a fraction of a cycle of the motion. Paris[11] has attempted to determine the two-dimensional transfer function for combined linear and simple harmonic image motion. Paris,[11] and Hendeberg and Welander[12] have considered the modification of the transfer function for linear motion by a low-efficiency shutter, neglecting the effects described in Chapter I-2 of this thesis.

In this chapter we shall also neglect the effects described in Chapter I-2, i.e., we shall assume that the

optical transfer function is independent of time. As should be clear from Chapter I-2, this assumption is strictly valid only in comparatively rare cases. However, it is a reasonably good approximation if the effective exposure time is appreciably greater than the time it takes for the shutter to open and close, or if we are restricted to low spatial frequencies and small aberrations.

The Transfer Function and Spread Function for Image Motion

Under this assumption the expression for the transfer function for the exposure image becomes

$$\tilde{\Phi}_E = \tilde{\Phi}_L \tilde{\Phi}_F \tilde{\Phi}_M \qquad \text{I-3.1}$$

where the effective transfer function for image motion is

$$\tilde{\Phi}_M(\tilde{u},\tilde{v}) = t_e^{-1}\int_{-\infty}^{\infty} S \exp\left[-2\pi i(\tilde{u}\,\underline{u} + \tilde{v}\,\underline{v})\right] dt. \qquad \text{I-3.2}$$

Remembering that $\underline{u}$ and $\underline{v}$ are parametric expressions describing the path of image motion, we shall henceforth omit the bar.

In order to evaluate this integral at some spatial frequency $(\tilde{u}, \tilde{v})$ we can simplify our problem by rotating the reference axes through an angle ψ so that the new abscissa $\tilde{u}'$ is in the direction of $(\tilde{u}, \tilde{v})$. Then

$$\tilde{u}' = \tilde{u}\cos\psi + \tilde{v}\sin\psi, \qquad \tilde{v}' = 0$$
$$u' = u\cos\psi + v\sin\psi \qquad \text{I - 3.3}$$

and

$$\tilde{\Phi}_M(\tilde{u}', \psi) = t_e^{-1}\int_{-\infty}^{\infty} S\, exp(-2\pi i \tilde{u}'u')\, dt \qquad \text{I - 3.4}$$

If we change the variable of integration to u',

$$\tilde{\Phi}_M = \frac{\int_{-\infty}^{\infty}[S/\dot{u}']\, exp(-2\pi i \tilde{u}'u')\, du'}{\int_{-\infty}^{\infty}[S/\dot{u}']\, du'}, \qquad \text{I - 3.5}$$

which has the form of a normalized Fourier transform, and, consequently, the quantity in the square brackets represents an equivalent image-motion spread function. However, the form of this quantity depends on orientation, suggesting that it corresponds to a line-spread function rather than a point-spread function.

An equivalent point-spread function does exist, but it differs radically in character from the more familiar optical or emulsion spread functions. The latter functions can be represented by continuous, approximately bell-shaped surfaces overlaying the u,v plane, whereas the former is represented by a ribbonlike cylinder, the directrix of which is the path traced by the parametric variables $u(t)$ and $v(t)$ and the generatrix of which is a line segment perpendicular to and extending up from the u,v plane and having a length given by $S/\sqrt{u^2+v^2}$, that is, the shutter function divided by the magnitude of the velocity of the image motion along the path (see Fig. I-5).

The line-spread function in the ψ direction is

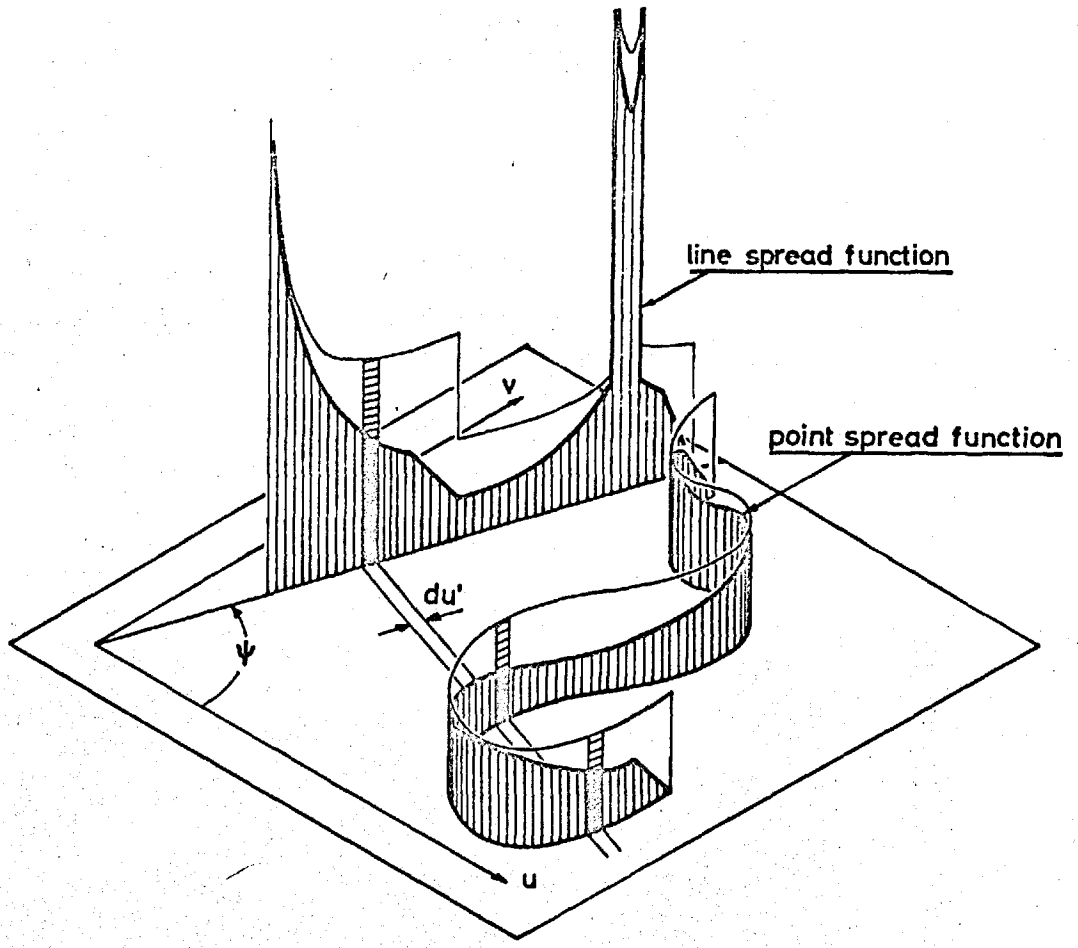


Figure I-5. Image Motion Spread Functions.

found by integrating this point-spread function in the perpendicular direction. That is, an element of area in the line-spread function of width du' is equal to the area of that portion of the ribbon representing the point-spread function intercepted by the interval du', this area being given by

$$dA = \frac{du'}{\cos(\psi - \phi)} \cdot \frac{S}{\sqrt{\dot{u}^2 + \dot{v}^2}} \qquad \text{I - 3.6}$$

where ϕ is the angle between the u axis and the tangent to the curve of the directrix. But

$$\begin{aligned} \dot{u}' &= \dot{u}\cos\psi + \dot{v}\sin\psi \\ &= \sqrt{\dot{u}^2 + \dot{v}^2}\cos(\psi - \phi), \end{aligned} \qquad \text{I - 3.7}$$

so the element of area of the line-spread function is represented by

$$dA = [S/\dot{u}']\,du', \qquad \text{I - 3.8}$$

in agreement with Eq. I-3.5.

It should be noted that the directrix may be a multi-valued function of u', in which case the line-spread function is the sum of all the intercepted elements (see Fig. I-5), and has an infinite discontinuity each value of u' corresponding to a bend and possibly a finite discontinuity at each value corresponding to a terminus. But the integral of the line-spread function exists (it is equal to the total area of the ribbon representing the point-spread function, which in fact is equal to the effective exposure time t_e) and there can be only two termini and a finite number of bends, so the validity of the Fourier transform expressed in Eq. I-3.5 is not jeopardized.

We shall now use Eqs. I-3.4 and I-3.5 to investigate a few special cases of image motion which are of practical importance.

Uniform Linear Motion

If $\dot{u}$ and $\dot{v}$ are constant with respect to time, then we have uniform linear motion having a velocity of magnitude

$$\dot{\ell} = \sqrt{\dot{u}_\ell^2 + \dot{v}_\ell^2}, \qquad \text{I-3.9}$$

in a direction

$$\phi_\ell = \arctan(\dot{v}_\ell / \dot{u}_\ell), \qquad \text{I-3.10}$$

In Eq. I-3.5

$$\dot{u}_\ell = \dot{\ell} \cos(\psi - \phi_\ell) \qquad \text{I-3.11}$$

is a constant with respect to u', so $\tilde{\Phi}_M$ becomes simply the normalized Fourier transform of the shutter function. If S is assumed to be unity during the exposure, then

$$\begin{aligned} \tilde{\Phi}_M &= \mathcal{S}(\pi \tilde{u}' \dot{u}_\ell' t_e) \\ &= \mathcal{S}[\pi \tilde{u}' \ell \cos(\psi - \phi_\ell)], \end{aligned} \qquad \text{I-3.12}$$

where $\ell = \dot{\ell} t_e$ is the distance the image travels in the exposure time. The corresponding point-spread function is a rectangle of length ℓ and height $1/\ell$. The two functions are shown in Fig. I-6.

If the shutter function is a trapezoid with a rise time t_r, then

$$\tilde{\Phi}_M = \mathcal{S}(\pi \tilde{u}' \dot{u}_\ell' t_e)\, \mathcal{S}(\pi \tilde{u}' \dot{u}_\ell' t_r), \qquad I-3.13$$

because a trapezoid is the convolution of two rectangles. If the shutter function is S-shaped during the rise time, that is, the convolution of a rectangle and a circle (focal-plane shutter),

$$\tilde{\Phi}_M = \mathcal{S}(\pi \tilde{u}' \dot{u}_\ell t_e) \left[\frac{2 J_1(\pi \tilde{u}' \dot{u}_\ell' t_r)}{(\pi \tilde{u}' \dot{u}_\ell' t_r)} \right]. \qquad I-3.14$$

Thus for uniform linear motion, the image-motion transfer function can be factored into a term dependent only on the effective exposure time regardless of the so-called shutter efficiency, and a term dependent only

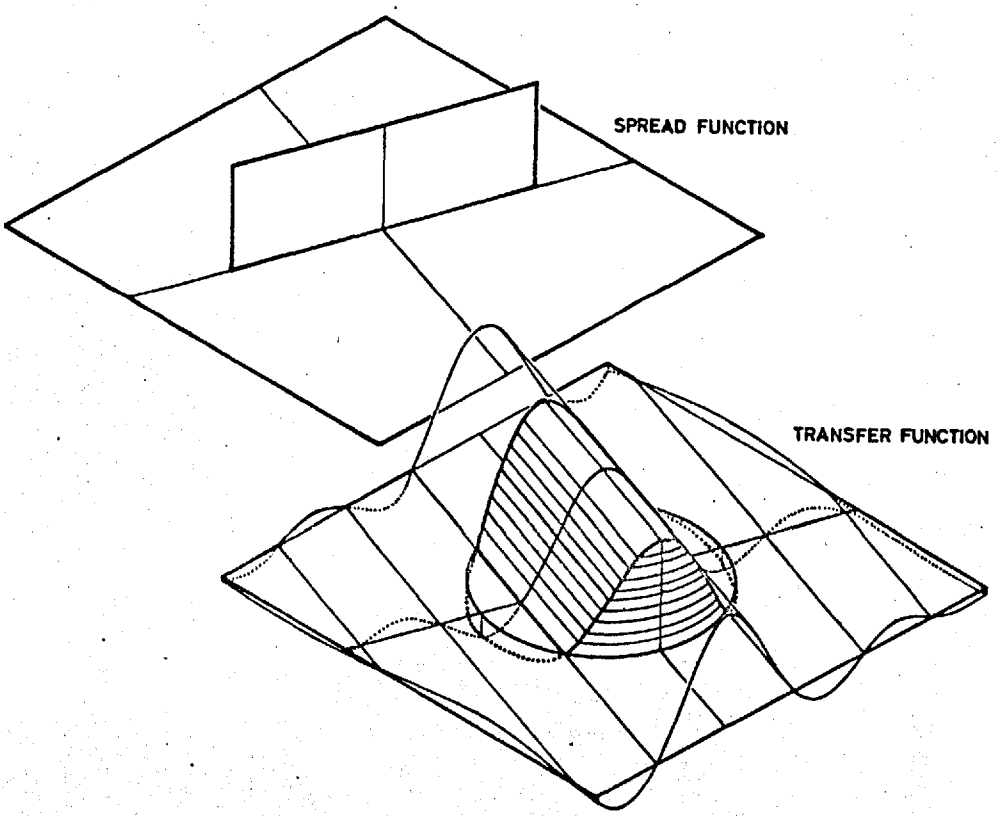


Figure I-6. Spread Function and Transfer Function for Uniform Linear Motion and a Rectangular Shutter Function. The cylinder represents the boundary of the frequency domain limited by optical diffraction.

on the rise time. The first is the Fourier transform of a rectangular function and the second the Fourier transform of the derivative of the shutter function during the rise time.

This result is valid only if the optical transfer function $\tilde{\mathcal{F}}_L$ is independent of time and no other form of image motion is present.

Simple Harmonic Motion

Next to uniform linear motion the most important elementary kind of image motion is that resulting from vibration in the camera system. In its simplest form this is simple harmonic motion, described by

$$u = u_a \sin(2\pi\nu t + \alpha), \quad v = v_a \sin(2\pi\nu t + \alpha), \qquad \text{I} - 3.15$$

where u_a and v_a are the component amplitudes of the motion (assumed constant), ν is the frequency of the vibration in cycles per second, and α is a phase term defining the position of the image in its path at time $t = 0$. The path of the image motion is a straight line making an angle

$$\phi_a = \arctan(v_a/u_a) \qquad \text{I} - 3.16$$

with the u axis and the amplitude is

$$a = \sqrt{u_a^2 + v_a^2}. \qquad \text{I} - 3.17$$

On rotation of coordinates we obtain

$$u' = u_a' \sin(2\pi\nu t + \alpha) \qquad \text{I - 3.18}$$

where

$$u_a' = a\cos(\gamma - \phi_a) \qquad \text{I - 3.19}$$

Substituting into Eq. I-3.4 we obtain

$$\tilde{\Phi}_M = t_e^{-1} \int_{-\infty}^{\infty} S \exp[-2\pi i \tilde{u}' u_a' \sin(2\pi\nu t + \alpha)]\, dt. \qquad \text{I - 3.20}$$

It is convenient to normalize the variable by $\hat{t} = t/t_e$; thus

$$\tilde{\Phi}_M = \int_{-\infty}^{\infty} S \exp[-2\pi i \tilde{u}' u_a' \sin(2\pi m\hat{t} + \alpha)]\, d\hat{t}, \qquad \text{I - 3.21}$$

where $m = \nu t_e$ is the number of cycles of vibration occurring in the effective exposure time. If now we assume that the shutter function is constant during the exposure time (rectangular shutter function), then we obtain

$$\tilde{\Phi}_M = \int_{-\frac{1}{2}}^{\frac{1}{2}} \exp\left[-2\pi i \tilde{u}' u_a' \sin(2\pi m \hat{t} + \alpha)\right] d\hat{t} \qquad I-3,22$$

where the origin of $\hat{t}$ is taken in the center of the exposure time.

It can be shown that, for integral values of m, Eq. I-3,22 reduces to the Bessel function

$$\tilde{\Phi}_M = J_o(2\pi \tilde{u}' u_a'), \qquad I-3,23$$

which is independent both of α and of the order of m. Also, for continuous values of m,

$$\lim_{m \to \infty} \tilde{\Phi}_M = J_o(2\pi \tilde{u}' u_a'). \qquad I-3,24$$

The latter expression, given by Scott[8], is of limited practical value, however, because in most cases not more than a few cycles of vibration will occur in the exposure time, and in this region the general dependence on m and on α (an uncontrolled parameter) may make $\tilde{\Phi}_M$ in any particular instance deviate considerably from that predicted by Eqs. I-3,23 or I-3,24.

Figure I-7 shows the manner in which $\tilde{\Phi}_M$ varies with α and with m for m ranging from 0 to 2. In this range, at least, it is clear that $\tilde{\Phi}_M$ depends strongly on both α and m. Figure I-8 shows this dependence more clearly, at least for the main lobe of $\tilde{\Phi}_M$.

Several features are apparent. The general trend is for $\tilde{\Phi}_M$ to approach the limiting value specified by Eq. I-3.24, but, for small values of m, $\tilde{\Phi}_M$ rises as m decreases, approaching unity as m approaches zero. The reason for the rise is, of course, the fact that, if the exposure time is less than one period, the distance the image travels during the exposure diminishes with m, and the blur is less. At integral values of m we have the nodes predicted by Eq. I-3.23. Between $m = 1$ and $m = 2$ is another node, of which more will be said later. Between the nodes, $\tilde{\Phi}_M$ can take on a range of values dependent on α.

The dependence on α is quite important. Although the amplitude and frequency of the vibration might conceivably be specified in advance, the reference phase

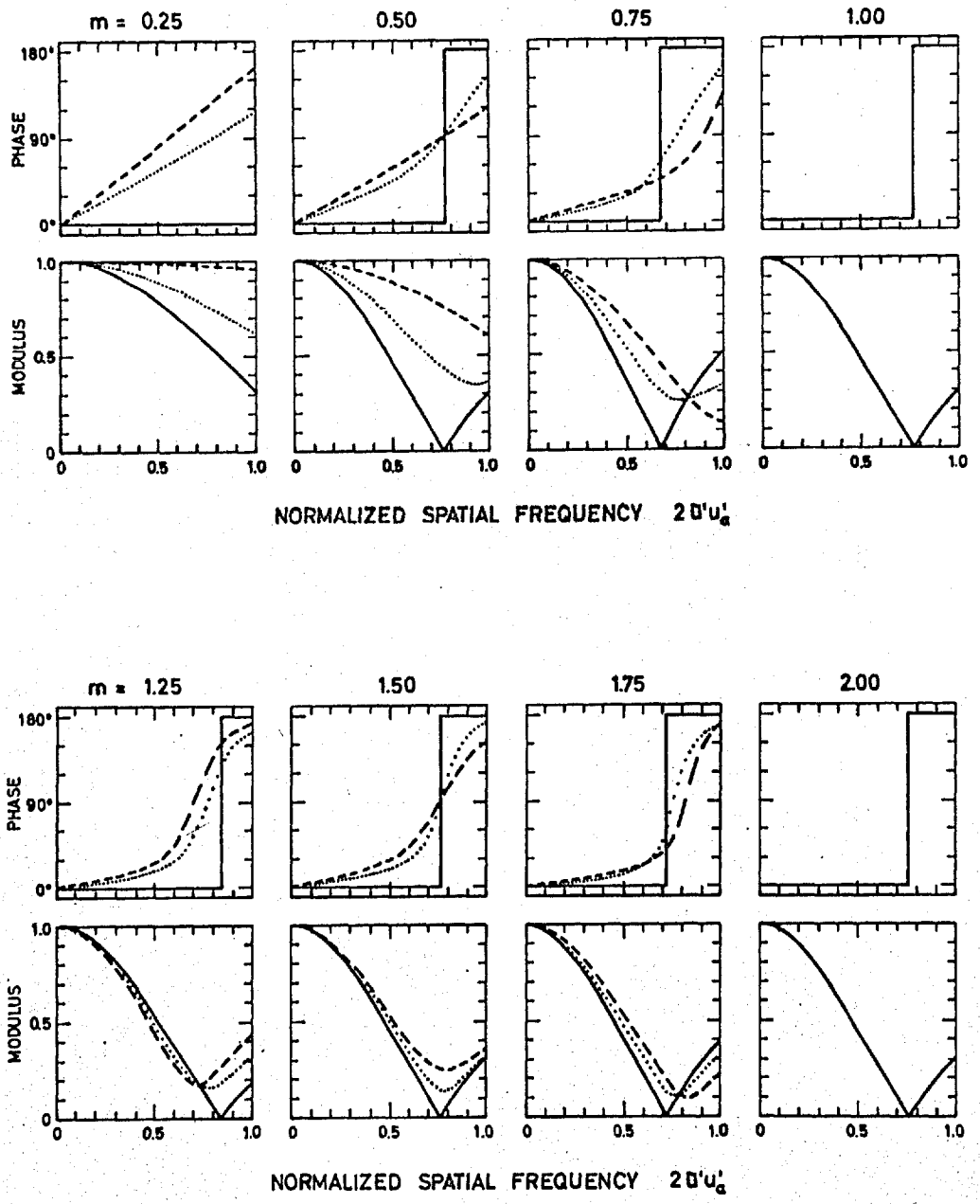


Figure I-7. Transfer Functions for Simple Harmonic Image Motion. The values of m indicate the number of cycles in the effective exposure time. The shutter function is rectangular. The solid curve is for $\alpha = o$, the dotted curve is for $\alpha = \pi/4$, and the dashed curve is for $\alpha = \pi/2$.

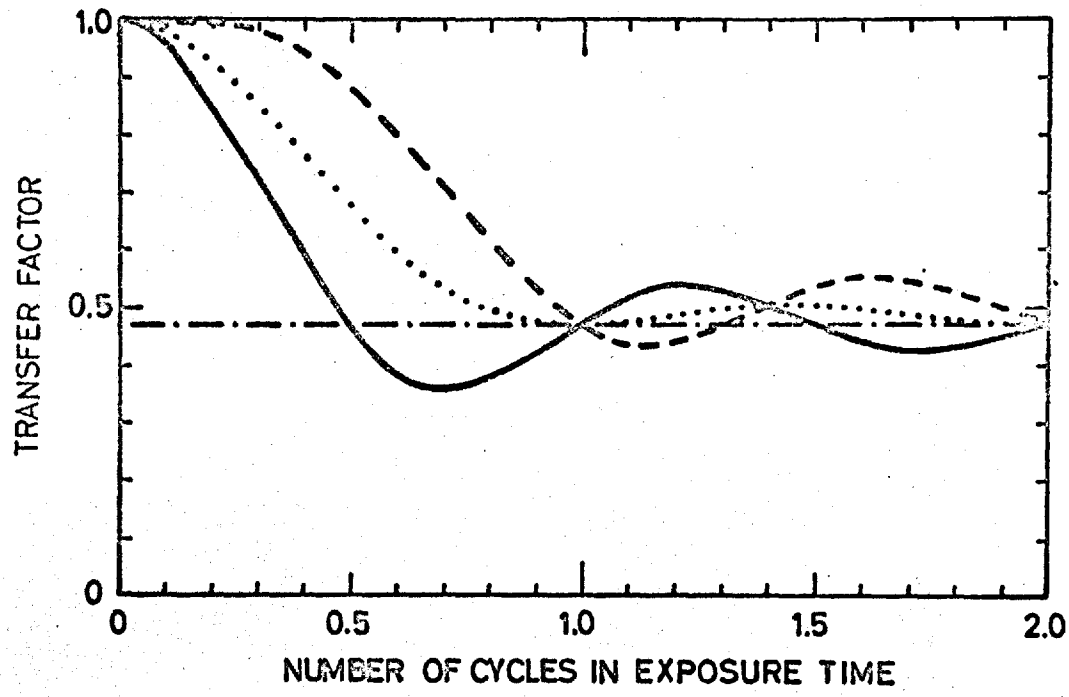


Figure I-8. Variation of the Simple Harmonic Image Motion Transfer Function with m and α, for a Normalized Spatial Frequency Value of 0.5. The shutter function is rectangular. The solid curve is for $\alpha = 0$, the dotted curve is for $\alpha = \pi/4$, and the dashed curve is for $\alpha = \pi/2$.

α is unpredictable in any given case, all values of being equally likely. Thus the best that can be done in describing or predicting $\tilde{\Phi}_M$ is to specify its probability distribution. In the present case this distribution has an upper and a lower limit, and it is more probable that a given value of $\tilde{\Phi}_M$ will be near one or the other of these limits than that it will lie midway between.

If instead of a rectangular shutter function, a triangular one is assumed, the same nodes occur at integral values of m (Eq. I-3,23), and the same limiting function applies (Eq. I-3,24), but the function converges much more rapidly to the limiting value. Figure I-9 shows the result obtained for the same parameters as Fig. I-8. It can be seen that, except for values of less than about 0.7, the value specified by the limiting function is quite a reasonable prediction for $\tilde{\Phi}_M$ for any value of α, at least for the main lobe of the transfer function.

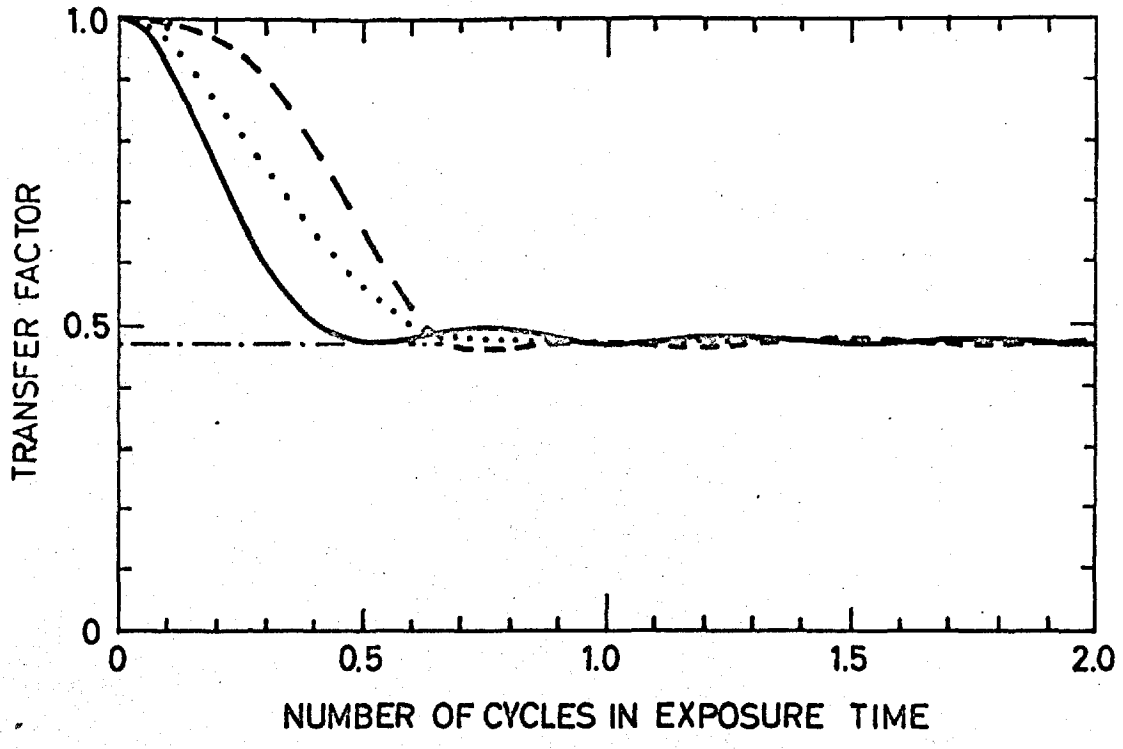


Figure I-9. Same as Fig. I-8 except that the shutter function is triangular. The solid curve is for $\alpha = o$, the dotted curve is for $\alpha = \pi/4$, and the dashed curve is for $\alpha = \pi/2$.

Combined Motions

In the general problem of combined motions

$$u = u_1 + u_2 + \dots, \qquad v = v_1 + v_2 + \dots, \qquad \text{I-3.25}$$

from which we obtain

$$u' = u_1' + u_2' + \dots, \qquad \text{I-3.26}$$

and

$$\dot{u}' = \dot{u}_1' + \dot{u}_2' + \dots. \qquad \text{I-3.27}$$

It should be noted that each pair of functions in I-3.25 is associated with a particular set of parameters, but because the orientation in general differs from pair to pair, the distribution of parameters in Eqs. I-3.26 and I-3.27 will vary with ψ.

The equivalent line-spread function is given by

$$S/\dot{u}' = S/(\dot{u}_1' + \dot{u}_2' + \dots), \qquad \text{I-3.28}$$

from which it can be seen that the resultant spread function is not a linear combination, product, or convolution of the separate component spread functions, and so the equivalent image-motion transfer function

cannot be a simple combination of the component transfer functions. It, of course, is given by

$$\tilde{\mathcal{E}}_M = t_e \int_{-\infty}^{\infty} S \exp[-2\pi i \tilde{u}'(u_1' + u_2' + \ldots)]\, dt \qquad \text{I-3.29}$$

The simplest case of combined motions is the sum of two or more uniform linear motions. Clearly the resultant is itself a uniform linear motion, and no further discussion is necessary.

The next most simple case of combined motions is that resulting from two or more vibrations of the same frequency. Then

$$u = u_a \sin(2\pi\nu t + \alpha), \quad v = v_a \sin(2\pi\nu t + \alpha + \beta), \qquad \text{I-3.30}$$

from which we obtain

$$u' = u_a' \sin(2\pi\nu t + \alpha'), \qquad \text{I-3.31}$$

where

$$\alpha' = \alpha + \beta', \quad \tan\beta' = \frac{(v_a/u_a)\sin\beta \tan\psi}{1 + (v_a/u_a)\cos\beta \tan\psi}, \qquad \text{I-3.32}$$

and

$$u_a' = \left(u_a^2 \cos^2\psi + v_a^2 \sin^2\psi + 2\, u_a v_a \sin\psi \cos\psi \cos\beta \right)^{1/2} \quad \text{I-3.33}$$

Equation I-3.31 has the same form as Eq. I-3.18, and thus the preceding analysis for a single source of vibration applies here as well for each direction, but in general the variation of $\tilde{\Phi}_M$ with ψ is different.

Of course, if $\beta = 0$, the two components of motion are in phase and the resultant motiont is indistinguishable from that resulting from a single source of vibration, and the previous analysis applies exactly.

On the other hand, if $\beta = \frac{1}{2}\pi$ and $u_a = v_a = a$, the path of motion is a circle and

$$u' = a \sin(2\pi\nu t + \alpha + \psi),$$

where the amplitude is independent of orientation but the reference phase is a linear function of ψ.

Figure I-10 illustrates these two extreme cases along with a general intermediate case for $m = 1$. The three different point-spread functions have ident-

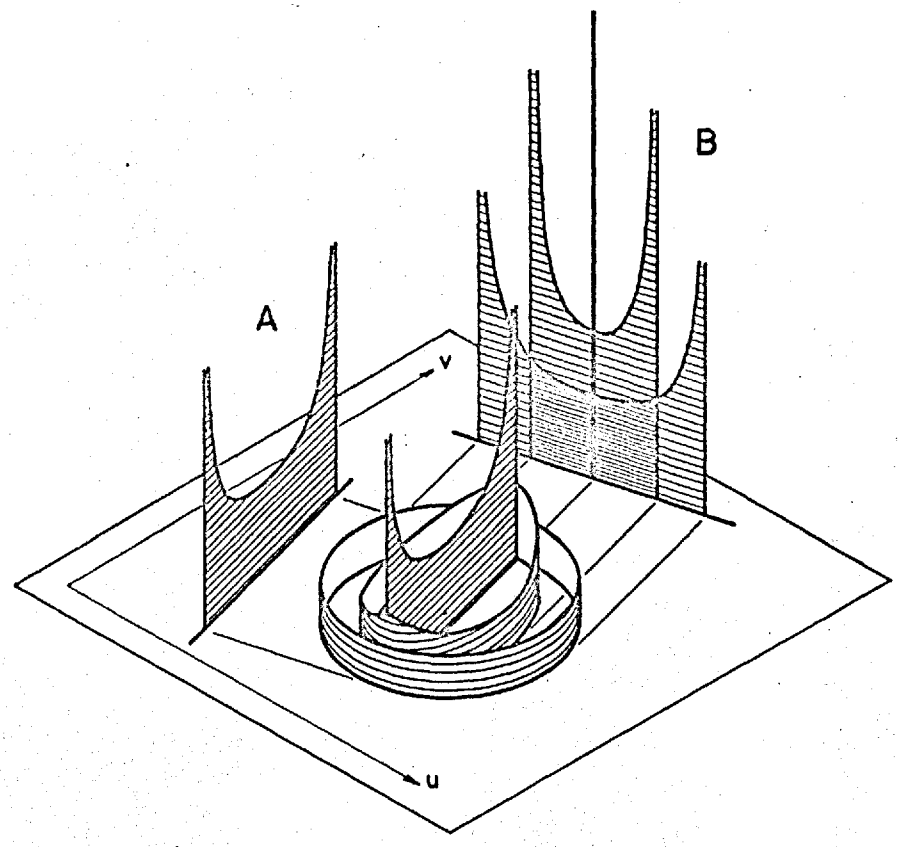


Figure I-10. Spread Functions for Combined Simple Harmonic Motions of the Same Frequency. Exactly one cycle in the exposure time is assumed. The shutter function is rectangular.

ical line-spread functions in the direction of A but different ones in the direction of B.

The more general problems of the combination of linear and simple harmonic motions and combined vibratory motions of different frequencies do not have any corresponding simplicity. Equations I-3.28 and I-3.29 must be applied directly.

Approximation for Small Degradations

One of the principal uses of transfer function theory is in the establishment of performance tolerances, that is, the determination of the magnitude of a source of degradation when the degradation is assumed to be just tolerable. For this purpose we can use an approximation which is valid for small values of the total degradation, the approximation being the use of the first three terms of the power series expansion for the exponential expression in the integrand of Eq. I-3.4.

First, however, let us establish as our coordinate reference for the spread function its center of gravity. Thus, before shifting our coordinates, we let

$$u' = \bar{u}' + u'' \qquad \text{I-3.35}$$

and substitute in Eq. I-3.4.

Shifting our coordinates by making $\bar{u}' = o$, we obtain

$$\tilde{\Phi}_M = t_e^{-1} \int_{-\infty}^{\infty} S \exp(-2\pi i \tilde{u}' u'') \, dt. \qquad \text{I-3.36}$$

The effect of shifting the coordinates is only the

introduction of a phase term which is a linear function of the spatial frequency.

We now make our approximation, obtaining

$$\tilde{\Phi}_M \approx t_e^{-1} \int_{-\infty}^{\infty} S\left[1 - 2\pi i \tilde{u}' \tilde{u}'' - 2\pi^2 \tilde{u}'^2 u''^2\right] dt$$

$$= 1 - 2\pi^2 \tilde{u}'^2 \left[t_e^{-1} \int_{-\infty}^{\infty} S\, u''^2 dt\right]. \qquad \text{I-3.37}$$

This expression is free of imaginary terms, thus indicating that, to the extent that the approximation is valid, asymmetry of the spread function is of no importance, and any central section of the transfer function is a parabola. It is always pessimistic, actual values of the function being somewhat higher.

A transfer function which corresponds to no degradation of the image is simply equal to unity for all spatial frequencies. Thus the second term of Eq. I-3.37 is conveniently considered as the loss L_M associated with the transfer function:

$$\tilde{\Phi}_M = 1 - L_M \qquad \text{I-3.38}$$

where

$$L_M = 2\pi \tilde{u}'^2 \left[t_e^{-1} \int_{-\infty}^{\infty} S\, u''^2\, dt \right], \qquad \text{I-3.39}$$

By changing the variable of integration to u', as in Eq. I-3.5, Eq. I-3.39 becomes

$$L_M = 2\pi \tilde{u}'^2 \left[\int_{-\infty}^{\infty} (S/\dot{u}')\, u''^2\, du' \right], \qquad \text{I-3.40}$$

and it becomes apparent that the quantity in square brackets is the second moment of the spread function.

If we assume a rectangular shutter function and the image motion is uniform linear motion, that is, $u' = \dot{u}_l' t$ then $\bar{u}' = 0$ and

$$L_M = \tfrac{1}{6} \pi^2 \tilde{u}'^2 u_l'^2 \qquad \text{I-3.41}$$

If the image motion is simple harmonic, that is

$$u' = u_a' \sin(2\pi\nu t + \alpha),$$

then

$$\tilde{u}' = u_a' \sin\alpha\, \mathfrak{S}(\pi m), \qquad \text{I-3.42}$$

and

$$L_m' = \pi^2 \tilde{a}'^2 u_a'^2 \left\{ [1 - \mathcal{G}^2(\pi m)] + (\cos 2\alpha)[\mathcal{G}^2(\pi m) - \mathcal{G}(2\pi m)] \right\}. \qquad \text{I-3.43}$$

The quantity inside the curly brackets expresses the dependence on m , the number of cycles of vibration in the effective exposure time, and α , the reference phase. It is plotted in Fig. I-11, which on comparison with Fig. I-8, shows that the dependence is quite adequately accounted for.

The quantity in the first set of square brackets gives the average dependence on m . The quantity in the second set of brackets defines the amplitude by which the loss fluctuates with α about the average value given by the first quantity.

The total expression in the curly brackets goes to zero for all α as m approaches zero and to unity as m approaches infinity. It also goes to unity for integral values of m . The intermediate node previously mentioned occurs when the quantity in the second set of

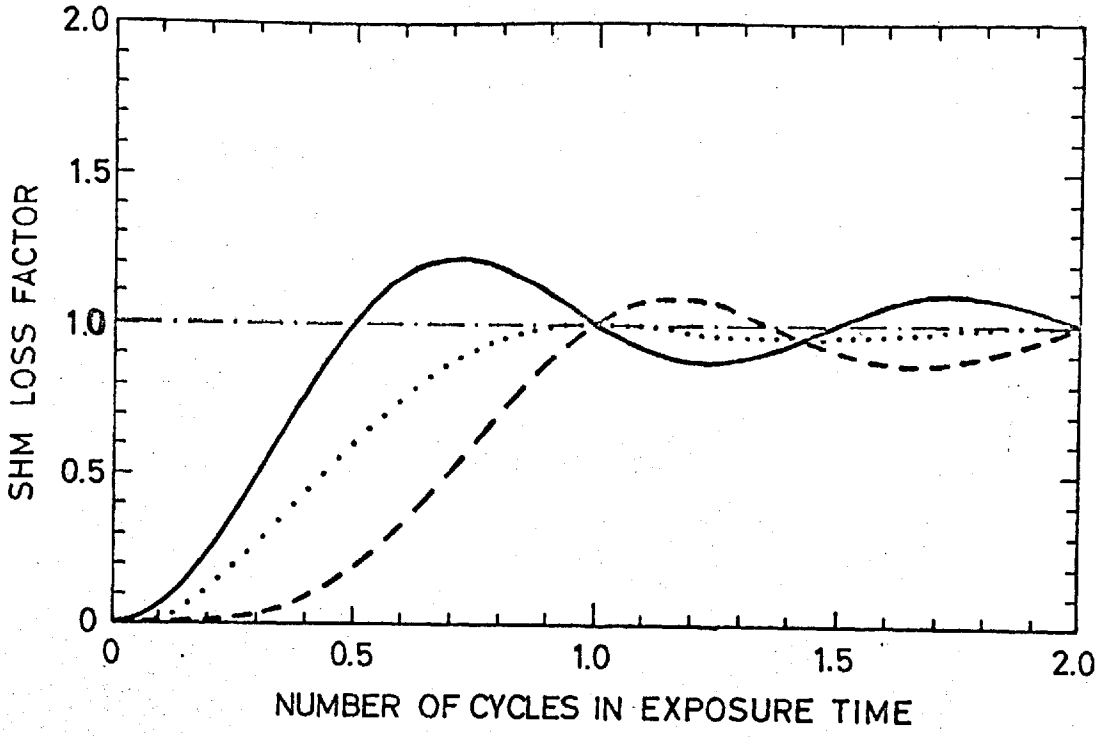


Figure I-11. Dependence of the Transfer Function Loss Factor on m and α (see equation (I-3.43)). The solid curve is for $\alpha = o$, the dotted curve is for $\alpha = \pi/4$, and the dashed curve is for $\alpha = \pi/2$. This figure should be compared with Fig. I-8.

square brackets is zero, that is, when $\tan \pi m = \pi m$. Other intermediate nodes will occur at larger values of m which satisfy this relationship.

If the image motion is the sum of uniform linear and simple harmonic motion, that is

$$u' = \dot{u}_l' t + u_a \sin(2\pi \nu' t + \alpha),$$

then

$$L_M = \frac{1}{6} \pi^2 \tilde{u}'^2 \dot{u}_l'^2 + \pi^2 \tilde{u}'^2 u_a'^2 \left\{ \left[1 - \mathfrak{S}^2(\pi m) \right] + (\cos 2\alpha) \left[\mathfrak{S}^2(\pi m) - \mathfrak{S}(2\pi m) \right] \right\} + 2\pi^2 \tilde{u}'^2 \dot{u}_l' u_a' \left\{ \cos \alpha \left[(\pi m)^{-1} \left(\mathfrak{S}(\pi m) - \cos \pi m \right) \right] \right\}. \qquad \text{I-3.44}$$

Here it is apparent that the first term is the loss which would result from the linear motion alone, the second is the loss which would result from the simple harmonic motion alone, and the third the loss resulting from the interaction between the two.

The interaction loss also depends on m and α, as shown by the expression inside the curly brackets. This expression is shown in Fig. I-12, where it can be seen from the nodes present that the interaction term vanishes when the quantity inside the square brackets is zero, that is, when $\tan \pi_m = \pi_m$, the same condition which applies to the intermediate nodes of the pure vibration dependence. Thus, at these values of m the image motion transfer function is independent of the reference phase α. The interaction term, however, does not, in general, vanish at integral values of m.

It is instructive to observe how the transfer function varies with m and α as the proportion between the linear component and the vibration component is varied, keeping the total loss constant (for $m \to \infty$).

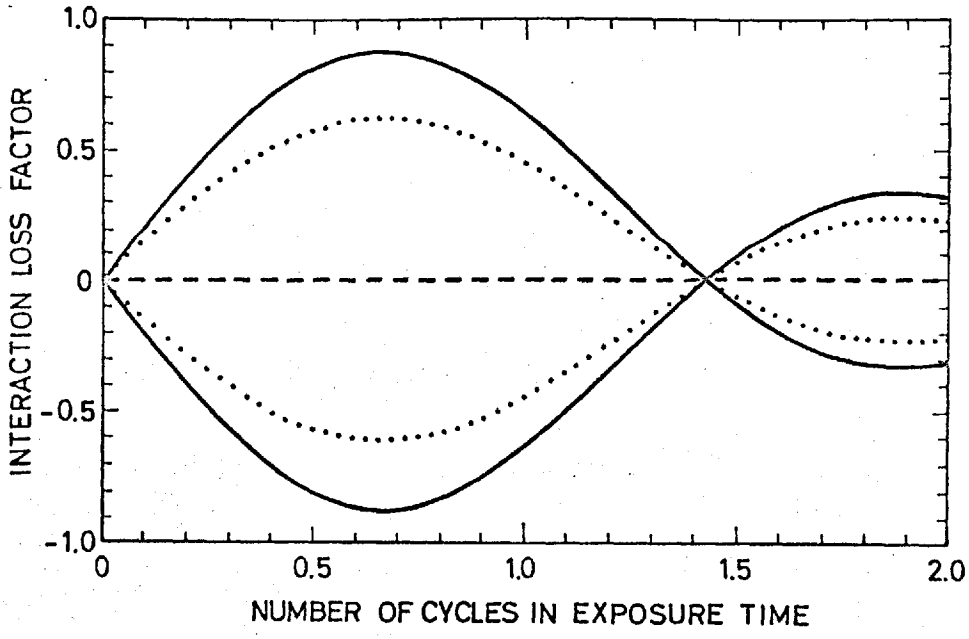


Figure I-12. Interaction Loss Factor for Uniform Linear Image Motion combined with Simple Harmonic Motion (see equation (I-3.44)). The factor 2 is incorporated in the figure. The solid curve is for $\alpha = 0$ or π, the dotted curve for $\alpha = \pi/4$ or $3\pi/4$, and the dashed curve for $\alpha = \pi/2$.

Figure I-13 shows this for a total loss of 0.2 with the proportion ranging from pure linear motion to pure vibratory motion.

The most striking feature is the dominant effect of the interaction term for even small amounts of vibration added to the linear image motion. This of course reflects the fact that when the vibratory component travels in the same direction as the linear component the blur is increased, whereas when it travels in the opposite direction the blur is reduced or even cancelled.

As was stated previously, α is a parameter over which there is normally no control, all values of α being equally likely. Thus, even if all other parameters are specified, no specific value of the transfer function can be predicted unless the value of m corresponds to one of the nodes as seen in Fig. I-13. Instead, the transfer function may take on a range of values which is best described by a probability distribution, and a tolerance level can be established

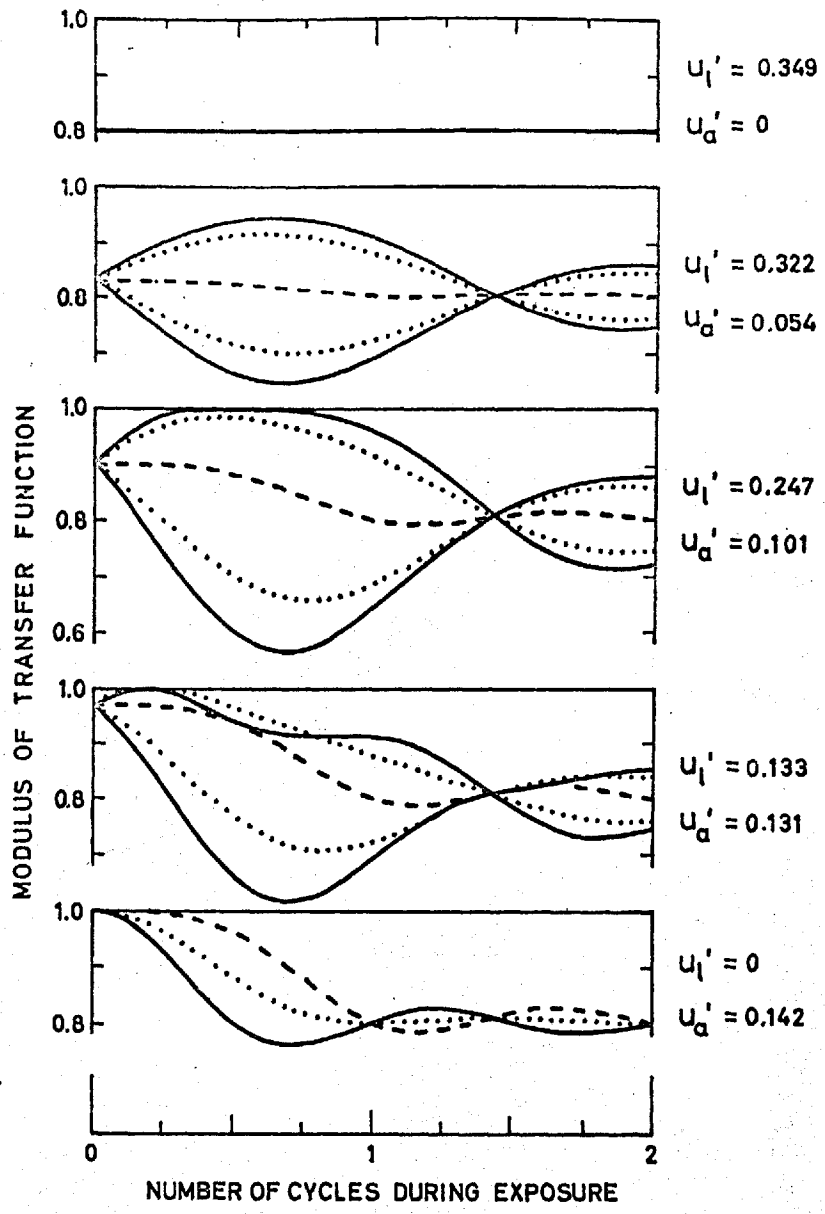


Figure I-13. Variation of the Transfer Function for combined uniform linear image motion and simple harmonic motion as the proportion of the two components is varied from pure linear motion to pure simple harmonic motion. The solid curve is for $\alpha = 0$ or π, the dotted curve for $\alpha = \pi/4$ or $3\pi/4$, and the dashed curve for $\alpha = \pi/2$.

on the basis of the statistical performance required. In the present case a fair approximation to the correct procedure can be made by using absolute values for the quantities in square brackets in Eq. I-3.44 and choosing a value of α to conform to the statistical performance required. Thus, no photograph will be worse than the value obtained by setting $\alpha = 0$, and if the peak-to-peak amplitude of the vibration does not exceed the distance of linear motion, then 75% of the photographs taken will be better than the value obtained by setting $\alpha = \frac{1}{4}\pi$ and 50% for $\alpha = \frac{1}{2}\pi$.

If the image motion is the sum of two simple harmonic motions of differing frequency, that is,

$$u' = u_{a_1}' \sin(2\pi\nu_1 t + \alpha_1) + u_{a_2}' \sin(2\pi\nu_2 t + \alpha_2),$$

then

$$L_M = \pi^2 \tilde{u}'^2 u_{a1}'^2 \left\{ \left[1 - \mathcal{G}^2(\pi m_1) \right] + (\cos 2\alpha_1) \left[\mathcal{G}^2(\pi m_1) - \mathcal{G}(2\pi m_1) \right] \right\}$$
$$+ \pi^2 \tilde{u}'^2 u_{a2}'^2 \left\{ \left[1 - \mathcal{G}^2(\pi m_2) \right] + (\cos 2\alpha_2) \left[\mathcal{G}^2(\pi m_2) - \mathcal{G}(2\pi m_2) \right] \right\}$$
$$+ 2\pi^2 \tilde{u}'^2 u_{a1}' u_{a2}' \left\{ \cos(\alpha_1 - \alpha_2) \left[\mathcal{G}(\pi m_1 - \pi m_2) - \mathcal{G}(\pi m_1)\, \mathcal{G}(\pi m_2) \right] \right.$$
$$\left. + \cos(\alpha_1 + \alpha_2) \left[\mathcal{G}(\pi m_1)\, \mathcal{G}(\pi m_2) - \mathcal{G}(\pi m_1 + \pi m_2) \right] \right\}.$$

I-3,45

Here again the first two terms are the losses which would result from each component of image motion by itself and the third term is the loss resulting from interaction between the two.

This example is more difficult to describe than the previous examples because we now have six parameters, two of which are independent random variables. In particular the probability distribution of L_M becomes much more complicated.

In order to get a general idea of the properties of Eq. I-3,45 we must look more closely at the interaction term. This term itself contains all six parameters, but they are conveniently grouped. The expressions in the square brackets describe the dependence on both m_1 and m_2. Figure I-14 shows the manner in which the

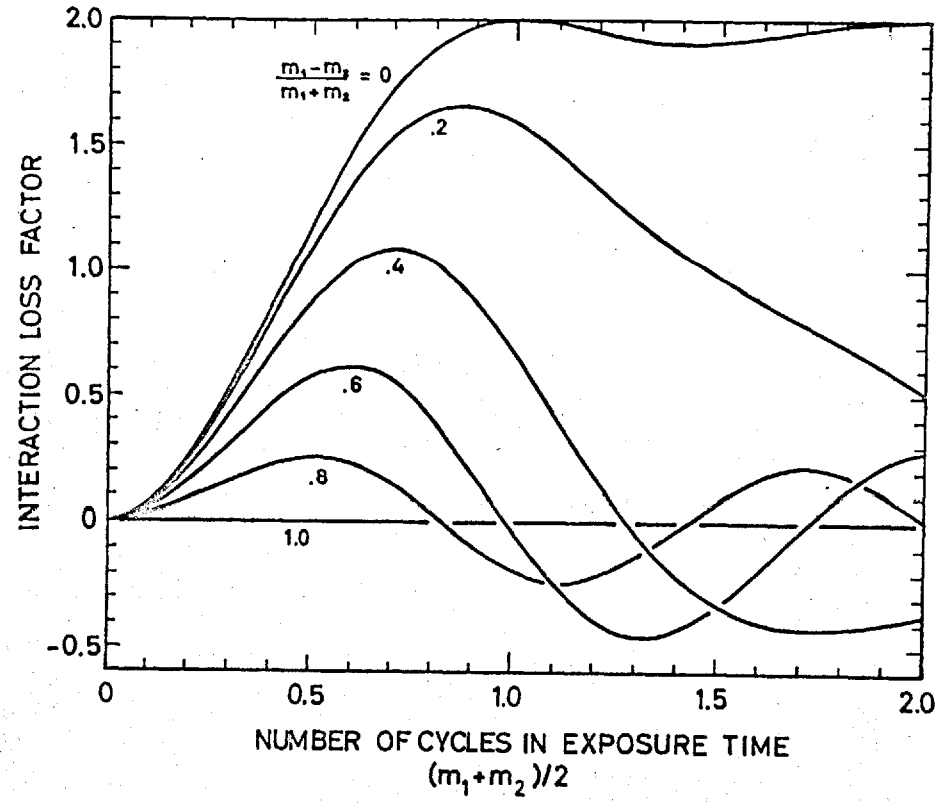


Figure I-14. Interaction loss factor for combined simple harmonic motions of different frequencies. This is for the expression in the first square brackets of the interaction term in equation (I-3.45). The dependence on α_1 and α_2 is not shown.

quantity in the first square brackets varies and Fig. I-15 the second. The parameters m_1 and m_2 have been transformed to $(m_1-m_2)/(m_1+m_2)$ and $\frac{1}{2}(m_1+m_2)$ in the figures.

It is clear that the magnitudes of the expressions generally increase as the difference between the frequencies decreases. This means that the interaction between the two components reaches its maximum when the two frequencies become the same. If the two frequencies are the same, then the loss is a maximum when the motions are in phase with each other, that is, $\alpha_1 = \alpha_2 = \alpha$, in which case the loss can be expressed as in Eq. I-3,43, where $u_a' = u_{a1}' + u_{a2}'$. For all other frequency and phase combinations, the loss will be diminished. Thus, if a tolerance is to be established for a combination of two different frequencies, then a safe maximum limit for the loss can be set by assuming that both frequencies are equal to their average and that they are in phase with each other. This procedure is, of course, a short-cut and may be excessively conserv-

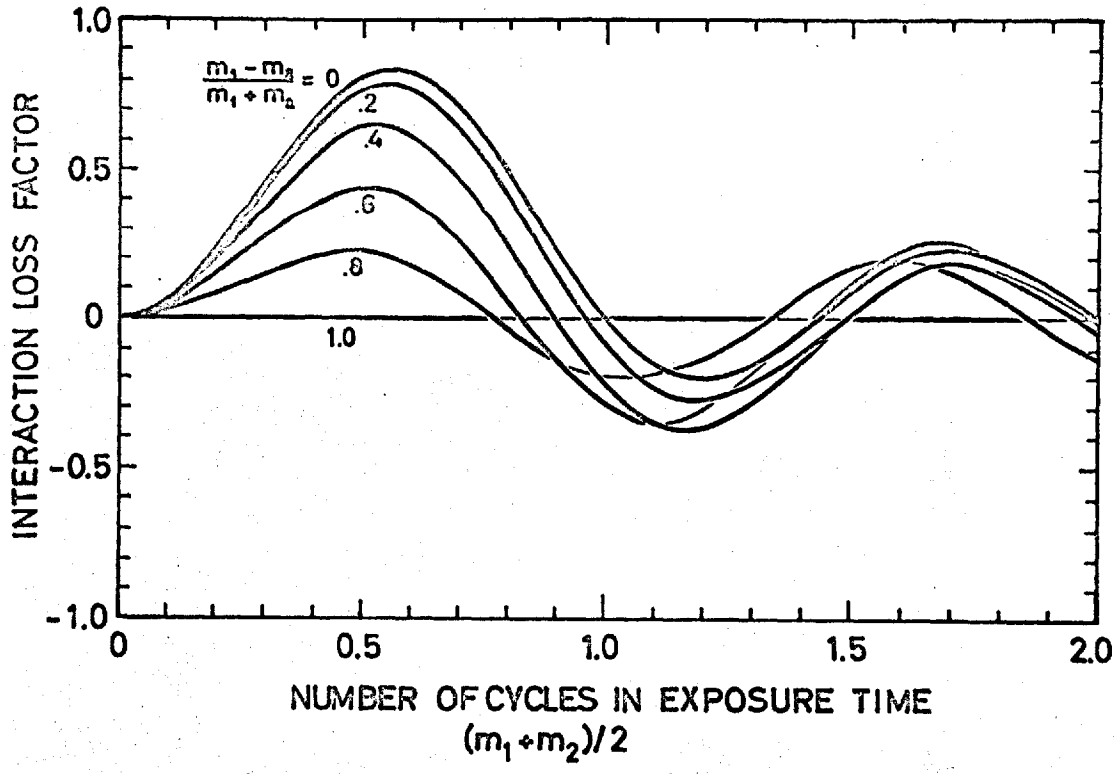


Figure I-15. Interaction loss factor for combined simple harmonic image motions of different frequencies. This is for the expression in the second square brackets of the interaction term in equation (I-3.45). The dependence on α_1 and α_2 is not shown.

ative if the frequency difference is large. The correct procedure is to calculate the probability distributions from which the tolerance can be determined according to the statistical performance required.

If the image motion is the sum of any number of uniform linear motions and simple harmonic motions, then the loss can be calculated with the terms contained in Eqs. I-3.44 and I-3.45. First all linear motions can be combined into a resultant linear motion and the vibrations for each frequency combined into a single resultant term. Then the loss will consist of a term for the linear motion as given by the first term in Eq. I-3.44, a term for each vibration as given by the second term in Eq. I-3.44, or either of the first two terms in Eq. I-3.45, an interaction term between the linear motion and each of the vibrations as given by the third term in Eq. I-3.44, and an interaction term for every pair of vibrations, as given by the third term in Eq. I-3.45.

The general procedure followed in this chapter can easily be extended to other kinds of image motion,

if desired. The two kinds which have been used in this chapter are, however, the ones of greatest practical interest.

PART II. INTRODUCTION

The first part of this thesis dealt with two time-dependent factors which are significant in determining the image-forming properties of a photographic system. The remainder of this thesis deals with some of the properties of images formed by an optical system looking at the object through a turbulent medium.

The principal effect of a turbulent medium is to distort the transmitted wavefront for each point in the object so that, for each object point, the wavefront reaching the optical system contains random structure in both phase and amplitude.

In the main we will not concern ourselves here with the details of the manner in which the phase and amplitude variations arise,[13,14] but with the manner in which they interact with the optical system in determining the properties of the images formed. However, a few words might be said in a qualitative way about the source of the phase and amplitude variations so that their significance might be more easily understood.

First of all, turbulence by itself does not result in wavefront disturbances.[15] This will occur only if inhomogeneities in refractive index are present, and in the case of the atmosphere, the latter result almost entirely from thermal inhomogeneities.[16] The turbulence establishes the random structure of the inhomogeneities, and consequently the structure of the random disturbances in the transmitted wavefront.

The immediate effect on the wavefront passing through the inhomogeneous medium is that only phase disturbances are imposed. However, as the wave propagates, the energy is redistributed along the wavefront because of the random "focus-defocus" effect of the phase disturbances, and this gives rise to the random amplitude component. Thus if the turbulent portion of the medium is near the optical system, only phase disturbances will be significant in the received wavefront,[17] unless the turbulence is severe. If amplitude fluctuations are observed, then they

must arise from distant turbulent structure.[17] Of course, both near and distant turbulence may, and generally will, be present, in which case the amplitude structure comes only from the distant turbulence, whereas the phase structure arises from both.

Another point is of considerable importance. If the disturbance were literally in the entrance pupil of the optical system, then, at any given moment the wavefronts from all points in the object would be affected (in phase alone) identically, and, to the extent that the optical system is isoplanatic, stationarity in the image plane will be achieved for each instantaneous image as well as for the average image. However, if the disturbance is distant from the entrance pupil, then wavefronts from separated points in the object will pass through different portions of the inhomogeneous medium before reaching the entrance pupil, and the image at any given moment will not be stationary. Thus for the instantaneous image, Fourier transform theory is not valid. However,

if the random structure of the disturbance is statistically stationary, then the average image will again be stationary, and as long as we restrict ourselves to considering stationary random structures and average properties of the image, we are free to use Fourier transform theory.

Historically, the principal concern with the effects of wavefront propagation through a turbulent medium has been that of astronomers,[18] partly because they are accustomed to using well-corrected, large optical instruments looking through a very long air path, and partly because they deal primarily with isolated stars, which are in effect discrete point sources, and the effects of atmospheric turbulence are most noticeable for such objects.

In observing the effects of atmospheric turbulence on star images, they have distinguished three image properties.[18] These are 1) scintillation, or fluctuations in the total intensity of the star image, 2) agitation, or moving about of the star image, and 3) blur, or

spreading of the star image beyond the diffraction pattern which would be expected from the optical system in the absence of turbulent structure. Terminology in this area is anything but consistent among astronomers,[18] but the general character of the three distinguishable properties is commonly agreed upon.

Unfortunately, these three properties do not in general depend on the random structure of the received wavefront alone, but also on the properties of the optical system used to form the image. For example, it has been observed that at least scintillation and agitation vary with the size of the entrance pupil.[19,20] The object of the remainder of this thesis is to investigate the manner in which the optical system and the random structure of the received wavefront interact in producing the above three phenomena.

It will be noted that the tilde notation will be only sparingly used in this part. This is because the emphasis here is on operations on the pupil function, in which displacement corresponds to spatial frequency

in the image plane. Thus we will use x', y' to represent displacements in the pupil, and it will be understood that these variables, being normalized in the pupil, can be identified with the reduced spatial frequency variables $\tilde{u}$, $\tilde{v}$ in the image plane.

A word is in order about the statistical averages with which we shall deal.

The principal random function of concern is the random portion of the incident wavefront which varies with both space and time. The area of the pupil can be considered to be a kind of window through which we can see a finite portion of a hypothetical infinite disturbed wavefront which coincides with our observed wavefront where we observe it. Also, we can observe it for only a finite period of time, but we can assume that our spatially infinite hypothetical wavefront is also infinite in time, coinciding with our observed wavefront during the time of observation. This hypothetical infinite wavefront is itself a member of a set, or ensemble, of all possible hypothetical wavefronts which could be

generated by the same random process which produced our observed wavefront.

Now the fundamental statistical properties (averages) are obtained in principle by averaging over the ensemble. If we assume that the generating process is strictly stationary, then the averages we obtain are independent of space and time. Furthermore, we can invoke the ergodic hypothesis, which means that the averages obtained over any (infinite) sample are the same as the corresponding ensemble averages. Moreover, for any sample, the averages over space alone or over time alone are also the same as the ensemble averages. Throughout the following material all averages indicated by angular brackets $\langle\ \rangle$ represent ensemble averages, although in most cases they can be interpreted as time or space averages wherever appropriate.

CHAPTER II-1. THE PUPIL FUNCTION WITH RANDOM COMPONENT

If the medium between the object and the optical system were perfectly homogeneous, then the wavefront emanating from any point on the object would, on entering the optical system, be coincident with a sphere centered on the object point and of uniform amplitude a_o. The essential function of the optical system is to change the curvature of each of the transmitted wavefronts so that each of them converges to a point in the image space providing a total image which is a point-for-point map of the object. However, it can only do this imperfectly, and apart from changing the curvature, the optical system modifies the incoming wavefront in both amplitude and phase according to its pupil function so that the transmitted wave can be described by

$$a = a_o f_i = a_o T_i e^{iKW_i}, \qquad \text{II-1.1}$$

where both τ_L and W_L are functions of position in the pupil, τ_L describing the imposed amplitude alterations, including the boundary of the clear aperture, and W_L describing the aberration of the transmitted wavefront.

If the medium between the object and the optical system is turbulent, containing random inhomogeneities of refractive index, then in general the wavefront reaching the optical system contains random disturbances in both amplitude and phase,[13,14] and can be represented by

$$a_R = a_0 e^{\alpha + i\beta} \qquad \text{II} - 1,2$$

where both α and β are functions of position in the wavefront and also functions of time. On being transmitted by the optical system the wave can be described by

$$a = a_0 (e^{\alpha + i\beta}) f_L . \qquad \text{II} - 1,3$$

On comparison with equation II-1.1, it can be seen that the effect of the turbulent medium can be considered as a time-dependent modifier of the pupil function, or better as a random component of the effective pupil

function. Thus we can define the random part of the pupil function as

$$f_R = e^{\alpha + i\beta}, \qquad \text{II} - 1.4$$

We can now assume simple statistical properties for the two real random functions α and β and still retain a reasonable physical model. The assumptions we shall make are as follows:

1) Both α and β are continuous functions of space and time. This property is not essential for many of the image-forming properties we shall obtain, but it is reasonable if the disturbances have arisen from the turbulent structure of the medium.

2) Both α and β are at least locally stationary in the statistical sense. This assumption corresponds to that of local isoplanarity in the linear treatment of the image-forming properties of aberrated optical systems.

3) The random phase function β has a mean value of zero. Again, this assumption is not often necessary, but it is reasonable if the disturbances have arisen from the turbulent structure of the medium.

4) The random amplitude function α has a non-zero mean.[1] This is necessary for the conservation of energy, as will be discussed in the following section.

5) Both α and β have normal (Gaussian) probability densities. A convincing theoretical justification for this assumption would be difficult to obtain, but we can say that observable properties derived with this assumption are at least not incompatible with experimental observations.

We shall use the common symbols σ_α, σ_β, σ_α^2, σ_β^2 to represent the standard deviations and variances of α and β, and also ρ_α, ρ_β, $\rho_{\alpha\beta}$ to represent the

---

[1] A superscript in square brackets indicates an entry under NOTES at the end of part II.

normalized auto- and cross-correlation functions of α' and β .

Returning now to the question of the mean value of α , we first note that in the absence of turbulence the intensity of the light incident on the optical system is given by

$$E_o = a_o^2 \qquad \text{II} - 1.5$$

When turbulence is present, the "instantaneous" intensity is given by

$$E = |a_R|^2 = a_o^2 e^{2\alpha} \qquad \text{II} - 1.6$$

the mean value of which is

$$\langle E \rangle = a_o^2 \langle e^{2\alpha} \rangle , \qquad \text{II} - 1.7$$

which, however, must be equal to E_o if the energy is to be conserved. Therefore

$$\langle e^{2\alpha} \rangle = 1 . \qquad \text{II} - 1.8$$

If we separate α into the sum of its mean value and the fluctuations with zero mean, that is

$$\alpha = \langle \alpha \rangle + \alpha' , \qquad \text{II - 1.9}$$

then

$$\langle e^{2\alpha} \rangle = e^{2\langle \alpha \rangle} \langle e^{2\alpha'} \rangle . \qquad \text{II - 1.10}$$

If we assume a normal probability density for α, then (see Appendix III),

$$\langle e^{2\alpha'} \rangle = e^{2\sigma_\alpha^2} \qquad \text{II - 1.11}$$

Substituting equations II-1.10 and II-1.11 into equation II-1.8, we find

$$e^{2(\langle \alpha \rangle + \sigma_\alpha^2)} = 1 , \qquad \text{II - 1.12}$$

from which we conclude that

$$\langle \alpha \rangle = - \sigma_\alpha^2 . \qquad \text{II - 1.13}$$

The validity of the above model as far as amplitude fluctuations are concerned is indicated in figure II-1. It

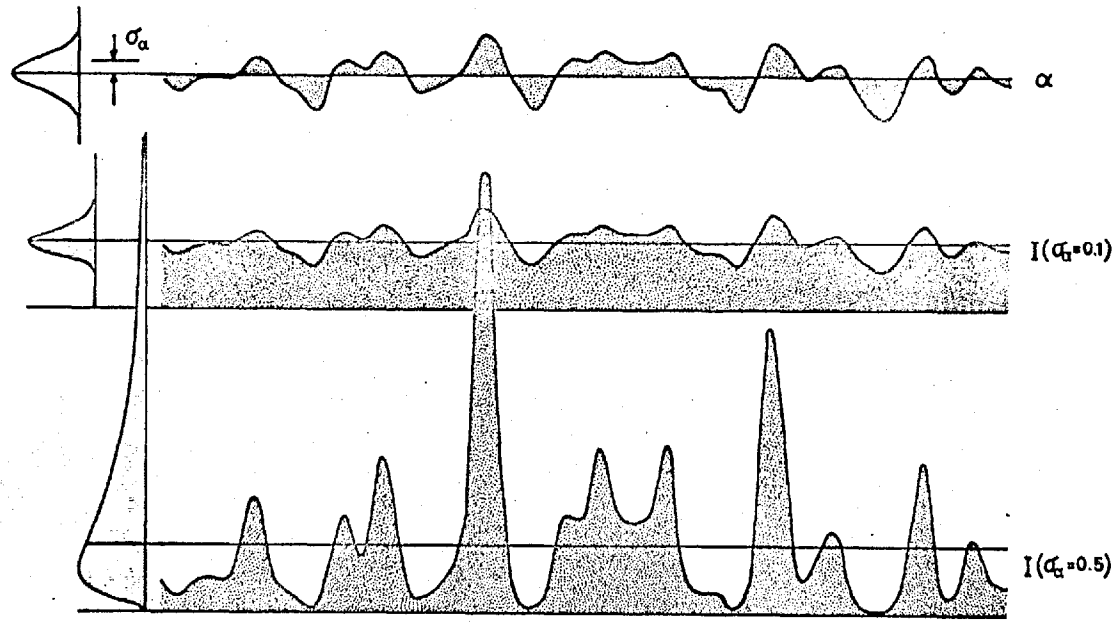


Figure II-1. Variation of the intensity function $I=e^{2\alpha}$ and its probability density for several values of σ_α ; assuming a normal probability density for α.

was constructed from a table of random normal deviates. The top line shows the assumed fluctuations in α. The middle line shows the observable fluctuations in intensity for an assumed value of $\sigma_\alpha = 0.1$ generated by the function shown in the top line. The bottom line shows the corresponding intensity fluctuations if $\sigma_\alpha = 0.5$. The probability density distributions of the three functions are shown on the left. Qualitatively the model fits experimental observations in that when scintillation (intensity fluctuations in star images) is moderately severe there occur bright flashes of intensity several times the average intensity in the image.[19] This indication that the model is reasonable is attributable to the exponential form e^{α} rather than to the assumption of a normal distribution in α, but the assumption of normality is at least consistent with the experimental data which has come to the author's attention.[14]

CHAPTER II-2. SCINTILLATION

As described in the introduction, one of the phenomena familiar to astronomers and attributable to atmospheric turbulence is the fluctuation in the intensity of a star image as observed by the eye or a photoelectric detector in the image plane of a telescope. It has been observed that the magnitude of the fluctuations depends among other things on the size of the telescope aperture.[19,20]

In this section we shall obtain a reasonably simple relation between the essential statistical variable describing scintillation and the assumed statistical properties of the random component of the pupil function. This relation contains as parameters the essential (non-statistical) properties of the pupil function, so that the effect on the scintillation of the shape as well as the size of the pupil, and including obscurations and apodization as well, can readily be determined.

The essential statistical variable describing scintillation is the coefficient of variation of the observable fluctuations of the light contained in the

total image of a point source, or star. This is the ratio of the standard deviation of the fluctuations to the mean value of the signal obtained.

The total instantaneous power in the image at any given moment is equal to the total instantaneous power in the exit pupil. Thus

$$I = a_o^2 \iint_{-\infty}^{\infty} |f|^2 dx\,dy = a_o^2 \iint_{-\infty}^{\infty} T_L^2 e^{2\alpha} dx\,dy. \qquad \text{II-2.1}$$

Its mean value is given by

$$\langle I \rangle = \left\langle a_o^2 \iint_{-\infty}^{\infty} T_L^2 e^{2\alpha} dx\,dy \right\rangle, \qquad \text{II-2.2}$$

which, because $e^{2\alpha}$ is the only random variable involved, becomes

$$\langle I \rangle = a_o^2 \iint_{-\infty}^{\infty} T_L^2 \langle e^{2\alpha} \rangle dx\,dy = a_o^2 \iint_{-\infty}^{\infty} T_L^2 dx\,dy$$

$$= a_o^2 T, \qquad \text{II-2.3}$$

where

$$T = \iint_{-\infty}^{\infty} T_L^2 dx\,dy \qquad \text{II-2.4}$$

is the total transmittance (area multiplied by mean transmittance) of the pupil.

The square of the total instantaneous power in the image is given by

$$I^2 = \left[a_o^2 \iint_{-\infty}^{\infty} |f|^2 dx\, dy\right]^2 = a_o^4 \iint_{-\infty}^{\infty} \iint_{-\infty}^{\infty} |f_1|^2 |f_2|^2 dx_1\, dy_1\, dx_2\, dy_2 . \quad \text{II-2.5}$$

By changing the variables so that

$$x_1 = x + \frac{x'}{2} \qquad y_1 = y + \frac{y'}{2}$$
$$x_2 = x - \frac{x'}{2} \qquad y_2 = y - \frac{y'}{2} \qquad \text{II-2.6}$$

we obtain

$$I^2 = a_o^4 \iint_{-\infty}^{\infty} \iint_{-\infty}^{\infty} |f_1|^2 |f_2|^2 dx\, dy\, dx'\, dy' . \qquad \text{II-2.7}$$

At this point it is worth noting that the variables $x, y ; x', y'$ can be identified with the same variables involved in the autocorrelation integral of the pupil function which is used to obtain the optical transfer function. In particular, x' and y' are associated with spatial frequencies in the image plane. The use to which

we shall put this identification will occur later.

The mean value of I^2 is given by

$$\langle I^2 \rangle = \iint_{-\infty}^{\infty} \iint_{-\infty}^{\infty} T_{L1}^2 T_{L2}^2 \langle |a_{R1}|^2 |a_{R2}|^2 \rangle \, dx\, dy\, dx'\, dy'$$

$$= \iint_{-\infty}^{\infty} \left[\iint_{-\infty}^{\infty} T_{L1}^2 T_{L2}^2 \, dx\, dy \right] \langle |a_{R1}|^2 |a_{R2}|^2 \rangle \, dx'\, dy' . \qquad \text{II-2.8}$$

Let us define the normalized function

$$D_o \equiv \frac{\iint_{-\infty}^{\infty} T_{L1}^2 T_{L2}^2 \, dx\, dy}{\iint_{-\infty}^{\infty} T_L^2 \, dx\, dy} = \frac{\iint_{-\infty}^{\infty} T_{L1}^2 T_{L2}^2 \, dx\, dy}{T} , \qquad \text{II-2.9}$$

whereupon

$$\langle I^2 \rangle = T \iint_{-\infty}^{\infty} D_o \langle |a_{R1}|^2 |a_{R2}|^2 \rangle \, dx'\, dy' . \qquad \text{II-2.10}$$

Now, letting V represent the coefficient of variation, we have

$$V^2 = \frac{\langle I^2\rangle - \langle I\rangle^2}{\langle I\rangle^2} = \frac{\langle I^2\rangle - a_o^4 T^2}{a_o^4 T^2}$$

$$= \frac{\iint_{-\infty}^{\infty} D_o \langle |a_{R1}|^2 |a_{R2}|^2\rangle \, dx'dy' - a_o^4 T}{a_o^4 T}$$

$$= \frac{\iint_{-\infty}^{\infty} D_o \left[\frac{\langle |a_{R1}|^2 |a_{R2}|^2\rangle}{a_o^4}\right] dx'dy' - T}{T} \qquad \text{II-2.11}$$

Noting that

$$T^2 = \left[\iint_{-\infty}^{\infty} T_L^2 \, dx\,dy\right]^2 = \iint_{-\infty}^{\infty}\iint_{-\infty}^{\infty} T_{L1}^2 T_{L2}^2 \, dx_1 dy_1 dx_2 dy_2$$

$$= \iint_{-\infty}^{\infty} \left[\iint_{-\infty}^{\infty} T_{L1}^2 T_{L2}^2 \, dx\,dy\right] dx'dy'$$

$$= T \iint_{-\infty}^{\infty} D_o \, dx'dy', \qquad \text{II-2.12}$$

we see that T can also be represented as

$$T = \iint_{-\infty}^{\infty} D_o \, dx'dy'. \qquad \text{II-2.13}$$

Therefore we obtain

$$V^2 = \frac{\iint_{-\infty}^{\infty} D_o \left[\frac{\langle |a_{R1}|^2 |a_{R2}|^2 \rangle}{a_0^4} - 1 \right] dx'dy'}{\iint_{-\infty}^{\infty} D_o \, dx'dy'}$$

$$= \frac{\iint_{-\infty}^{\infty} D_o \, B_s \, dx'dy'}{\iint_{-\infty}^{\infty} D_o \, dx'dy'}, \qquad \text{II - 2.14}$$

where B_s is exclusively dependent on the random part of the pupil function and D_o is exclusively dependent on the non-random part.

Let us look more closely at the function D_o, (eq. II-2.9) which we shall call the core function. In the absence of apodization, T_L merely describes the boundary of the pupil, and is assumed to be constant within the pupil. The integrand $T_{L1}^2 T_{L2}^2$ then describes two overlapping pupils displaced with respect to each other, and is identically zero everywhere except in the region they hold in common. Therefore we see that, in the absence of apodization, the core function D_o is

identical to the aberration-free transfer function of the optical system, whether the optical system is free of aberrations or not. This is of course reasonable, since the aberrations affect the distribution of power in the image and not the total power itself.

When apodization is present, the identification of the core function with a form of transfer function is lost, except in a qualitative sense. However, it should be noted that the integrations indicated in eq. II-2.14 can be taken as over the spatial frequency domain of the system transfer function.

Now let us consider the function

$$B_s = \frac{\langle |a_{R1}|^2 |a_{R2}|^2 \rangle}{a_0^4} - 1 . \qquad \text{II-2.15}$$

First let us note that

$$\langle |a_{R1}|^2 |a_{R2}|^2 \rangle = \langle a_0^2 e^{2\alpha_1} a_0^2 e^{2\alpha_2} \rangle = a_0^4 \langle e^{2(\alpha_1 + \alpha_2)} \rangle , \qquad \text{II-2.16}$$

Now if $p(\alpha)$ is Gaussian, then $p(\alpha_1+\alpha_2)$ is also Gaussian (see appendix II), and we obtain (see appendix III)

$$\langle e^{2(\alpha_1+\alpha_2)}\rangle = e^{4\langle\alpha\rangle}\langle e^{2(\alpha_1'+\alpha_2')}\rangle$$

$$= e^{4\sigma_\alpha^2\rho_\alpha} \qquad \text{II-2,17}$$

so

$$B_s = e^{4\sigma_\alpha^2\rho_\alpha} - 1, \qquad \text{II-2,18}$$

which is a very simple function of the statistical properties of the incident wavefront, and is of course independent of its phase variation. It is a function of (x',y') because the autocorrelation function ρ_α is a function of (x',y'). Note that B_s is a maximum when ρ_α is a maximum, and goes to zero when ρ_α goes to zero. In other words, the shape of B_s is similar to that of ρ_α. In fact, if we normalize B_s to its value at (0,0), that is,

$$B_{SN} = \frac{B_s(x', y')}{B_s(0,0)} = \frac{e^{4\sigma_\alpha^2 \rho_\alpha} - 1}{e^{4\sigma_\alpha^2} - 1} \qquad \text{II-2.17}$$

then we find that B_{SN} approaches ρ_α as σ_α^2 approaches zero. Figure II-2 shows how B_{SN} varies with ρ_α for several assumed values of σ_α^2.

We see, therefore, that the squared coefficient of variation can be obtained as a normalized, weighted integral of the core function D_α, which is similar if not identical to the aberration-free transfer function of the optical system (in any case, it covers the same range in (x', y')), and where the weighting function B_s, a simple function of the amplitude disturbances, is similar to the correlation function ρ_α, having the same scale in (x', y') as ρ_α. Note in particular that, if the transfer function of the optical system has a finite limit in (x', y'), as it must in any real optical system, the core function has the same limit so that the integrals involved need by taken only over a finite range.

In examining the effect of the optical system on the

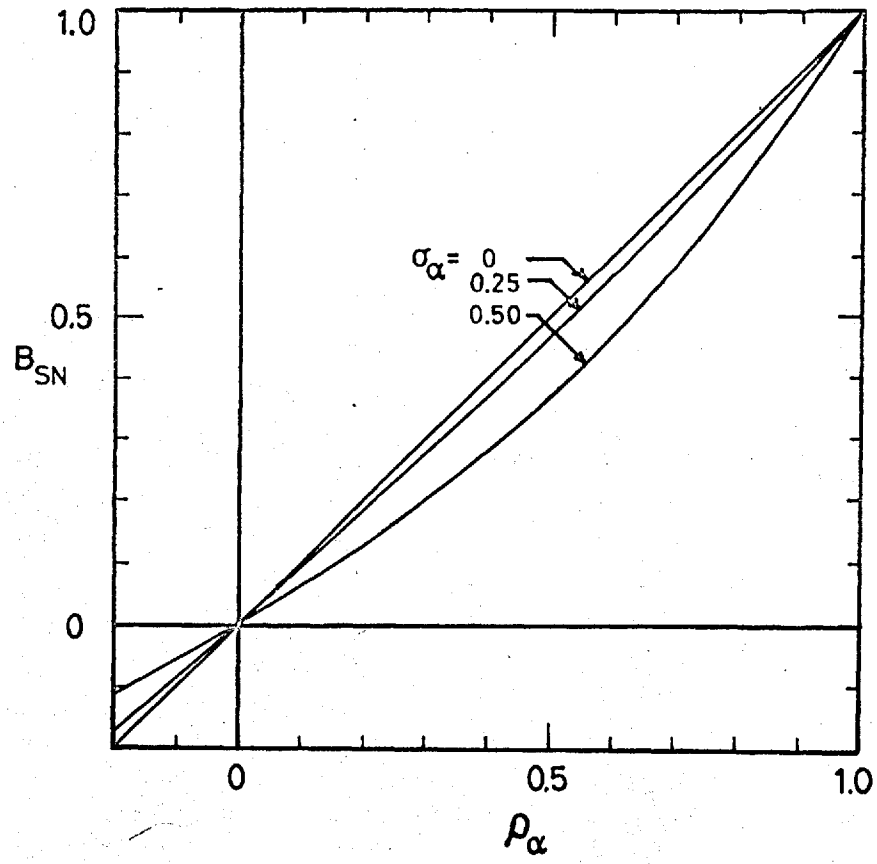


Figure II-2. Variation of the normalized weighting function B_{SN} with the autocorrelation function ρ_α for several values of σ_α (see eq. II-2.19).

intensity fluctuations in the image we must first be concerned with the relative scale of the functions D_o and B_s. Because of its usual circular symmetry and sharp outer boundary, the radius of the pupil of the optical system can usually be used to represent its characteristic length.

The characteristic length of the weighting function B_s or the autocorrelation function ρ_x is not so clearly distinguished because of the general lack of common identifying features for the class of possible functions. In fact, there are two scale factors generally used in turbulence theory,[21] the microscale, and the integral scale. The microscale is determined by the size of the smallest eddies formed, and the integral scale is determined by the larger eddies, the sizes of which depend on the boundary conditions of the situation in which turbulence is created. In the present context the integral scale would be the radius of the cylinder of unit height which would contain the same volume as the correlation function under consideration.

If we assume a continuously varying index gradient in the medium and a minimum eddy size, then any possible autocorrelation function must have a zero derivative at the origin and must be parabolic in the vicinity of the origin. One way of defining the microscale in terms of the properties of the autocorrelation function is the radius at which the function ceases to be parabolic, although this is necessarily vague because the transition is gradual. Another way would be to use a measure of the parabola, such as the radius at which the parabolic approximation would drop to 1/2 (the radius at which the parabolic approximation would drop to zero would generally be a little too large to be consistent with the meaning of microscale). A third way would be to use a characteristic length of a function which is used to generate the form of the autocorrelation function.

For the functions we will use for illustrative purposes in this thesis we will only use a measure of the microscale and neglect explicit discussion of the integral scale, because we shall use simple mathematical models

in which the two are strictly related.

The ratio of the characteristic length Λ of the pupil function to the characteristic length of the autocorrelation function ρ_α is the relative scale factor $\mathcal{R}$. The manner in which the observable scintillation varies with aperture size is described by the dependence of V on $\mathcal{R}$. Let us consider two limiting cases.

As $\mathcal{R} \to 0$ the range of integration is limited to the central peak of the autocorrelation function ρ_α, which approaches unity, and $B_s \to e^{4\sigma_\alpha^2} - 1$, which is independent of (x', y') and can therefore be taken outside the integral so that

$$V^2 \longrightarrow V_0^2 = e^{4\sigma_\alpha^2} - 1 . \qquad \text{II-2.20}$$

The manner in which V_0 varies with σ_α is shown in fig. II-3.

The general expression for V^2 can therefore be written as

$$V^2 = V_0^2 \frac{\iint_{-\infty}^{\infty} D_0 B_{SN}\, dx'dy'}{\iint_{-\infty}^{\infty} D_0\, dx'dy'} = V_0^2 C_S^2 , \qquad \text{II-2.21}$$

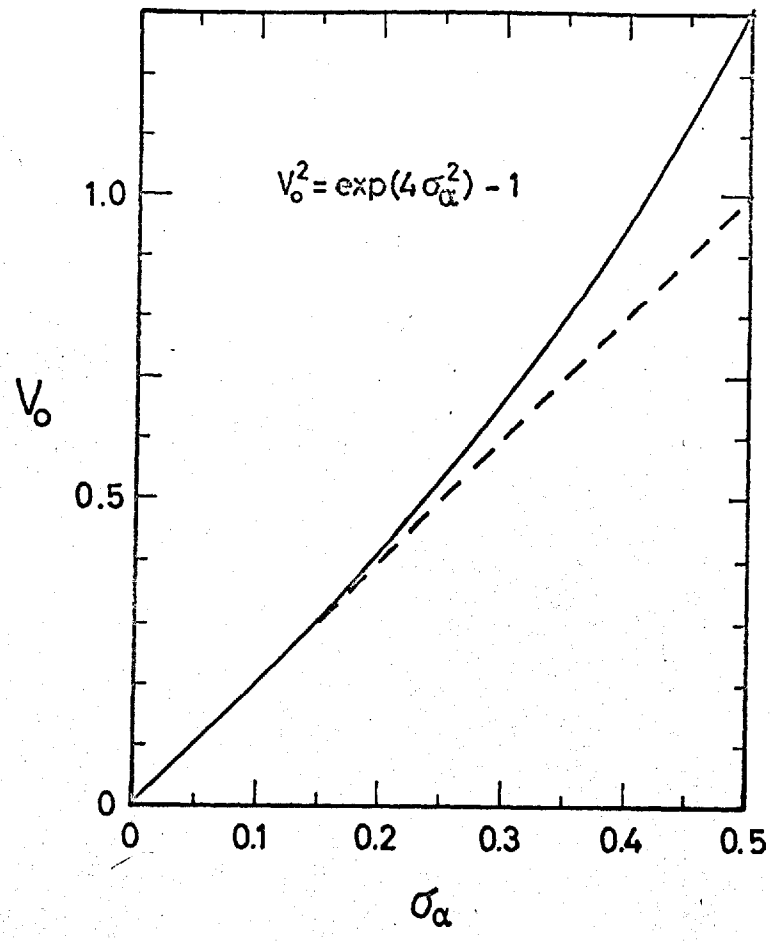


Figure II-3. Dependence of the coefficient of variation for scintillation on the standard deviation of α.

The function C_s as a function of R describes the manner in which the scintillation varies with the size of the pupil, and is practically independent of σ_α for small values of σ_α (see figure II-2).

In the case where $R \to \infty$, the range of integration is essentially limited by the range of ρ_α, and D_σ approaches unity over this range. The volume represented by the upper integral then approaches the volume under the function B_{SN} (approximately the volume under ρ_α, although slightly less), but diminishing at a rate of $1/R^2$. Now if we assume that the volume under ρ_α is finite, then V must fall off at least as fast as $1/R$ for large values of R.

A hypothetical illustrative example will serve to show the manner in which V depends on R and at the same time demonstrate the effects of a central obscuration and of apodization.

The assumed autocorrelation function ρ_α and its equation are shown in figure II-4. It is assumed to be isotropic. The special property characterizing this

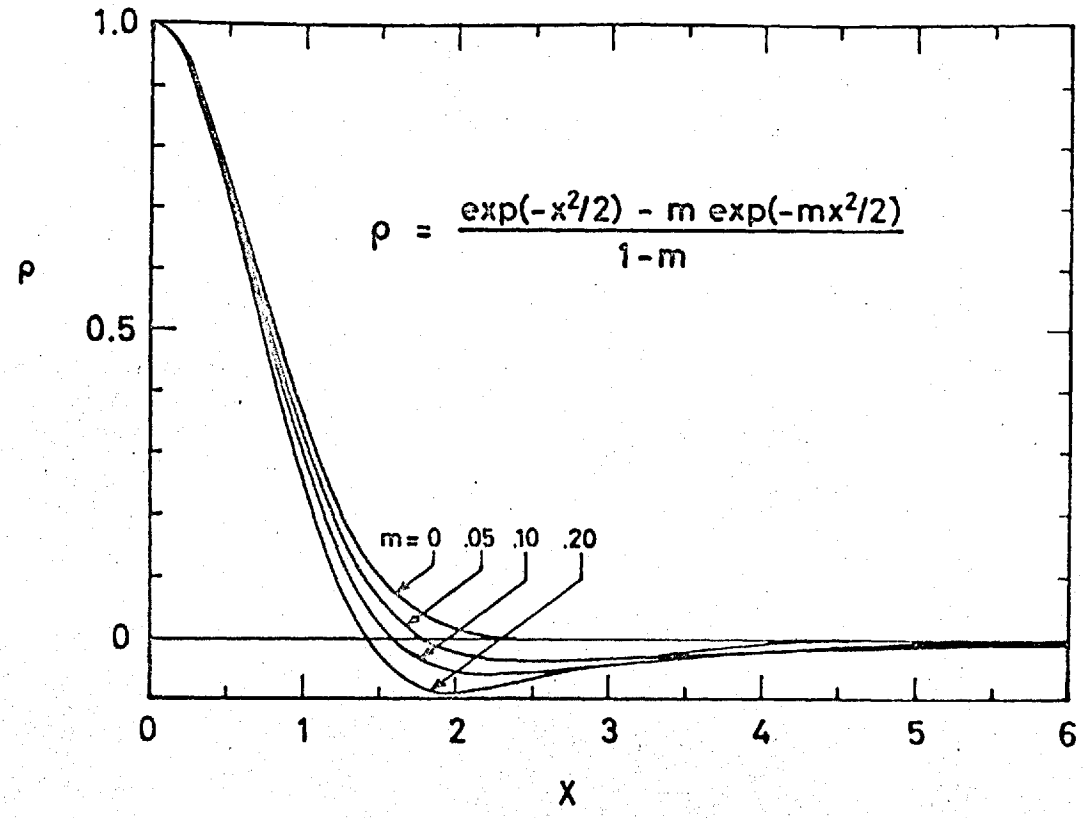


Figure II-4. Family of hypothetical autocorrelation functions for α .

function is that its integral is zero. Tatarski[14] claims that this is a necessary property, but he bases his argument on the conservation of energy, whereas we have shown that the conservation of energy is satisfied by merely assuming a non-zero mean for α. However, his experimental determination of ρ_α indicates the dip below zero which is necessary if the integral is to be zero, so we shall also assume that he is correct. The value of the parameter m for the curve in figure II-4 which most nearly conforms to his experimental determination is 0.1, and this is assumed in the remainder of the illustration.

The core functions which we shall assume are those for a circular aperture, annular apertures with obscuration ratios of 0.5 and 0.9, and effectively Gaussian apodization. The core functions, shown in figure II-5, are all assumed to have the maximum radius of the aperture as the characteristic length.

The results are shown in figure II-6, where the aperture scale factor is the ratio of the radius of the

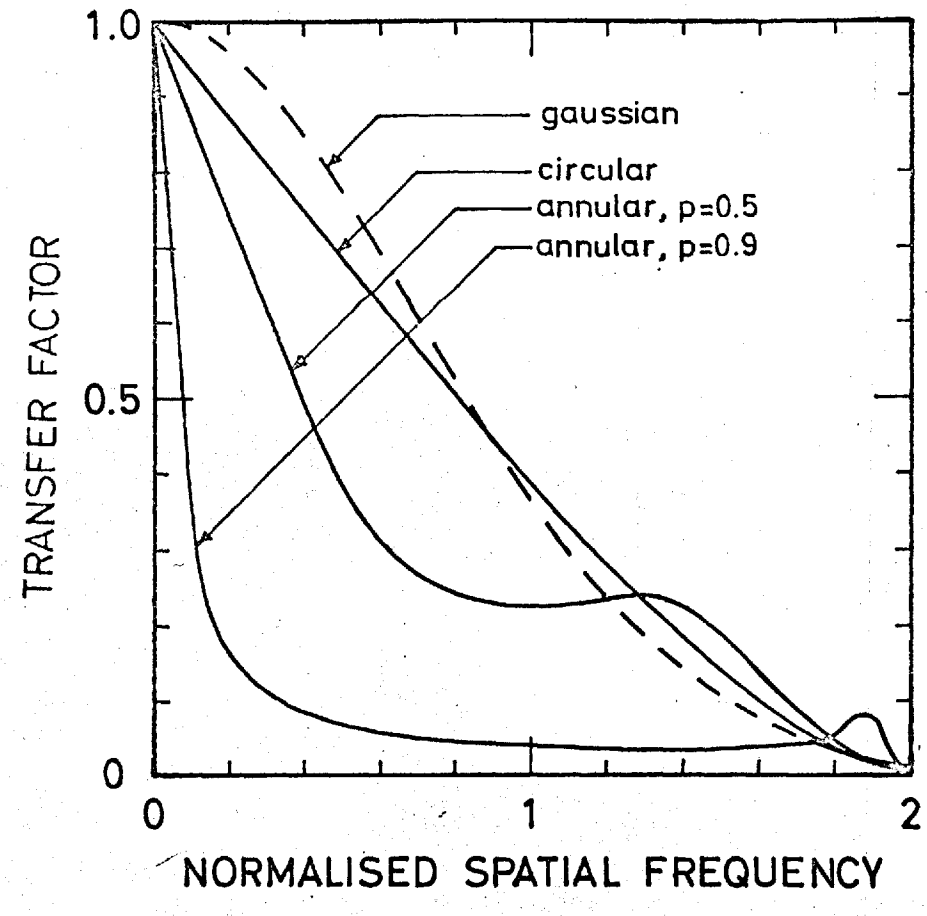


Figure II-5. Core functions for several types of aperture, all having circular symmetry and equal maximum dimensions.

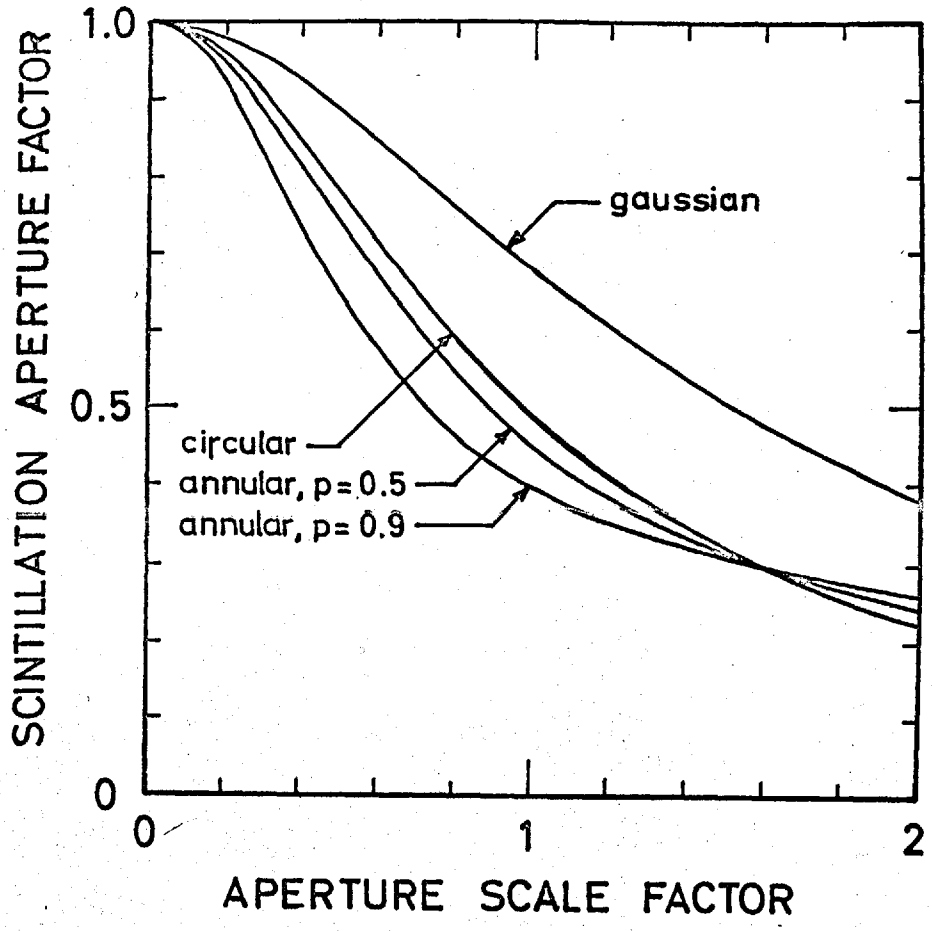


Figure II-6. Variation in scintillation with aperture size for the several types of aperture shown in Figure II-5.

aperture to the standard deviation of the Gaussian function used to generate ρ_α , and the scintillation aperture factor is C_s. It can be seen that scintillation is reduced by using an annular aperture (at least if the radius of the aperture is of the order of the microscale of ρ_α or less), but the reduction is small for obscuration ratios up to 0.5 and only moderate for ratios as great as 0.9 whereas the amount of light received is 75% and 19% respectively relative to the unobscured circular aperture, so that the deliberate obstruction of the circular aperture to reduce the observed scintillation is not generally to be recommended. On the other hand, apodizing the aperture results in a significant increase in the scintillation.

It is interesting to observe how well the assumed function fits observed experimental data. Figure II-7 shows observations made by Protheroe[20] averaged for summer and winter along with the best fitting curves. It can be seen that the general qualitative character is confirmed, but that the experimental data does not fall off

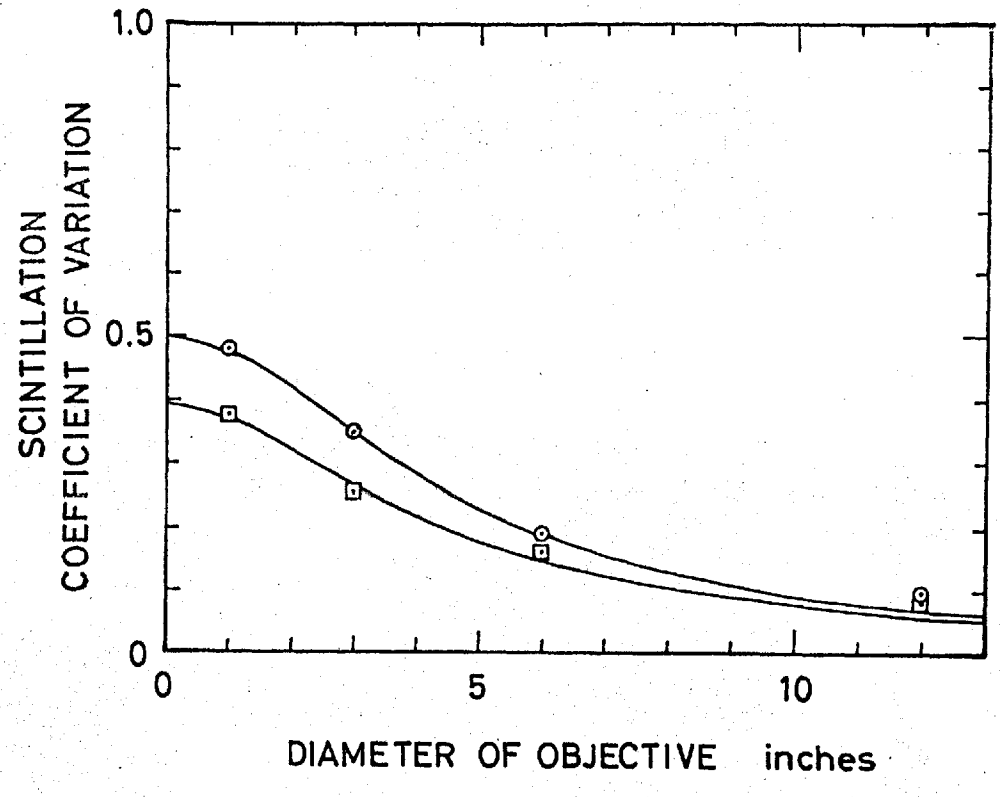


Figure II-7. Best fit of the model to experimental data reported by Protheroe.[20] The general character of the model is in agreement with the experimental observations.

as rapidly as $1/\kappa$. This is undoubtedly due to the fact that the macroscale of the turbulence must be larger than the maximum reported aperture of 12 inches and the experimental data does not go out far enought to indicate the limiting slope. The apparent discrepancy is also compounded by the fact that each point represents the average of a number of observations in which the scale of the turbulence may have varied appreciably. However, it is unlikely that the function used to describe ρ_α in the illustration is a really good representation. It was used primarily because it is a simple two-parameter function which has the approximate characteristics to be expected.

CHAPTER II-3. AGITATION

Another observable property of an image formed by an optical system looking through a turbulent medium is that the image appears to dance about, at least for small apertures. As in the case of scintillation, this property, which we shall call agitation, also decreases as the size of the aperture increases.[18] A measure of agitation is the radius of gyration (r.m.s. deviation) of the centroid of the image from its mean position.

In order to develop an expression for the agitation, let us first develop the relationship between the position of the centroid of the instantaneous image and the pupil function. We shall do this by using as an intermediary the transfer function of the system.

First consider the relationship between the transfer function $\tilde{\Phi}$ and the spread function Φ :

$$\tilde{\Phi}(x',y') = \frac{\iint_{-\infty}^{\infty} \Phi(u,v)\, e^{-2\pi i(ux'+vy')}\, du\, dv}{\iint_{-\infty}^{\infty} \Phi(u,v)\, du\, dv}$$

$$= \frac{1}{T} \iint_{-\infty}^{\infty} \Phi(u,v)\, e^{-2\pi i(ux'+vy')}\, du\, dv . \qquad \text{II - 3.1}$$

If we differentiate $\tilde{\Phi}$ with respect to x', we obtain

$$\frac{\partial \tilde{\Phi}}{\partial x'} = -\frac{2\pi i}{T} \iint_{-\infty}^{\infty} u\, \Phi(u,v)\, e^{-2\pi i(ux'+vy')}\, du\, dv. \qquad \text{II} - 3,2$$

Evaluating the derivative at $x', y' = 0$ we obtain

$$\left.\frac{\partial \tilde{\Phi}}{\partial x'}\right|_{x',y'=0} = -\frac{2\pi i}{T} \iint_{-\infty}^{\infty} u\, \Phi(u,v)\, du\, dv, \qquad \text{II} - 3,3$$

but the u-coordinate of the centroid is

$$\bar{u} = \frac{1}{T} \iint_{-\infty}^{\infty} u\, \Phi(u,v)\, du\, dv, \qquad \text{II} - 3,4$$

so we obtain

$$\bar{u} = -\frac{1}{2\pi i} \left.\frac{\partial \tilde{\Phi}}{\partial x'}\right|_{x',y'=0}. \qquad \text{II} - 3,5$$

The same argument holds for the v-coordinate:

$$\bar{v} = -\frac{1}{2\pi i} \left.\frac{\partial \tilde{\Phi}}{\partial y'}\right|_{x',y'=0}. \qquad \text{II} - 3,6$$

Now consider the relationship between the transfer function $\tilde{\Phi}$ and the pupil function $f(x,y) = \tau' e^{i\kappa W'}$, where $\tau' = \tau_L e^{\alpha}$, $W' = W_L + W_R$, and $W_R = \beta/\kappa = \beta/(2\pi/\lambda)$:

$$\tilde{\Phi}(x', y') = \frac{\iint_{-\infty}^{\infty} f(x + \frac{x'}{2}, y + \frac{y'}{2}) f^*(x - \frac{x'}{2}, y - \frac{y'}{2}) \, dx \, dy}{\iint_{-\infty}^{\infty} |f|^2 dx \, dy}$$

$$= \frac{1}{T} \iint_{-\infty}^{\infty} f_1 f_2^* \, dx \, dy \qquad \text{II - 3.8}$$

Differentiating, we obtain

$$\frac{\partial \tilde{\Phi}}{\partial x'} = \frac{1}{T} \iint_{-\infty}^{\infty} \left(f_1 \frac{\partial f_2^*}{\partial x'} + f_2^* \frac{\partial f_1}{\partial x'} \right) dx \, dy \qquad \text{II - 3.9}$$

Noting that $\frac{\partial f_1}{\partial x'} = \frac{1}{2} \frac{\partial f_1}{\partial x}$ and $\frac{\partial f_2^*}{\partial x'} = -\frac{1}{2} \frac{\partial f_2^*}{\partial x}$,

$$\frac{\partial \tilde{\Phi}}{\partial x'} = \frac{1}{2T} \iint_{-\infty}^{\infty} \left(f_2^* \frac{\partial f_1}{\partial x} - f_1 \frac{\partial f_2^*}{\partial x} \right) dx \, dy$$

$$= \frac{1}{2T} \iint_{-\infty}^{\infty} e^{i\kappa(W_1' - W_2')} \left[\tau_2' \frac{\partial \tau_1'}{\partial x} - \tau_1' \frac{\partial \tau_2'}{\partial x} + i\kappa \tau_1' \tau_2' \left(\frac{\partial W_1'}{\partial x} + \frac{\partial W_2'}{\partial x} \right) \right] dx \, dy .$$

II - 3.10

Evaluating at $x', y' = 0$,

$$\left.\frac{\partial \tilde{\mathcal{E}}}{\partial x'}\right|_{x',y'=0} = \frac{1}{2\mathcal{T}} \iint_{-\infty}^{\infty} 2\,i\,K\,\mathcal{T}'^{2} \frac{\partial W'}{\partial x}\,dx\,dy = \frac{2\pi i}{\lambda \mathcal{T}} \iint_{-\infty}^{\infty} \mathcal{T}'^{2} \frac{\partial W'}{\partial x}\,dx\,dy, \qquad \text{II} - 3,11$$

and correspondingly

$$\left.\frac{\partial \tilde{\mathcal{E}}}{\partial y'}\right|_{x',y'=0} = \frac{2\pi i}{\lambda \mathcal{T}} \iint_{-\infty}^{\infty} \mathcal{T}'^{2} \frac{\partial W'}{\partial y}\,dx\,dy, \qquad \text{II} - 3,12$$

and we finally obtain

$$\bar{u} = -\frac{1}{\lambda \mathcal{T}} \iint_{-\infty}^{\infty} \mathcal{T}'^{2} \frac{\partial W'}{\partial x}\,dx\,dy$$

$$\bar{v} = -\frac{1}{\lambda \mathcal{T}} \iint_{-\infty}^{\infty} \mathcal{T}'^{2} \frac{\partial W'}{\partial y}\,dx\,dy \qquad \text{II} - 3,13$$

as the required relationship between the coordinates of the centroid and the pupil function. It is of interest to note that, although these equations apply to the centroid of the diffraction image, exactly

the same equations apply to the geometrical image, so that even though the predicted image distributions are different for the two types of image, their centroids are the same.

We are not, however, interested in the absolute coordinates of the centroid but the deviations resulting from the random component of the pupil function. Noting that

$$\bar{u} = -\frac{1}{\lambda T}\iint_{-\infty}^{\infty} T_L^2\, e^{2\alpha}\left(\frac{\partial W_L}{\partial x} + \frac{\partial W_R}{\partial x}\right) dx\,dy \qquad \text{II} - 3.14$$

and

$$\bar{u}_L = -\frac{1}{\lambda T}\iint_{-\infty}^{\infty} T_L^2\, \frac{\partial W_L}{\partial x}\, dx\,dy, \qquad \text{II} - 3.15$$

we obtain

$$\bar{u}_R = \bar{u} - \bar{u}_L = -\frac{1}{\lambda T}\iint_{-\infty}^{\infty} T_L^2\, \frac{\partial W_L}{\partial x}\,(e^{2\alpha}-1)\,dx\,dy - \frac{1}{\lambda T}\iint_{-\infty}^{\infty} T_L^2\, \frac{\partial W_R}{\partial x}\, e^{2\alpha}\, dx\,dy$$
$$= A + B \qquad \text{II} - 3.16$$

and correspondingly

$$\bar{v}_R = \bar{v} - \bar{v}_L = -\frac{1}{\lambda T}\iint_{-\infty}^{\infty} T_L^2 \frac{\partial W_L}{\partial y}(e^{2\alpha}-1)\,dx\,dy - \frac{1}{\lambda T}\iint_{-\infty}^{\infty} T_L^2 \frac{\partial W_R}{\partial y} e^{2\alpha}\,dx\,dy$$
$$= C + D. \qquad \text{II} - 3.17$$

Before proceeding further let us verify that

$$\langle \bar{u}_R \rangle = \langle \bar{v}_R \rangle = 0 \qquad \text{II} - 3.18$$

First note that

$$\langle \bar{u}_R \rangle = -\frac{1}{\lambda T}\iint_{-\infty}^{\infty} T_L^2 \frac{\partial W_L}{\partial x}\left(\langle e^{2\alpha}\rangle - 1\right)dx\,dy - \frac{1}{\lambda T}\iint_{-\infty}^{\infty} T_L^2 \left\langle e^{2\alpha}\frac{\partial W_R}{\partial x}\right\rangle dx\,dy. \qquad \text{II} - 3.19$$

We have already demonstrated that $\langle e^{2\alpha}\rangle = 1$, so the first integral vanishes. Let us consider next the term

$$\left\langle e^{2\alpha}\frac{\partial W_R}{\partial x}\right\rangle = \frac{e^{-2\sigma_\alpha^2}}{K}\left\langle e^{2\alpha'}\frac{\partial \beta}{\partial x}\right\rangle. \qquad \text{II} - 3.20$$

If the joint probability density $p(\alpha', \beta)$ is known, then

$$\left\langle e^{2\alpha'}\beta \right\rangle = \iint_{-\infty}^{\infty} e^{2\alpha'}\beta\, p(\alpha', \beta)\, d\alpha\, d\beta, \qquad \text{II} - 3,21$$

If we assume a joint Gaussian distribution,

$$p(\alpha', \beta) = \frac{1}{2\pi\sigma_\alpha \sigma_\beta (1-\rho_{\alpha\beta}^2)^{1/2}} \exp\left[- \frac{\left(\frac{\alpha'}{\sigma_\alpha}\right)^2 + \left(\frac{\beta}{\sigma_\beta}\right)^2 - 2\rho_{\alpha\beta}\left(\frac{\alpha}{\sigma_\alpha}\right)\left(\frac{\beta}{\sigma_\beta}\right)}{2(1-\rho_{\alpha\beta}^2)} \right]$$

$$\text{II} - 3,22$$

then

$$e^{-2\sigma_\alpha^2} \left\langle e^{2\alpha'}\beta \right\rangle = 2\sigma_\alpha \sigma_\beta \rho_{\alpha\beta}. \qquad \text{II} - 3,23$$

Noting that

$$\frac{\partial \beta}{\partial x} = \lim_{\Delta x \to 0} \frac{\beta\left(x + \frac{\Delta x}{2}\right) - \beta\left(x - \frac{\Delta x}{2}\right)}{\Delta x}, \qquad \text{II} - 3,24$$

then

$$\left\langle e^{2\alpha \frac{\partial W_R}{\partial x}} \right\rangle = \frac{1}{K} \lim_{\Delta x \to 0} \frac{e^{-2\sigma_\alpha^2} \left\langle e^{2\alpha' \beta(x+\frac{\Delta x}{2})} \right\rangle - e^{-2\sigma_\alpha^2} \left\langle e^{2\alpha' \beta(x-\frac{\Delta x}{2})} \right\rangle}{\Delta x}$$

$$= \frac{1}{K} \lim_{\Delta x \to 0} \frac{2\sigma_\alpha \sigma_\beta \left[\rho_{\alpha\beta}\left(\frac{\Delta x}{2}\right) - \rho_{\alpha\beta}\left(\frac{\Delta x}{2}\right) \right]}{\Delta x}$$

$$\equiv 0, \qquad \text{II} - 3.2.5$$

and the second integral vanishes, verifying that

$$\langle \bar{u}_R \rangle = \langle \bar{v}_R \rangle = 0 \qquad \text{II} - 3.26$$

Returning now to the problem of the radius of gyration, we note that the square of the radius of gyration is given by

$$\sigma_M^2 = \frac{1}{2} \langle \bar{u}_R^2 + \bar{v}_R^2 \rangle, \qquad \text{II} - 3.27$$

and

$$\bar{u}_R^2 = (A+B)^2 = A^2 + B^2 + 2AB$$
$$\bar{v}_R^2 = (C+D)^2 = C^2 + D^2 + 2CD \qquad \text{II} - 3.28$$

so

$$\bar{u}_R^2 + \bar{v}_R^2 = (A^2 + C^2) + (B^2 + D^2) + 2(AB + CD), \qquad \text{II} - 3.29$$

and

$$\sigma_M^2 = \frac{1}{2}\left[\langle A^2 + C^2 \rangle + \langle B^2 + D^2 \rangle + 2\langle AB + CD \rangle\right], \qquad \text{II} - 3.30$$

Using the same technique for converting a squared integral into a double integral that we used in the section on scintillation, we obtain

$$A' = \frac{1}{\lambda^2 T^2}\iint_{-\infty}^{\infty}\iint_{-\infty}^{\infty} T_{L1}^2\, T_{L2}^2\, \frac{\partial W_{L1}}{\partial x}\,\frac{\partial W_{L2}}{\partial x}\,(e^{2\alpha_1}-1)(e^{2\alpha_2}-1)\,dx\,dy\,dx'\,dy',$$

$$C' = \frac{1}{\lambda^2 T^2}\iint_{-\infty}^{\infty}\iint_{-\infty}^{\infty} T_{L1}^2\, T_{L2}^2\, \frac{\partial W_{L1}}{\partial y}\,\frac{\partial W_{L2}}{\partial y}\,(e^{2\alpha_1}-1)(e^{2\alpha_2}-1)\,dx\,dy\,dx'\,dy',$$

$$A^2+C^2 = \frac{1}{\lambda^2 T^2}\iint_{-\infty}^{\infty}\iint_{-\infty}^{\infty} T_{L1}^2\, T_{L2}^2\left(\frac{\partial W_{L1}}{\partial x}\,\frac{\partial W_{L2}}{\partial x}+\frac{\partial W_{L1}}{\partial y}\,\frac{\partial W_{L2}}{\partial y}\right)(e^{2\alpha_1}-1)(e^{2\alpha_2}-1)\,dx\,dy\,dx'\,dy'.$$

II - 3.31

Averaging,

$$\langle A^2+C^2\rangle = \frac{1}{\lambda^2 T^2}\iint_{-\infty}^{\infty}\left[\iint_{-\infty}^{\infty} T_{L1}^2\, T_{L2}^2\left(\frac{\partial W_{L1}}{\partial x}\,\frac{\partial W_{L2}}{\partial x}+\frac{\partial W_{L1}}{\partial y}\,\frac{\partial W_{L2}}{\partial y}\right)dx\,dy\right]\langle(e^{2\alpha_1}-1)(e^{2\alpha_2}-1)\rangle\,dx'\,dy',$$

II - 3.32

and in the same way,

$$\langle B^2+D^2\rangle = \frac{1}{\lambda^2 T^2}\iint_{-\infty}^{\infty}\left[\iint_{-\infty}^{\infty} T_{L1}^2\, T_{L2}^2\,dx\,dy\right]\left\langle e^{2(\alpha_1+\alpha_2)}\left(\frac{\partial W_{R1}}{\partial x}\,\frac{\partial W_{R2}}{\partial x}+\frac{\partial W_{R1}}{\partial y}\,\frac{\partial W_{R2}}{\partial y}\right)\right\rangle dx'\,dy',$$

II - 3.33

and

$$\langle AB+CD\rangle = \frac{1}{\lambda^2 T^2}\int_{-\infty}^{\infty}\int\left[\int_{-m}^{\infty}\int \tau_{11}\,\tau_{12}\,dx\,dy\right]\left\langle e^{2\alpha_2}\left(e^{2\alpha_1}-1\right)\left(\frac{\partial W_{11}}{\partial x}\frac{\partial W_{R2}}{\partial x}+\frac{\partial W_{11}}{\partial y}\frac{\partial W_{R2}}{\partial y}\right)\right\rangle dx'\,dy'.$$

II - 3.34

Consider the average within the last integral. It can be expanded as

$$\left\langle \frac{\partial W_{11}}{\partial x}\left(e^{2(\alpha_1+\alpha_2)}\frac{\partial W_{R2}}{\partial x}-e^{2\alpha_2}\frac{\partial W_{R2}}{\partial x}\right)+\frac{\partial W_{11}}{\partial y}\left(e^{2(\alpha_1+\alpha_2)}\frac{\partial W_{R2}}{\partial y}-e^{2\alpha_2}\frac{\partial W_{R2}}{\partial y}\right)\right\rangle$$

$$=\frac{\partial W_{11}}{\partial x}\left[\left\langle e^{2(\alpha_1+\alpha_2)}\frac{\partial W_{R2}}{\partial x}\right\rangle-\left\langle e^{2\alpha_2}\frac{\partial W_{R2}}{\partial x}\right\rangle\right]+\frac{\partial W_{11}}{\partial y}\left[\left\langle e^{2(\alpha_1+\alpha_2)}\frac{\partial W_{R2}}{\partial y}\right\rangle-\left\langle e^{2\alpha_2}\frac{\partial W_{R2}}{\partial y}\right\rangle\right].$$

II - 3.35

Noting that if α has a Gaussian probability density distribution, then $(\alpha_1+\alpha_2)$ has also (see Appendix II), and each of the four averages in the last expression is identically zero by the argument leading to equation II-3.25. Therefore $\langle AB+CD\rangle$ is zero, and

$$\sigma_M^2 = \langle A^2+C^2\rangle + \langle B^2+D^2\rangle \qquad \text{II - 3.36}$$

Let us consider each of these terms separately. The average within the integral of the first term becomes

$$\left\langle e^{2(\alpha_1+\alpha_2)} - e^{2\alpha_2} - e^{2\alpha_1} + 1 \right\rangle$$
$$= \left\langle e^{2(\alpha_1+\alpha_2)} \right\rangle - \left\langle e^{2\alpha_2} \right\rangle - \left\langle e^{2\alpha_1} \right\rangle + 1, \qquad \text{II} - 3.37$$

which, by the equations II-1.8 and II-2.17, becomes

$$e^{4\sigma_\alpha^2/\rho_\alpha} - 1 = B_S, \qquad \text{II} - 3.38$$

the same function obtained in the case of scintillation. Defining

$$D_0' \equiv \frac{\int\int_{-\infty}^{\infty} \tau_{L1}^2 \tau_{L2}^2 \left(\frac{\partial (W_{L1}/\lambda)}{\partial x} \frac{\partial (W_{L2}/\lambda)}{\partial x} + \frac{\partial (W_{L1}/\lambda)}{\partial y} \frac{\partial (W_{L2}/\lambda)}{\partial y} \right) dx\,dy}{\int\int_{-\infty}^{\infty} \tau_L^2\, dx\,dy}, \qquad \text{II} - 3.39$$

which is related to the function D_o obtained in the case of scintillation, we obtain for the first term

$$\langle A^2 + C^2 \rangle = \frac{\iint_{-\infty}^{\infty} D_o' B_1 \, dx'dy'}{\iint_{-\infty}^{\infty} D_o \, dx'dy'} \qquad \text{II} - 3,40$$

This term describes the effect of the interaction between the amplitude fluctuations and the properties of the optical system, and is independent of the random phase fluctuations. Note also that it vanishes if the optical system is free of aberrations.

Finally, let us consider the remaining term

$$\langle B^2 + D^2 \rangle = \frac{1}{\lambda^2 T^2} \iint_{-\infty}^{\infty} \left[\iint_{-\infty}^{\infty} T_{L1}^2 T_{L2}^2 \, dx dy \right] \langle e^{2(\alpha_1 + \alpha_2)} \left(\frac{\partial W_{R1}}{\partial x} \frac{\partial W_{R2}}{\partial x} + \frac{\partial W_{R1}}{\partial y} \frac{\partial W_{R2}}{\partial y} \right) \rangle \, dx'dy'.$$

$$\text{II} - 3,41$$

The integral within the square brackets is the numerator of the function D_o, and we can define the function

$$B_M = \left\langle e^{2(\nu_1+\nu_2)} \left(\frac{\partial(W_{R1}/\lambda)}{\partial x} \frac{\partial(W_{R2}/\lambda)}{\partial x} + \frac{\partial(W_{R1}/\lambda)}{\partial y} \frac{\partial(W_{R2}/\lambda)}{\partial y} \right) \right\rangle$$

II - 3.42

so that we obtain

$$\langle B^2 + D^2 \rangle = \frac{\iint_{-\infty}^{\infty} D_0 B_M \, dx' dy'}{\iint_{-\infty}^{\infty} D_0 \, dx' dy'} \qquad \text{II - 3.43}$$

which is independent of the aberrations of the optical system.

The general solution of the average represented by B_M involves a four-point correlation which is much too complicated to be included here. However, if we assume that the amplitude and phase functions are essentially uncorrelated, as is very often approximately the case when the larger part of the phase variations come about by turbulence near the aperture of the system, whereas the amplitude variations can only come about from distant turbulence, then

B_M factors into two parts, one dependent on the amplitude variations and the other on the phase variations:

$$B_M = \left\langle e^{2(\alpha_1+\alpha_2)} \right\rangle \left\langle \frac{\partial(W_{R1}/\lambda)}{\partial x}\frac{\partial(W_{R2}/\lambda)}{\partial x} + \frac{\partial(W_{R1}/\lambda)}{\partial y}\frac{\partial(W_{R2}/\lambda)}{\partial y} \right\rangle. \qquad \text{II} - 3.44$$

The first factor, which is familiar by now, reduces to

$$\left\langle e^{2(\alpha_1+\alpha_2)} \right\rangle = e^{4\sigma_\alpha^2 \rho_\alpha}. \qquad \text{II} - 3.45$$

The second factor also reduces to a simple expression as follows. In terms of β it becomes

$$\frac{1}{4\pi^2}\left\langle \frac{\partial\beta_1}{\partial x}\frac{\partial\beta_2}{\partial x} + \frac{\partial\beta_1}{\partial y}\frac{\partial\beta_2}{\partial y} \right\rangle, \qquad \text{II} - 3.46$$

and noting that

$$\frac{\partial\beta_1}{\partial x} = \lim_{\Delta x \to 0} \frac{\beta\left(x + \frac{x'}{2} + \frac{\Delta x}{2}\right) - \beta\left(x + \frac{x'}{2} - \frac{\Delta x}{2}\right)}{\Delta x} \qquad \text{II} - 3.47$$

and

$$\frac{\partial \beta_2}{\partial x} = \lim_{\Delta x \to 0} \frac{\beta\left(x - \frac{x'}{2} + \frac{\Delta x}{2}\right) - \beta\left(x - \frac{x'}{2} - \frac{\Delta x}{2}\right)}{\Delta x} \qquad \text{II} - 3,48$$

we obtain

$$\left\langle \frac{\partial \beta_1}{\partial x} \frac{\partial \beta_2}{\partial x} \right\rangle = \lim_{\Delta x \to 0} \frac{\left\langle \left[\beta\left(x + \frac{x'}{2} + \frac{\Delta x}{2}\right) - \beta\left(x + \frac{x'}{2} - \frac{\Delta x}{2}\right)\right]\left[\beta\left(x - \frac{x'}{2} + \frac{\Delta x}{2}\right) - \beta\left(x - \frac{x'}{2} - \frac{\Delta x}{2}\right)\right]\right\rangle}{(\Delta x)^2}$$

$$= \lim_{\Delta x \to 0} \frac{1}{(\Delta x)^2}\left[\left\langle \beta\left(x + \frac{x'}{2} + \frac{\Delta x}{2}\right)\beta\left(x - \frac{x'}{2} + \frac{\Delta x}{2}\right)\right\rangle + \left\langle \beta\left(x + \frac{x'}{2} - \frac{\Delta x}{2}\right)\beta\left(x - \frac{x'}{2} - \frac{\Delta x}{2}\right)\right\rangle \right.$$
$$\left. - \left\langle \beta\left(x + \frac{x'}{2} + \frac{\Delta x}{2}\right)\beta\left(x - \frac{x'}{2} - \frac{\Delta x}{2}\right)\right\rangle - \left\langle \beta\left(x + \frac{x'}{2} - \frac{\Delta x}{2}\right)\beta\left(x - \frac{x'}{2} + \frac{\Delta x}{2}\right)\right\rangle\right]$$

$$= \lim_{\Delta x \to 0} \frac{1}{(\Delta x)^2}\left[\sigma_\beta^2 \rho_\beta(x') + \sigma_\beta^2 \rho_\beta(x') - \sigma_\beta^2 \rho_\beta(x' - \Delta x) - \sigma_\beta^2 \rho_\beta(x' + \Delta x)\right]$$

$$= \sigma_\beta^2 \lim_{\Delta x \to 0} \frac{1}{(\Delta x)^2}\left[2\rho_\beta(x') - \left\{\rho_\beta(x' - \Delta x) + \rho_\beta(x' + \Delta x)\right\}\right], \qquad \text{II} - 3,49$$

Expanding the last two terms in a Taylor's series

we obtain

$$\left\langle \frac{\partial \beta_1}{\partial x} \frac{\partial \beta_2}{\partial x} \right\rangle = \sigma_\beta^2 \lim_{\Delta x \to 0} \frac{1}{(\Delta x)^2}\left[2\rho_\beta(x') - \left\{2\rho_\beta(x') + (\Delta x)^2 \frac{\partial^2 \rho_\beta(x')}{\partial x^2} + \text{higher order terms}\right\}\right]$$

$$= -\sigma_\beta^2 \frac{\partial^2 \rho_\beta(x')}{\partial x^2}, \qquad \text{II} - 3,50$$

and similarly for η , so

$$\left\langle \frac{\partial(W_{R1}/\lambda)}{\partial x}\frac{\partial(W_{R2}/\lambda)}{\partial x} + \frac{\partial(W_{R1}/\lambda)}{\partial y}\frac{\partial(W_{R2}/\lambda)}{\partial \eta} \right\rangle = -\frac{\sigma_\alpha^2}{4\pi^2}\left(\frac{\partial^2\rho_\alpha}{\partial x^2}+\frac{\partial^2\rho_\alpha}{\partial y^2}\right) = -\frac{\sigma_\alpha^2}{4\pi^2}\nabla^2\rho_\alpha . \qquad \text{II} - 3,51$$

In terms of the wavefront fluctuation,

$$-\frac{\sigma_\alpha^2}{4\pi^2}\nabla^2\rho_\alpha = -\left(\frac{\sigma_w}{\lambda}\right)^2\nabla^2\rho_w . \qquad \text{II} - 3,52$$

The complete expression for $\mathcal{B}_M$ then becomes

$$\mathcal{B}_M = -\left(\frac{\sigma_w}{\lambda}\right)^2 e^{4\sigma_\alpha^2\rho_\alpha}\nabla^2\rho_w . \qquad \text{II} - 3,53$$

For illustrative purposes we shall assume that only phase variations are present. Then

$$\sigma_M^2 = -\frac{1}{2}\left(\frac{\sigma_w}{\lambda}\right)^2 \frac{\iint_{-\infty}^{\infty} D_o\,\nabla^2\rho_w\,dx'dy'}{\iint_{-\infty}^{\infty} D_o\,dx'dy'} \qquad \text{II} - 3,54$$

Assuming that ρ_w is parabolic near the origin, it may be represented in the vicinity of the origin by

$$\rho_w \approx 1 - \frac{1}{2}R^2(x'^2+y'^2) \qquad \text{II} - 3,55$$

As in the case of scintillation, $\mathcal{R}$ is the relative scale factor, the ratio of the radius of the pupil function, assumed circular in perimeter, to the microscale factor of the autocorrelation function, which is the value of x' and y' for which the parabolic approximation for ρ_w drops to a value of 1/2. We then obtain, in the vicinity of the origin,

$$\nabla^2 \rho_w = -2\mathcal{R}^2, \qquad \text{II} - 3.56$$

which is its peak value at the origin. Normalizing $\nabla^2 \rho_w$, we obtain

$$\sigma_M^2 = \mathcal{R}^2 \left(\frac{\sigma_w}{\lambda}\right)^2 \frac{\iint_{-\infty}^{\infty} D_o B_{MN}\, dx'dy'}{\iint_{-\infty}^{\infty} D_o\, dx'dy'} = \mathcal{R}^2 \left(\frac{\sigma_w}{\lambda}\right)^2 C_M^2, \qquad \text{II} - 3.57$$

where

$$B_{MN} = -\frac{\nabla^2 \rho_w}{2\mathcal{R}^2}. \qquad \text{II} - 3.58$$

Now σ_M is measured in the same units as u and v, which are reduced from actual spatial coordinates by the factor $h/f\lambda$, where h is the radius of the

aperture and f the distance to the image plane. It can be expressed in angular units by multiplying by λ/d:

$$\sigma_{MA} = \left(\frac{\sigma_W}{d/f}\right) C_M = \left(\frac{\sigma_W}{r_R}\right) C_M \qquad \text{II-3,59}$$

and noting that C_M goes to unity for vanishingly small apertures, the maximum angular radius of gyration of the image is given by the ratio of the standard deviation of the wavefront disturbances to the microscale of the disturbances. We note that, if we assume a microscale of about 25 mm., we get a radius of gyration of 1 arc-second for $\sigma_W = 1.25 \times 10^{-4}$ mm., or about $\frac{1}{4}\lambda$ of visible light, which is of the right order of magnitude.

The manner in which the agitation diminishes as the scale factor $\mathcal{R}$ increases is described by the function C_M, and it is clear from its formal similarity to the function C_S for scintillation that its qualitative variation, and the effects of modifying

the transmission of the pupil by obscuration or apodization, is very similar to that obtained in the case of scintillation.

In order to provide a numerical example, we must assume a form for the autocorrelation function ρ_W. In the case of phase variations there is no reason to assume that ρ_W does not fall off monotonically. We could use the Gaussian function

$$\rho_W = e^{-\frac{1}{2}r'^2}, \quad r'^2 = \frac{1}{R_n^2}(x'^2+y'^2) \qquad \text{II} - 3,60$$

but this function falls off too quickly to be realistic. Another function which is very nearly as easy to handle is the hyperbolic secant:

$$\rho_W = \frac{2}{e^{r'}+e^{-r'}} = \operatorname{sech} r' \qquad \text{II} - 3,61$$

Both functions are shown in figure II-8. It is clear that the latter function is more desirable.

Its Laplacian is obtained as

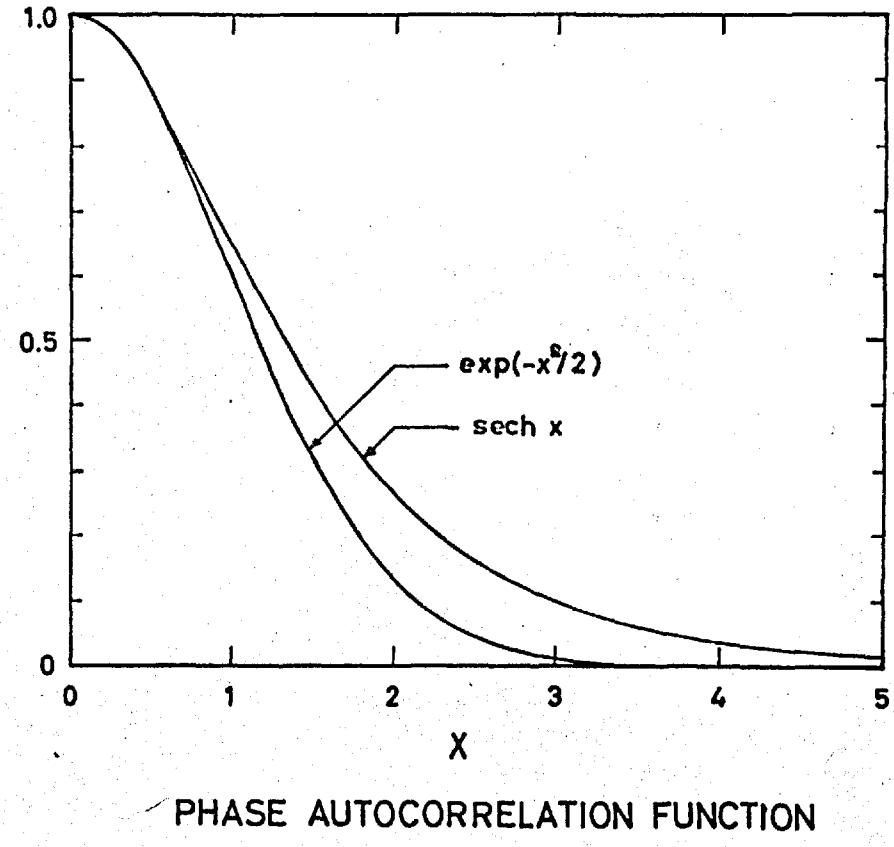


Figure II-8. Hypothetical phase autocorrelation functions.

$$\nabla^2 \rho_w = \operatorname{sech} r' \left\{ 1 - 2 \operatorname{sech} r' - \operatorname{sech} r' \left(\frac{\sinh r'}{r'} \right) \right\}, \quad \text{II-3.62}$$

and by substituting this in the integral and evaluating C_M as a function of $\mathcal{R}$ for the various pupil functions we used in the case of scintillation, we obtain the curves shown in figure II-7. The expectation that the dependence of agitation on the geometry of the pupil would be similar to that of scintillation is borne out (see fig. II-6).

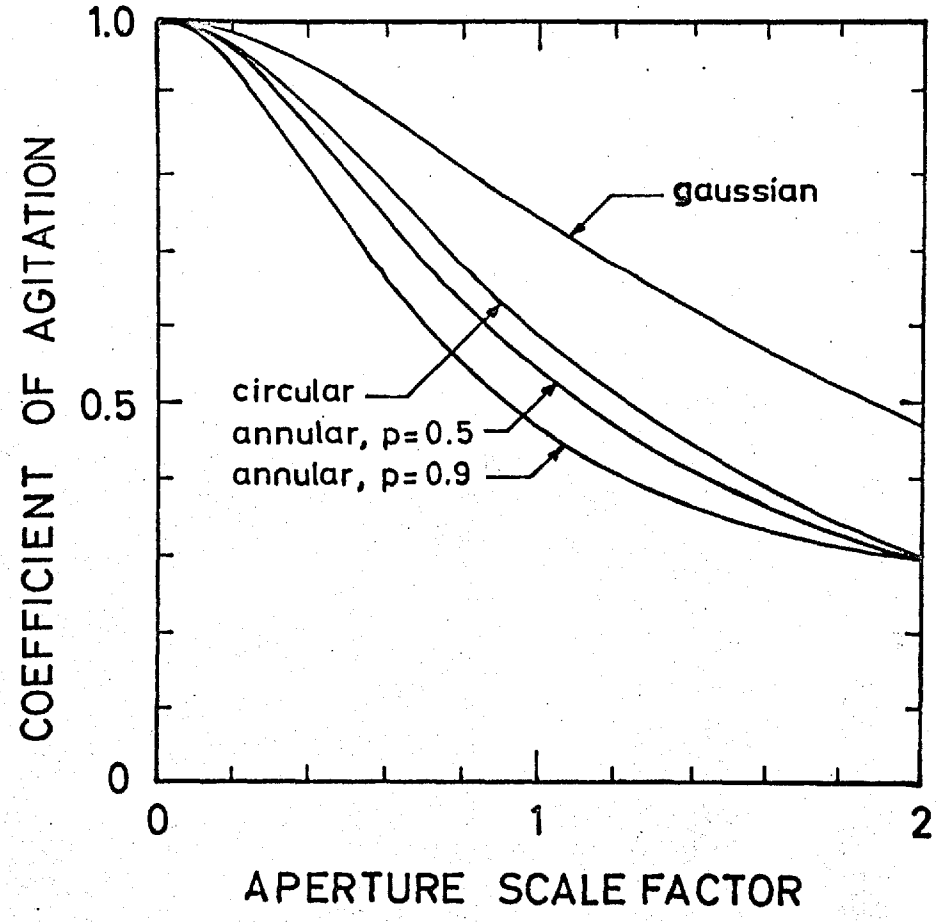


Figure II-9. Variation in agitation with aperture size for several types of aperture.

CHAPTER II-4. BLUR

The third phenomenon observed in stellar images obtained when the observation is through a turbulent medium is that the form of the image is distorted, and the average, or time-integrated, image is blurred by the random amplitude and phase structure of the wavefront incident on the optical system. Other investigators [13,14] have tried, usually with some awkwardness, to obtain the blur effects by calculating the properties of the spread function obtained. Difficulties have usually arisen because the process involves convolutions which are difficult to identify when an equation is being developed. We will show that by approaching the problem by calculating the effective transfer function, the expressions obtained easily factor into independent components, and indirectly the properties of the spread function are simply obtained.

Also, whereas scintillation and agitation are only significant when isolated star images are involved,

the spread function or transfer function describing the blur of the star image is directly applicable to the understanding of the images of extended objects.

First, let us consider the expression for the transfer function when the pupil function contains a random component:

$$\tilde{\Phi}(x',y') = \frac{1}{T}\iint_{-\infty}^{\infty} f_1 f_2^{*}\,dx\,dy = \frac{1}{T}\iint_{-\infty}^{\infty}\left(T_{L1}T_{L2}\,e^{iK(W_{L1}-W_{L2})}\right)\left(e^{(\alpha_1+\alpha_2)+i(\beta_1-\beta_2)}\right)dx\,dy. \qquad \text{II - 4.1}$$

Taking the average,

$$\langle\tilde{\Phi}(x',y')\rangle = \frac{1}{T}\iint_{-\infty}^{\infty}\left(T_{L1}T_{L2}\,e^{iK(W_{L1}-W_{L2})}\right)\langle e^{(\alpha_1+\alpha_2)+i(\beta_1-\beta_2)}\rangle\,dx\,dy, \qquad \text{II - 4.2}$$

but, because the random functions involved are assumed to be stationary, the mean value under the integral is independent of x and y, and it can be taken outside the integral:[2]

$$\langle\tilde{\Phi}\rangle = \langle e^{(\alpha_1+\alpha_2)+i(\beta_1-\beta_2)}\rangle\left[\frac{1}{T}\iint_{-\infty}^{\infty} T_{L1}T_{L2}\,e^{iK(W_{L1}-W_{L2})}\,dx\,dy\right]. \qquad \text{II - 4.3}$$

The quantity inside the square brackets is seen to be the intrinsic transfer function of the optical system, so the average quantity outside can be considered to be the equivalent transfer function of the turbulent medium:

$$\tilde{\Phi}_R = \langle e^{(\alpha_1+\alpha_2)+i(\beta_1-\beta_2)} \rangle = \langle e^{\gamma+i\delta} \rangle, \qquad \text{II} - 4.4$$

Let us consider this average. First it can be seen that γ and δ are uncorrelated whether α and β are correlated or not.[3] Consider the correlation function

$$\langle \gamma\delta \rangle = \langle (\alpha_1+\alpha_2)(\beta_1-\beta_2) \rangle = \langle \alpha_1\beta_1 + \alpha_2\beta_1 - \alpha_1\beta_2 - \alpha_2\beta_2 \rangle$$
$$= [\langle \alpha_1\beta_1 \rangle - \langle \alpha_2\beta_2 \rangle] + [\langle \alpha_2\beta_1 - \langle \alpha_1\beta_2 \rangle] \qquad \text{II} - 4.5$$

but because of stationarity,

$$\langle \alpha_1\beta_1 \rangle = \langle \alpha_2\beta_2 \rangle$$

and

$$\langle \alpha_2 \beta_1 \rangle = \langle \alpha_1 \beta_2 \rangle ,$$

so the correlation function $\langle \gamma\delta \rangle$ is identically zero. The equivalent transfer function then becomes

$$\tilde{\Phi}_R = \langle e^{\gamma} \rangle \langle e^{i\delta} \rangle . \qquad \text{II} - 4.6$$

That is, it factors into two component equivalent transfer functions, one for the amplitude variations and the other for the phase variations. Let us consider each of these separately.

First we note that γ can be separated into a mean value and a fluctuation with zero mean:

$$\gamma = \langle \gamma \rangle + \gamma' = 2\langle \alpha \rangle + (\alpha_1' + \alpha_2') . \qquad \text{II} - 4.7$$

Then

$$\langle e^{\gamma} \rangle = e^{2\langle \alpha \rangle} \langle e^{\gamma'} \rangle = e^{-2\sigma_\alpha^2} e^{\frac{1}{2}\sigma_\gamma^2} . \qquad \text{II} - 4.8$$

The variance of γ' is related to the variance of α

by

$$\sigma_\gamma^2 = \langle \gamma'^2 \rangle = \langle (\alpha_1' + \alpha_2')^2 \rangle = \langle \alpha_1'^2 + \alpha_2'^2 + 2\alpha_1'\alpha_2' \rangle$$
$$= \langle \alpha_1'^2 \rangle + \langle \alpha_2'^2 \rangle + 2\langle \alpha_1 \alpha_2 \rangle = 2\sigma_\alpha^2 (1+\rho_\alpha)$$

II - 4.9

so

$$\langle e^{\gamma} \rangle = e^{-\sigma_\alpha^2 (1-\rho_\alpha)} \qquad \text{II - 4.10}$$

The phase factor is given by

$$\langle e^{i\delta} \rangle = e^{-\frac{1}{2}\sigma_\delta^2}, \qquad \text{II - 4.11}$$

but

$$\sigma_\delta^2 = \langle \delta^2 \rangle = \langle (\beta_1 - \beta_2)^2 \rangle = \langle \beta_1^2 + \beta_2^2 - 2\beta_1\beta_2 \rangle$$
$$= 2\sigma_\beta^2 (1-\rho_\beta), \qquad \text{II - 4.12}$$

so

$$\langle e^{i\delta} \rangle = e^{-\sigma_\beta^2 (1-\rho_\beta)} = e^{-4\pi^2 \left(\frac{\sigma_w}{\lambda}\right)^2 (1-\rho_w)}, \qquad \text{II - 4.13}$$

The equivalent transfer function for the turbulent medium is therefore

$$\underline{\tilde{\Phi}}_R = \left[e^{-\sigma_\alpha^2(1-\rho_\alpha)} \right] \left[e^{-\sigma_\beta^2(1-\rho_\beta)} \right], \qquad \text{II-4.14}$$

It is interesting to note that both the amplitude and the phase components have the same form [4], shown diagrammatically in figure II-10. However, something might be said about their relative contributions.

As discussed in the section on scintillation, and as can be seen in figure II-1, a value of $\sigma_\alpha = 0.5$ corresponds to fairly strong scintillation. The same value of σ_α reduces the effective transfer function by only 22% in the limit. On the other hand, the standard deviation in the wavefront which would produce the same degradation is less than $\lambda/12$, which is fairly weak. Therefore it is to be expected that phase, or wavefront, fluctuations will dominate over amplitude fluctuations in their combined effect on the blur.

Let us look at the form of the equivalent transfer

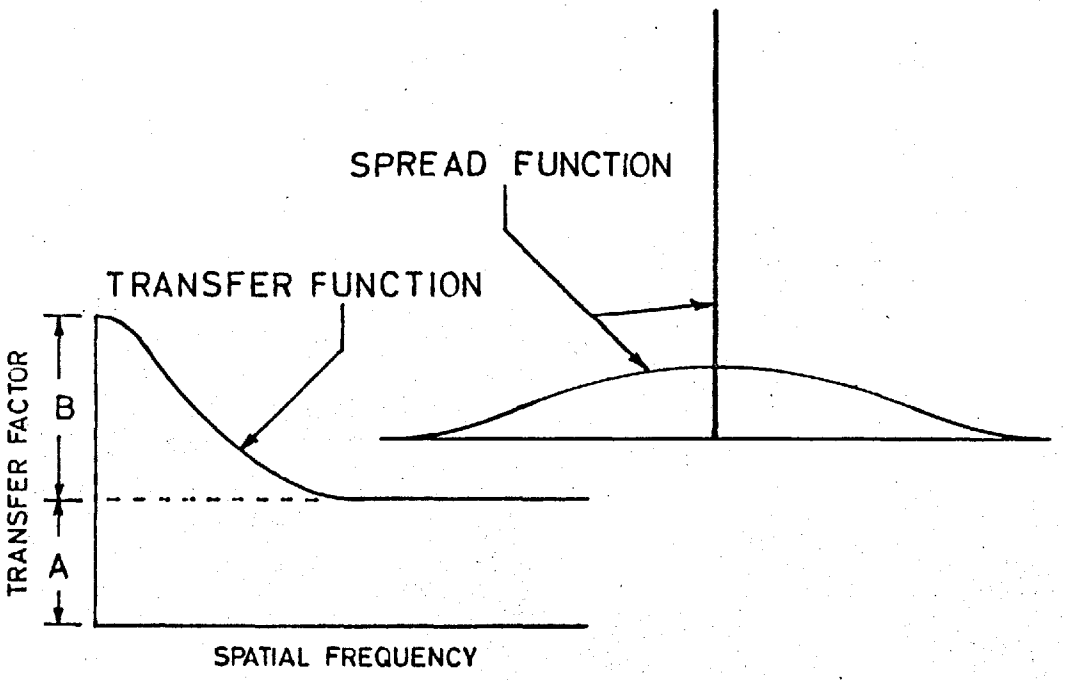


Figure II-10. General character of equivalent transfer function and corresponding spread function for a turbulent medium.

function more closely, taking the phase factor transfer function alone for simplicity.

As the displacement x', y' (to which the spatial frequency is proportional) approaches zero, the auto-correlation function ρ_n approaches 1, and the equivalent transfer function also approaches 1, as it should. As $x', y' \to \infty$, $\rho_n \to 0$ and the equivalent transfer function approaches a plateau of height $e^{-\sigma_\beta^2}$. Thus the equivalent transfer function can be interpreted as a uniform plateau surmounted by a bell-shaped function.

In terms of this picture the equivalent transfer function can be written as

$$\tilde{\Phi}_R = A + B\,Q(x', y') \qquad \text{II - 4.15}$$

where

$$A = e^{-\sigma_\beta^2}$$

$$B = 1 - A = 1 - e^{-\sigma_\beta^2}$$

$$Q = \frac{e^{\sigma_\beta^2 \rho_n} - 1}{e^{\sigma_\beta^2} - 1} \,. \qquad \text{II - 4.16}$$

The latter function is formally similar to the function B_{LN} (eq. II-2.17) in the section on scintillation, and likewise resembles the autocorrelation function ρ_R in that it is unity when $x', y' = 0$, approaches zero as $x', y' \to \infty$ and approaches the shape of ρ_S as $\sigma_\rho \to 0$.

The significance of the above interpretation of the equivalent transfer function is most dramatically shown in the inferred properties of the corresponding spread function. Taking the transfer function to be the sum of two functions, the corresponding spread function is the sum of the inverse transforms of the two component functions. The plateau transforms into a delta function, and the bell-shaped component transforms into another bell-shaped function. This is equivalent to saying that the star appears to the optical system as an undisturbed star of diminished intensity surrounded by a halo, as if a fraction of the light transmitted by the turbulent atmosphere were scattered, the rest remaining undisturbed. Moreover, the fraction of the total light which appears to be scattered is simply the factor B.

The image formed by the optical system likewise will appear to consist of the sum of two components, a core which is the delta function convolved with the spread function of the optical system, and a halo which is the bellshaped halo, also convolved with the optical system spread function. The ratio of the powers in the two components of the image is the same as that of the apparent object.

The core-and-halo effect will, however, only be apparent if the aperture of the telescope is considerably larger than the autocorrelation length of the phase structure. If they are about the same size, then the spread function for the optical system will be about the same size as the halo disc, and the appearance will be that of a slightly degraded spread function.

The relative power distribution between the core and the halo will in general depend on the wavelength of the light. This is most easily seen by considering the fraction A of the power remaining in the core,

$$A = e^{-\sigma_\beta^2} = e^{-4\pi^2\left(\frac{\sigma_w}{\lambda}\right)^2}, \qquad \text{II} - 4.17$$

from which it can be seen that the fraction remaining in the core approaches zero as $\lambda \to 0$, and approaches unity as $\lambda \to \infty$, if we assume that σ_w is constant. In general, of course, σ_w will also vary with the wavelength, but as the latter variation depends on the dispersion of the medium, it will except in very unusual cases by a very weak dependence, and the above variation of A, and by implication B, will be generally found to hold.

When the turbulence is sufficiently severe, the power in the core will be negligibly small, and the halo function Q will have significant values only over the range of x', y' for which the autocorrelation function can be represented by its parabolic approximation:

$$\rho_0 \approx 1 - \tfrac{1}{2} R^2 (x'^2 + y'^2), \qquad \text{II} - 4.18$$

in which case the equivalent transfer function is

given by

$$\tilde{\Phi}_R \approx e^{-\frac{1}{2}\sigma_\beta^2 \rho^2 (x'^2+y'^2)} \qquad \text{II} - 4.19$$

where it is seen that as the turbulence increases in severity, the transfer function becomes more nearly Gaussian in form, and consequently so does the halo. The core, of course, diminishes to the vanishing point.

In the form of equation II-4.19, the variables involved are reduced coordinates. If we restore them to real angular coordinates, we obtain

$$\tilde{\Phi}_R \approx e^{-\frac{1}{2}(4\pi^2)\left(\frac{\sigma_w}{\lambda R}\right)^2 \nu_A^2} \qquad \text{II} - 4.20$$

where ν_A is the angular spatial frequency, expressed in cycles per radian, and is obtained by multiplying the linear spatial frequency in the image plane by the distance from the exit pupil. Note that, when expressed in real angular coordinates the transfer function and spread function for severe turbulence are not only Gaussian, but also independent of the wavelength of the light.[15] In fact, this is the same result that would

be obtained by applying geometrical optical theory, and corresponds to the observation that, in ordinary aberration theory, the diffraction image approaches the geometrical image in appearance as the aberrations become severe.

One other point is interesting to note. The above expression for the transfer function corresponds to the transfer function for the image motion obtained in the section on agitation when C_M is unity; that is, for small apertures. In other words, when the aperture is small, the blur of the apparent (average) star is almost entirely due to the motion of the centroid. As the aperture of the optical system increases relative to the scale of the phase autocorrelation function, the apparent motion of the star decreases according to the function C_M, but the blur (halo function) of the apparent star is independent of the aperture size, and we observe that, as the aperture increases, the motion diminishes but the form of the instantaneous image becomes more distorted in such a way that the average

blur remains constant.

Again, in order to provide a quantitative illustration, an arbitrary form of phase autocorrelation function will be assumed. In fact, we will use the same function employed in the section on agitation, and figure II-11 shows the family of curves obtained for different values of σ_w . Note that the extent of the bell-shaped portion of the curve is essentially constant as long as the plateau is appreciable, after which it decreases in inverse proportion to σ_w .

It would be interesting to see how well this form of transfer function, containing as it does a single parameter arbitrary autocorrelation function, fits experimental data.[24] There are just the minimum possible number of two parameters involved in the transfer function, namely σ_w and h_R , and the results of one investigation[24] are shown in figure II-12. Except for one set of experimental data, the theoretical curves fit very well. The assumed values of σ_w and h_R are shown in the figure.

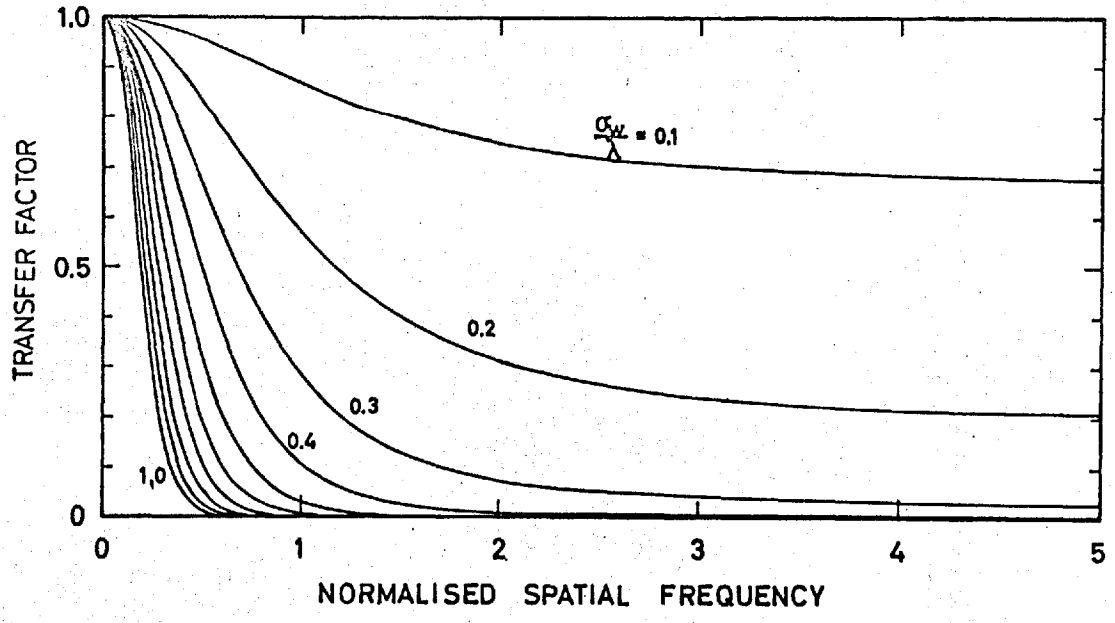


Figure II-11. Variation in equivalent transfer function with depth of wavefront disturbance. Autocorrelation function of wavefront disturbance is assumed constant.

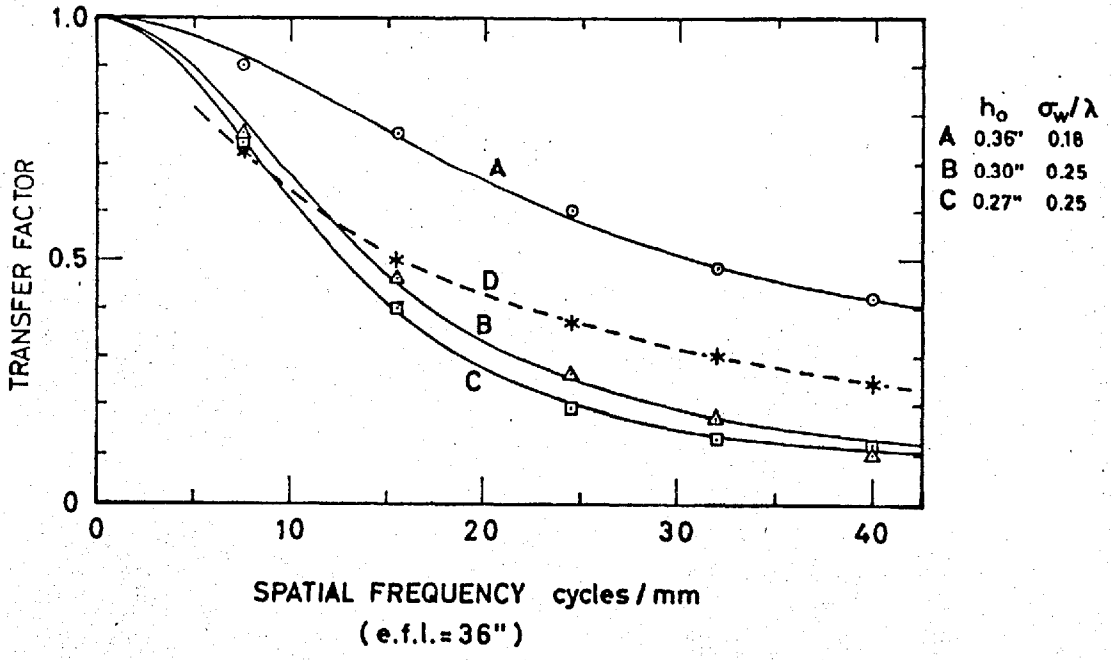


Figure II-12. Best fit of the equivalent transfer function model to experimental data obtained by Djurle and Bäck.[24] The data labelled D could not be fitted. For the others the assumed correlation length and the r.m.s. depth of the wavefront disturbances is indicated.

NOTES

[1] Both Chernov[13] and Tatarski[14] incorrectly assumed a mean value of zero for α. This however does not have any serious effect on the validity of their conclusions as to the statistical properties of the transmitted wavefront.

[2] Essentially the same result was obtained by Diederichs[23] in investigating the effect of imperfect polish on optical surfaces.

[3] The fact that γ and δ are uncorrelated was shown by Chernov.[13]

[4] Both Keller and Hardie[15] and Hufnagel and Stanley[22] obtained an expression for the equivalent transfer function having the same form as either of the factors in equation II-4.16, but the random variable involved was the refractive index fluctuation of the medium rather than the wavefront disturbance. Although Keller and Hardie obtained this function,

they did not recognize it as a transfer function but considered it to be merely a mathematical convenience. Hufnagel and Stanley, on the other hand, called it a transfer function, and obtained it in a most elegant manner. They did not, however, distinguish between amplitude and phase contributions.

SUMMARY

The intrinsic factors which determine the over-all performance of an image-forming system are the performance of the optical system and the performance of the detector. Other factors may or may not affect the over-all performance, and are therefore in this thesis called extrinsic. We have investigated in this thesis several of these extrinsic factors.

In part I we examined two important time-dependent factors affecting the performance of a photographic system, namely the effects of the shutter in obscuring part of the pupil of the optical system during the finite opening and closing time it requires, and the effects of motion of the image relative to the film during the exposure time. In chapter I-1 a general expression for the transfer function of the exposure image was derived, including as factors the film transfer function, the shutter function, the optical transfer function, and the motion of the image.

In chapter I-2 the effect of the shutter on the exposure-image transfer function in the absence of image motion was discussed, and it was shown that a focal-plane shutter introduces an effective shutter transfer function which is independent of the pupil size and of aberrations. It was also shown that a between-the-lens shutter acts as an incoherent apodizing agent which is capable of improving the performance in the presence of aberrations.

In chapter I-3 the effect of image motion was investigated, assuming that the shutter does not significantly modify the optical transfer function. A two-dimensional treatment was pursued throughout. The concept of an equivalent spread function for image motion was developed. Uniform linear motion and simple harmonic motion were discussed in detail, and elementary forms of combined motions were considered. An approximation appropriate for small degradations was derived, and expressions for all possible combinations of uniform linear motion and simple harmonic motion were obtained.

In part II we investigated the effects of a turbulent medium between the object and the optical system on the properties of point images. This part covered the classical phenomena of scintillation, agitation (image motion), and blur which have been observed for many years by astronomers.

In chapter II-1 the concept of a pupil function with a random component was developed, the random component being expressed as an exponential with a complex argument in which the amplitude fluctuations depend on the real part of the argument and the phase fluctuations on the imaginary part. It was shown that this exponential model with both the real and the imaginary parts assumed to have Gaussian probability densities is a reasonable one considering its simplicity.

In chapter II-2 scintillation was investigated and it was shown that the coefficient of variation of the total flux in a star image (ratio of the standard deviation to the mean value), which is a measure of scintillation, can be obtained as the square root of the

normalized integral in frequency space of the product of two functions, one of which depends only on the random amplitude fluctuations and the other, which was called the core function, only on the pertinent properties of the optical system. For illustrative purposes a function was assumed for the amplitude autocorrelation function and the dependence of scintillation on the size of the aperture was determined for several forms of aperture, including one which was apodized. The dependence for a circular aperture was then fitted to experimental data reported in the literature and it was seen that at least qualitatively the form of the dependence is quite reasonable.

In chapter II-3 agitation was investigated and it was shown that the radius of gyration of the centroid of a star image (r.m.s. deviation of the centroid from its mean position) can be obtained as ths square root of the sum of two normalized integrals, one of which depends only on the amplitude fluctuations, whereas the other depends on both amplitude and phase fluctuations.

The first depends also on the aberrations of the optical system and is zero if the aberrations are zero. The second is independent of the aberrations of the optical system. Both integrands are products of two factors, one of which depends on the random components alone and the other on the optical system alone. For illustrative purposes it was assumed that only phase fluctuations were present, and a function was assumed describing the phase autocorrelation function. The dependence of agitation on aperture size was then obtained for the same apertures used in the scintillation illustration, and it was seen that agitation depends on aperture size in much the same way that scintillation does.

Incidental developments obtained in chapter II-3 were (1) that the position of the centroid of any point spread function calculated by diffraction is the same as the position of the centroid calculated by geometrical optics, and (2) the position of the centroid of the spread function depends only on the gradient of the transfer

function at the origin, which, as shown in Appendix IV, is the gradient of the phase component of the transfer function at the origin.

In chapter II-4 blur was investigated, and it was shown that the effect of the turbulent medium on the average image is just as if the turbulent medium were an independent blurring factor with its own transfer function. Furthermore it was shown that the contributions from amplitude and phase fluctuations are independent of each other, the effective transfer function for the turbulent medium factoring into two parts, one dependent on amplitude fluctuations and the other on phase fluctuations, the form of the two being identical. This form is such that it may be separated into the sum of two parts, one a plateau of uniform height and the other a bell-shaped function determined by the autocorrelation function of the disturbance. The corresponding apparent spread function consists of a delta function added to a bell-shaped spread function; that is, the apparent star has an undisturbed core surrounded

by a halo. Again for illustrative purposes it was assumed that only phase fluctuations were present and that they had the same autocorrelation function as was used in the agitation illustration, and the resulting function was fitted to experimental data reported in the literature. Although one set of data could not be fitted, the others fitted very well, and again at least qualitatively the form of the dependence is quite reasonable.

APPENDIX I. GRAPHICAL-NUMERICAL INTEGRATION

In cases involving image motion where the optical transfer function cannot be assumed independent of the shutter, or where the path and velocity of motion cannot be reduced to simple functions, a direct integration based on equation I-1.20 is required. As in the chapter on image motion, it is convenient to rotate the reference coordinates. Thus,

$$\tilde{\Phi}_E(\tilde{u}', \psi) = \tilde{\Phi}_F\, t_e^{-1} \int_{-\infty}^{\infty} S\, \tilde{\Phi}_L \exp(-2\pi i \tilde{u}' u')\, dt. \qquad A\text{-}1.1$$

The evaluation of this integral depends on knowledge of the three functions S, $\tilde{\Phi}_L$, and u'. The first two will generally be determined empirically and they enter into the integrand directly. The third, however, will generally have to be calculated from a pair of parametric functions $u(t)$ and $v(t)$ by means of equation I-3.3.for each orientation and in addition must be transformed into the complex exponential function contained in the integrand.

The real and imaginary parts of this function can be readily calculated graphically from the basic functions $u(t)$ and $v(t)$ with the aid of a simple pair of transparent plastic overlays shown in fig. A-1. One overlay is for the real (cosine) component and the other for the imaginary (sine) component. Each overlay has marked on it a set of ten sawtooth waves, each representing one frequency component $\tilde{u}'$. Theoretically the waves should be sine and cosine waves, but the curved lines which that would require are difficult to scribe accurately, whereas straight lines are extremely simple. The "distortion" of the sine and cosine waves into sawtooth waves is exactly compensated by calibrating them with a non-linear scale.

In order to use these overlays the following procedure is required. A set of points is marked on a piece of paper representing the image plane, each point representing the position of the image at each of a set of equally-spaced intervals of time during the exposure. The points are labelled serially and the center of the

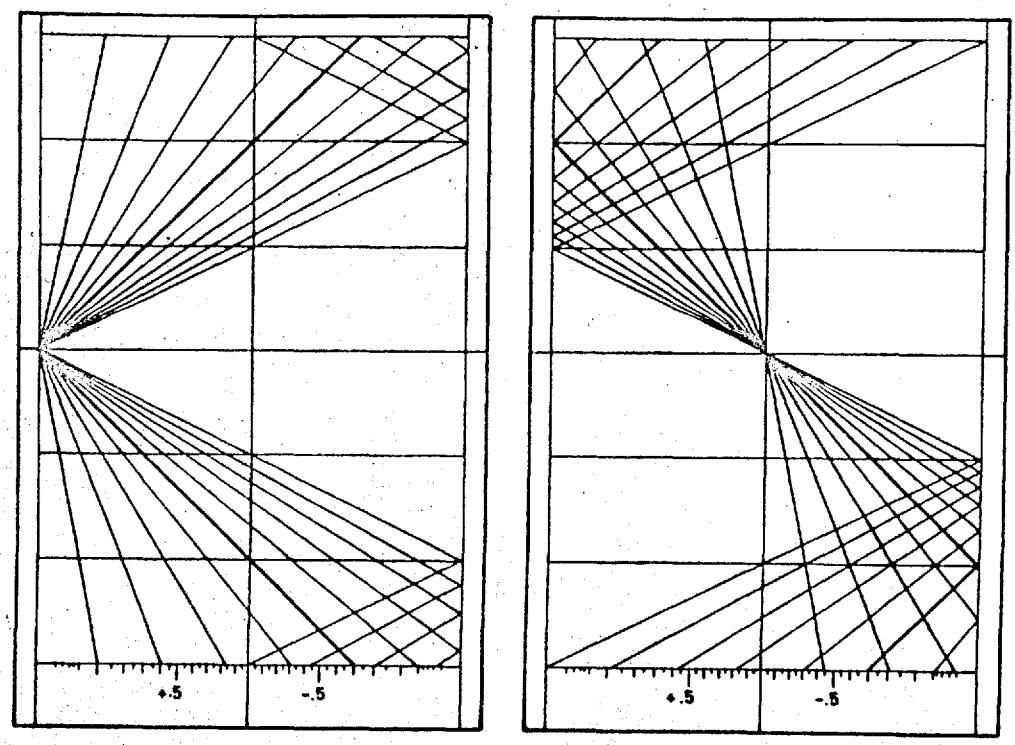


Figure A-1. Overlay pattern to aid in the calculation of equation A-1.1.

pattern is estimated and given a distinctive mark.

The piece of paper is fastened down on a drawing board with the chosen direction of scan ψ vertical. A T-square is then located so that, with either of the overlays riding against it, the horizontal center line of the overlay passes through the center mark of the pattern. The T-square is fastened to the board in this position.

Now if one of the overlays is slid horizontally along the T-square until a chosen point on the paper lies on a chosen "wave", the distance of that point from the vertical center line, measured on the non-linear scale, is equal to the value of the real or imaginary part of the complex function in the integrand of equation A-1.1.

If the non-linear scale had extended up through the pattern of waves, this value could have been read directly. However, such an extension of the scale would have cluttered the overlay to such an extent that it would have been difficult to set. Instead, the reading is obtained

by bringing a straight-edge (such as the overlay not in use at the moment, or a drawing triangle) up against the side of the overlay and, holding the straight-edge stationary, sliding the overlay vertically until the point on the paper is on the scale.

In this manner the complex function in the integrand can be evaluated for ten spatial frequencies at uniformly spaced intervals of time. Each value is multiplied by the corresponding values of S and $\tilde{\Phi}_L$ and the product is integrated. For other orientations ψ, the paper is simply rotated and the process repeated.

APPENDIX II. PROBABILITY DENSITY DISTRIBUTION OF THE SUM OF TWO NORMALLY DISTRIBUTED RANDOM VARIABLES WHICH ARE CORRELATED

It is usually stated in the literature that the sum of two normally distributed random variables is also a normally distributed random variable if the two component variables are uncorrelated. We shall show that this restriction is unnecessary, and that the sum is normally distributed even when the two componets are correlated.

Let the two componet variables be represented by $x(t)$ and $y(t)$ and the sum by $z(t) = x(t) + y(t)$. Each of the three variables can be assumed to have zero mean value without loss of generality (if each were separated into the sum of its mean value and its fluctuation with zero mean, the mean value of the sum would be equal to the sum of the mean values of the components, and the fluctuations with zero mean would be related as above). The variances and covariances would be

$$\langle x^2 \rangle = \sigma_x^2$$

$$\langle y^2 \rangle = \sigma_y^2$$

$$\langle xy \rangle = \sigma_x \sigma_y \rho$$

$$\langle z^2 \rangle = \langle (x+y)^2 \rangle = \langle (x^2 + y^2 + 2xy) \rangle = \sigma_x^2 + \sigma_y^2 + 2\sigma_x \sigma_y \rho = \sigma_z^2 \qquad \text{A-2.1}$$

where ρ is the cross-correlation of x and y.

The joint probability of x and y is

$$p(x, y) = \frac{1}{2\pi \sigma_x \sigma_y \sqrt{1-\rho^2}} \exp\left[- \frac{\left(\frac{x}{\sigma_x}\right)^2 + \left(\frac{y}{\sigma_y}\right)^2 - 2\rho \left(\frac{x}{\sigma_x}\right)\left(\frac{y}{\sigma_y}\right)}{2(1-\rho^2)} \right], \qquad \text{A-2.2}$$

The probability density of z is

$$p(z) = \int_{-\infty}^{\infty} p(x, y = z - x)\, dx$$

$$= \frac{1}{2\pi \sigma_x \sigma_y \sqrt{1-\rho^2}} \int_{-\infty}^{\infty} \exp\left[- \frac{\left(\frac{x}{\sigma_x}\right)^2 + \frac{(z-x)^2}{\sigma_y^2} - 2\rho \frac{x(z-x)}{\sigma_x \sigma_y}}{2(1-\rho^2)} \right] dx,$$

$$\text{A-2.3}$$

Consider the numerator in the exponent:

$$\frac{x^2}{\sigma_x^2} + \frac{(z-x)^2}{\sigma_y^2} - \frac{2\rho x(z-x)}{\sigma_x \sigma_y}$$

$$= \left[\frac{1}{\sigma_x^2} + \frac{1}{\sigma_y^2} + \frac{2\rho}{\sigma_x \sigma_y}\right] x^2 - 2\left[\frac{1}{\sigma_y^2} + \frac{\rho}{\sigma_x \sigma_y}\right] z x + \frac{z^2}{\sigma_y^2}$$

$$= \left[\frac{\sigma_z^2}{\sigma_x^2 \sigma_y^2}\right] x^2 - 2\left[\frac{\sigma_x^2 + \sigma_x \sigma_y \rho}{\sigma_x^2 \sigma_y^2}\right] z x + \frac{z^2}{\sigma_y^2}$$

$$= \frac{\sigma_z^2}{\sigma_x^2 \sigma_y^2}\left[x - \frac{\sigma_x^2 + \sigma_x \sigma_y \rho}{\sigma_z^2}\right]^2 + \frac{(1-\rho^2)}{\sigma_z^2} z^2$$

$$= a(x-h)^2 + k \qquad A-2.4$$

Then

$$p(z) = \frac{1}{2\pi \sigma_x \sigma_y \sqrt{1-\rho^2}} \exp\left[-\frac{k}{2(1-\rho^2)}\right] \int_{-\infty}^{\infty} \exp\left[-\frac{a(x-h)^2}{2(1-\rho^2)}\right] dx$$

$$= \frac{1}{2\pi \sigma_x \sigma_y \sqrt{1-\rho^2}} \left[\sqrt{\frac{2\pi(1-\rho^2)}{a}}\right] \exp\left[-\frac{k}{2(1-\rho^2)}\right]$$

$$= \frac{1}{\sqrt{2\pi}\,\sigma_z} e^{-\frac{1}{2}\left(\frac{z}{\sigma_z}\right)^2}, \qquad A-2.5$$

which is a normal distribution.

APPENDIX III. AVERAGES OF EXPONENTIAL FUNCTIONS OF NORMALLY DISTRIBUTED RANDOM VARIABLES.

Throughout the second part of the thesis average values of exponential functions of normally distributed random variables frequently occur. If $\zeta = \langle\zeta\rangle + \zeta'$ is such a variable and does not have a zero mean, then the mean should be subtracted first:

$$\langle e^{\zeta} \rangle = e^{\langle\zeta\rangle} \langle e^{\zeta'} \rangle , \qquad A-3.1$$

Thus without loss of generality we can assume that the random variable under consideration has zero mean. Then

$$\langle e^{\zeta} \rangle = \int_{-\infty}^{\infty} e^{\zeta} p(\zeta) d\zeta , \qquad A-3.2$$

where

$$p(\zeta) = \frac{1}{\sqrt{2\pi}\,\sigma_\zeta} e^{-\frac{1}{2}\left(\frac{\zeta}{\sigma_\zeta}\right)^2} , \qquad A-3.3$$

Thus

$$\langle e^{z} \rangle = \frac{1}{\sqrt{2\pi}\,\sigma_z} \int_{-\infty}^{\infty} e^{-\frac{1}{2\sigma_z^2}\left[z^2 - 2\sigma_z^2 z\right]} dz$$

$$= \frac{1}{\sqrt{2\pi}\,\sigma_z} \int_{-\infty}^{\infty} e^{-\frac{1}{2\sigma_z^2}\left[(z-\sigma_z)^2 - \sigma_z^4\right]} dz$$

$$= \frac{e^{\frac{1}{2}\sigma_z^2}}{\sqrt{2\pi}\,\sigma_z} \int_{-\infty}^{\infty} e^{-\frac{1}{2\sigma_z^2}(z-\sigma_z^2)^2} dz$$

$$= \frac{1}{\sqrt{2\pi}\,\sigma_z} \left[\sqrt{2\pi\sigma_z^2}\right] e^{\frac{1}{2}\sigma_z^2}$$

$$= e^{\frac{1}{2}\sigma_z^2}. \qquad \text{A - 3.4}$$

If the random function z is a linear combination of normally distributed random functions,

$$z = mx + ny \qquad \text{A - 3.5}$$

where m and n are arbitrary constants, then

$$\sigma_z^2 = \langle z^2 \rangle = \langle m^2x^2 + n^2y^2 + 2mn\,xy \rangle$$

$$= m^2\sigma_x^2 + n^2\sigma_y^2 + 2mn\,\sigma_x\sigma_y\rho. \qquad \text{A - 3.6}$$

Thus

$$\langle e^{2\alpha'} \rangle = e^{\frac{1}{2}(4\sigma_\alpha^2)} = e^{2\sigma_\alpha^2}$$

$$\langle e^{2(\alpha_1'+\alpha_2')} \rangle = e^{\frac{1}{2}(4\sigma_\alpha^2+4\sigma_\alpha^2+8\sigma_\alpha^2\rho_\alpha)} = e^{4\sigma_\alpha^2(1+\rho_\alpha)}$$

$$\langle e^{i\delta} \rangle = e^{\frac{1}{2}(-\sigma_\delta^2)} = e^{-\frac{1}{2}\sigma_\delta^2}$$

$$\langle e^{i(\beta_1-\beta_2)} \rangle = e^{\frac{1}{2}(-\sigma_\beta^2-\sigma_\beta^2+2\sigma_\beta^2\rho_\beta)} = e^{-\sigma_\beta^2(1-\rho_\beta)}, \qquad A-3,7$$

APPENDIX IV. DERIVATIVE OF THE TRANSFER FUNCTION AT THE ORIGIN

The transfer function is in general a complex function:

$$\tilde{\Phi} = |\tilde{\Phi}(x',y')|\, e^{i\phi(x',y')}, \qquad A-4.1$$

Its partial derivative with respect to x' is

$$\frac{\partial \tilde{\Phi}}{\partial x'} = e^{i\phi(x',y')}\left[\frac{\partial |\tilde{\Phi}|}{\partial x'} + i|\tilde{\Phi}|\frac{\partial \phi}{\partial x'}\right], \qquad A-4.2$$

Evaluated at the origin,

$$\left.\frac{\partial \tilde{\Phi}}{\partial x'}\right|_{x',y'=0} = e^{i\phi(0,0)}\left[\left.\frac{\partial |\tilde{\Phi}|}{\partial x'}\right|_{x',y'=0} + i|\tilde{\Phi}(0,0)|\left.\frac{\partial \phi}{\partial x'}\right|_{x',y'=0}\right], \qquad A-4.3$$

provided the component partial derivatives exist. Now

$$\phi(0,0) \equiv 0\,, \quad |\tilde{\Phi}(0,0)| \equiv 1\,, \qquad A-4.4$$

so

$$\left.\frac{\partial \tilde{\Phi}}{\partial x'}\right|_{x',y'=0} = \left.\frac{\partial |\tilde{\Phi}|}{\partial x'}\right|_{x',y'=0} + i \left.\frac{\partial \phi}{\partial x'}\right|_{x',y'=0} \qquad A-4.5$$

The derivative of the phase function ϕ is continuous at the origin, so it presents no problem. However, the derivative of the modulus is not continuous for an optical system having a sharply defined pupil boundary. It is continuous for $y' \neq o$, and has the value of zero at $x' = o$ for all $y' \neq o$ but when $y' = o$, it has a finite non-zero negative value in the limit when approached from the positive side of x', and an equal but positive value in the limit when approached from the negative side of x'. The nature of the discontinuity is similar to that frequently obtained in Fourier series analysis, and it seems reasonable to assume that the value at the origin can be defined as the average between the right and left limiting values. This value is identically zero, so the partial derivative with respect to x' becomes

$$\left.\frac{\partial \tilde{\Phi}}{\partial x'}\right|_{x',y'=o} = i \left.\frac{\partial \phi}{\partial x'}\right|_{x',y'=o} . \qquad A-4.6$$

The same argument holds for the partial derivative with respect to y', so the important conclusion can be made (see the section on agitation) that the position of the centroid of the image is determined by the gradient of the phase of the transfer function at the origin, irrespective of the shape of this phase function.

BIBLIOGRAPHY

(1) O.H.Schade, RCA Rev. 9, 5,245,490,653 (1948)

(2) P.B.Fellgett and E.H.Linfoot, Trans. Roy. Soc. A247, 369 (1955)

(3) E.L.O'Neill, Introduction to Statistical Optics, Addison-Wesley (1963)

(4) C.E.Shannon and W.Weaver, The Mathematical Theory of Communication, U. of Illinois Press (1963) (reprint)

(5) H.Frieser, Phot. Sci. Eng. 4, 324 (1960)

(6) M.E.Bechtel, Cornell Aeronautical Laboratory, Rept. No. VF-1260-P-2, Contract No. AF33(616)-8870 (1959)

(7) T.Asakura, J.Appl. Phys. Japan 30, 797 (1961)

(8) R.M.Scott, Phot. Sci. Eng. 3, 201 (1959)

(9) M.D.Rosenau, Photogrammetric Eng. 421 (June 1961)

(10) T.T.Chang, Cornell Aeronautical Laboratory, Rept. No. VF-1260-P-2, Contract No. AF33(616)-8870 (1959)

(11) D.P.Paris, Phot. Sci. Eng. 7, 233 (1963)

(12) L. O. Hendeberg and W.E.Welander, Appl. Opt. 2, 379 (1963)

(13) L.A.Chernov, Wave Propagation in a Random Medium, McGraw-Hill (1960)

(14) V.I.Tatarski, Wave Propagation in a Turbulent Medium, McGraw-Hill (1961)

(15) G.Keller and R.H.Hardie, Astronom. J. 59, 105 (1954)

(16) J. van Isacker, Quart. J. Roy. Meteor, Soc. 80, 251 (1954)

(17) D.H.Menzel, Proc. Symp. Solar Seeing 25, Rome (1962)

(18) J.R.Meyer-Arendt and C.B.Emmanuel, NBS Tech. Note 225 (1965)

(19) M.A.Ellison and H.Seddon, Monthly Not. Roy. Astron. Soc. 112, 73 (1952)

(20) N.M.Protheroe, Contrib. Perkins Observ., ser. 2, no. 4, 127 (1955)

(21) C.C.Lin, Statistical Theories of Turbulence, Princeton Univ. Press (1961)

(22) R.E.Hufnagel and N.R.Stanley, J. Opt. Soc. Am. 54, 52 (1964)

(23) E. Diederichs, Dissertation, 1961

(24) E. Djurle and A. Bäck, J. Opt. Soc. Am. 51, 1029 (1961)

ACKNOWLEDGMENTS

The author would like to acknowledge the advice and encouragement given by Dr. H. H. Hopkins, especially for suggesting the derivation of the equivalent transfer function for a focal-plane shutter, and also the help of his wife, Pamela, in typing the manuscript.

Journal of Research of the National Bureau of Standards Vol. 56, No. 5, May 1956 Research Paper 2672

Characteristics of an Image-Forming System

Roland V. Shack

Two general approaches to the analysis of an image-forming system are considered. One depends on the image of a point object and the other on the Fourier transform of this image. The two are developed independently and then coordinated, a practical characteristic function being determined for each approach. The relative merits of the two approaches are considered.

1. Introduction

For the past few years considerable energy has been expended in the search for an objective procedure for evaluating the quality of images formed by optical instruments. Existing procedures have been found to be not entirely satisfactory, and much work has been done in measuring previously unused physical parameters, which are objectively determinable, for the purpose of correlating them with the existing quality criteria.

The objection to this is that such an empirical, and therefore much limited, correlation eliminates only one of the faults of the existing criteria, and this is the subjectivity of their determination. Any other weakness is ignored.

A better approach is to analyze the image-forming process as a phenomenon, the aim being to characterize the process in as general and inclusive a way as possible, consistent with practical instrumentation. New criteria of image quality would of course be expected to be developed. Many approaches have been made in this direction also, and the present paper is to be included among them. However, in contrast to some of the published material, the emphasis here is on the practicality and usefulness of the results obtained rather than on mathematical rigor, although the treatment should be rigorous enough to include all essential factors.

Let us consider this matter of practical instrumentation. The heart of the test instrument is the photosensitive detector, for it is this which provides the data by which the tested instrument is evaluated. Three practical photosensitive detectors are available—the eye, the photographic emulsion, and the photocell.

The only test of image quality for which the eye is capable of making quantitative measurements is the resolving-power test. This test is rapid and relatively inexpensive, but the information obtained is incomplete, the precision is low, and the results are subject to variation from individual to individual.

A photographic detector allows quantitative measurements to be made under nonthreshold conditions, but time is required for processing, the processing conditions must be rigidly standardized, the granularity and diffusion in the emulsion affect the results, the response of the film is nonlinear with respect to incident flux, and, in the end, an additional sensing mechanism, such as a microdensitometer, is needed to reduce the emulsion properties to numerical values.

The photocell is probably the most satisfactory photosensitive detector for the test instrument. Within its proper operating range, its characteristics remain reasonably constant, its response is linear with respect to incident flux, its spectral response can be adjusted so as to approximate that of the eye, and its output can very easily produce graphical or numerical results. However, it must be used in conjunction with an aperture that limits the spatial integration of the detail in the image being examined, and there must be provision for relative displacement between the aperture and the image so that various portions of the image may be sampled.

It should also be pointed out that the report is illustrated throughout by the characteristics of an aberration-free system with a circular aperture in monochromatic light, diffraction being the sole source of image degradation, and the light from various points in the object space being noncoherent. This has been selected as an interesting and informative type of system, which real systems tend to approximate as their quality improves. It must be emphasized,

379949—56——1

however, that this is used for illustration only. The material covered applies to all types of images, assuming noncoherence.

There are two viewpoints from which image evaluation can be approached. One, the classical viewpoint, considers the point image to be the most fundamental element in an image process. Any object can be regarded as a summation of points, and its image as the summation of the corresponding point images. An evaluation of the point image would be sufficient to characterize the system.

The other viewpoint involves the concepts of Fourier analysis. Here the object is considered to be the summation of a set of sinusoidal waves distributed in the object plane. These component waves, differing from each other in amplitude, frequency, phase, and direction, are spatially distributed waves. That is, they are spatial, not temporal, sinusoidal variations in brightness throughout the object plane. The image consists of the summation of the images of these component waves. A description of the manner in which the optical system forms images of these component waves would also be sufficient to characterize the system.

In section 2 a way of describing the point image is developed, which can be obtained, at least in principle, from a variety of test objects, namely, a point, a fine line, a variable slit, and a knife edge. Data from any or all of these objects can be represented by a single common curve, which can be interpreted in terms of any of the objects.

Section 3 deals with the Fourier type of approach. An imaging system does not affect the frequency, direction, or sinusoidal character of the component waves. It can only affect their amplitudes and phase relationships. The function that describes these modifications as a function of the frequency and direction of the component waves is also characteristic of the imaging system. The test object required to obtain this information consists of a series of patterns in which the luminance varies sinusoidally, each pattern having a different spatial frequency and all oriented in the same direction. The directional variation can be obtained by rotating the test object with respect to the system being tested.

Section 4 is concerned with the coordination between the two viewpoints. It is shown that it is possible to obtain from either approach the characteristics for a periodic test object consisting of alternate dark and light stripes of equal width, such as is commonly used in resolving-power work. It is also shown that it is possible to transform the characteristics of either approach into those of the other.

Section 5 discusses the application to practice, the interpretation of the results and the relative merits of the two approaches.

The appendix gives the mathematical formulation of the diffraction images used as illustrations.

In the following mathematical statements, constant coefficients are ignored in the integrations unless otherwise indicated. The functions are assumed to be normalized after integration. Also, the object-plane coordinates are reduced to the image plane by application of the magnification.

2. Evaluation of the Point Image

2.1. General Image Formation with the Point Image

The point image is the flux-density distribution in the image plane when the object is a point source (figs. 1, 2). A general object can be considered to consist of a summation of points, and its image the summation of the corresponding point images.

Let $O(x,y)$=general object function,
$\varphi(x,y)$=point image function,
$I(x,y)$=general image function.

Then

$$I(x',y')=\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} O(x,y)\varphi(x'-x,y'-y)dxdy. \tag{1}$$

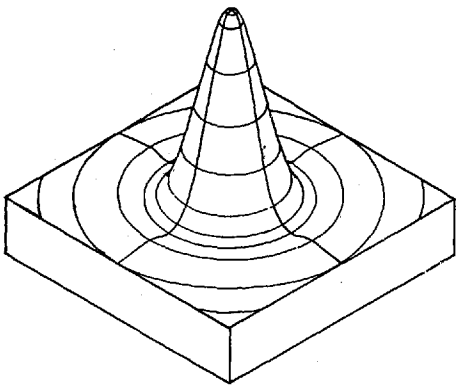


FIGURE 1. *Point-image model.*

In this model the vertical dimension represents flux density (per unit area). It is understood that the rings actually continue indefinitely, whereas only the first bright ring is shown completely.

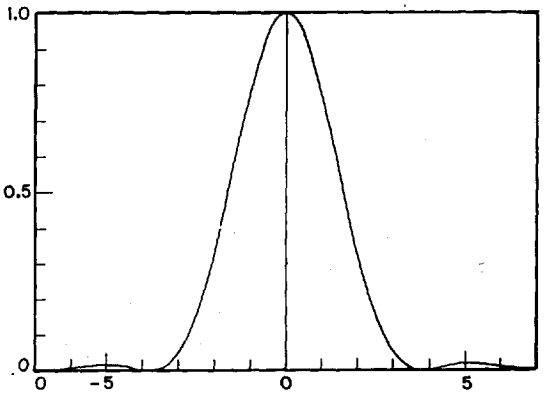


FIGURE 2. *Section of point image.*

The normalized distance in the image plane is measured in z-units, as explained in appendix 1.

The primed variables represent the displacement between the O and φ functions required for the integration. For each displacement, the integral, being a definite integral, establishes a specific value for I. The image function, I, then, is a function of the displacement involved in the integration. The space described by x',y', however, is the same as that described by x,y. O and I can be compared point for point.

The integration can also be written in such a form that the object function is the displaced function, that is,

$$I(x',y')=\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\varphi(x,y)O(x'-x,y'-y)dxdy. \tag{2}$$

Either of these forms is perfectly valid, and either may be transformed into the other, provided one recognizes that x,y in eq (1) is not the same as x,y in eq (2). To distinguish them, one might use subscripts on the symbols, but this would make the equations more confusing, and is not necessary if one understands the situation.

The image function is to be sampled with a space-integrating detector in the image plane, that is, a photocell behind an aperture.

Let $A(x,y)$=detector aperture transmission function,

$E(x,y)$=total flux passing through A as a function of the position of A.

Then

$$E(x'',y'')=\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}I(x',y')A(x''-x',y''-y')dx'dy'. \tag{3}$$

Combining this with eq (2) we obtain

$$E(x'',y'')=\int\int\int\int_{-\infty}^{\infty}\varphi(x,y)O(x'-x,y'-y)A(x''-x',y''-y')dxdydx'dy'. \tag{4}$$

From this it can be seen that the functional characteristics of O and A can be interchanged without affecting the measured flux, E. For example, suppose the object were a point source and the aperture a circular hole centered on the point image. The output from the photocell will indicate amount of flux passing through the hole. Then interchange the object and aperture. Now the object is a uniformly luminous disk with a reduced diameter equal to that of the previous aperture, and the new aperture is a minute hole, equal in size to the reduced geometric size of the previous point source. The output from the photocell will be the same as before.

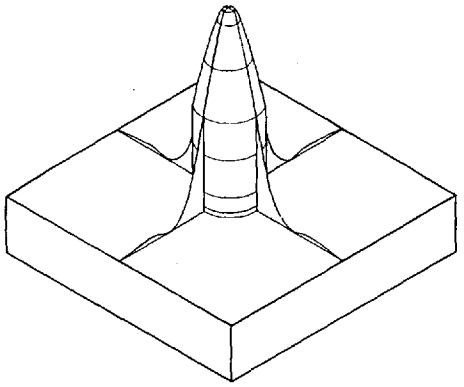


FIGURE 3. *Point-image evaluation—Hopkins' method.*

The solid represents the portion of the total flux that passes through the circular aperture.

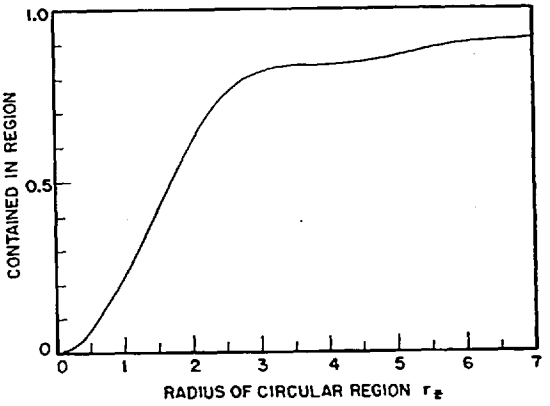


FIGURE 4. *Radial flux distribution of point image.*

This curve shows the normalized volume of the solid of figure 3 as a function of the radius of the limiting cylinder.

Mathematically, and in a more general sense, this situation is as follows:
If O is a point source, then

$$E(x'',y'')=\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}A(x''-x',y''-y')\left[\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\varphi(x,y)O(x'-x,y'-y)dxdy\right]dx'dy'$$
$$=\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\varphi(x',y')A(x''-x',y''-y')dx'dy', \tag{5}$$

and the flux detected depends on the nature of A.

On the other hand, if A is a point aperture, then

$$E(x'',y'')=\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\varphi(x,y)\left[\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}O(x'-x,y'-y)A(x''-x',y''-y')dx'dy'\right]dxdy$$
$$=\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\varphi(x,y)O(x''-x,y''-y)dxdy, \tag{6}$$

which is identical in form with eq (5), except that A is replaced by O. If O in eq (6) had the same functional characteristics as A in eq (5), then they would be mathematically indistinguishable, and the same E will be obtained from either. Also note that eq (6) has the same form as eq (2). As one would expect, the use of a point aperture would produce an undistorted map of the flux-density distribution in the image plane.

2.2. Determination of a Characteristic Function of the Point Image

It should be clear from the above that the function which distinguishes one image-forming system from another is the point-image function φ. The problem is to find some way of describing φ, which can be obtained experimentally and which provides significant information to the user.

The direct mapping of the flux-density distribution in the image of a point object with a point aperture is impractical because of the very low-energy levels involved. A practical method must involve in some manner the integration of the flux over an area.

One method, used by Hopkins [7] [1] is to measure the flux contained within successive concentric circular regions about the center of the point image (fig. 3). The ratio of the flux contained within a circular region to the total flux of the point image is plotted as a function of the radius of the circle (fig. 4).

[1] Figures in brackets indicate the literature references at the end of this paper.

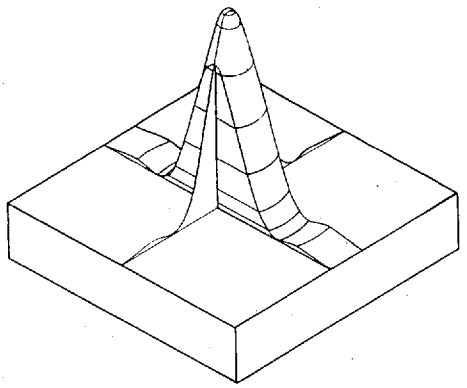


FIGURE 5. *Determination of point-image characteristic.*

The solid represents the portion of the total flux that passes through the receiver slit, the point image being centered on the slit.

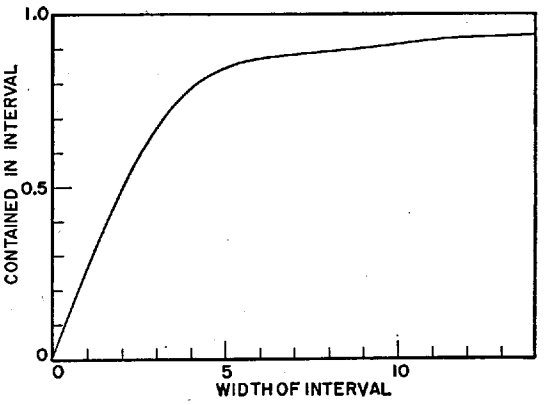


FIGURE 6. *Point-image characteristic.*

This curve shows the normalized volume of the solid in figure 5 as a function of the width of the slab.

This method has several advantages over mapping the flux-density distribution directly. Of course, there is an increase in the energy involved, which increases the signal-to-noise ratio of the measurements. Also, there is a reduction of a three-dimensional function to two dimensions, which is more convenient. Further, this method provides the user with an idea of the contrast with which small detail will be imaged, for, by the interchangeability of the object and aperture function, the resulting curve (fig. 4) can be considered to represent the flux density at the center of the image of a disk as a function of the size of the disk. This disk corresponds to a small object in the scene, later observed by the user.

This method does have disadvantages, however. It is satisfactory for systems in which the point image is radially symmetrical, but is insensitive to the presence of radial asymmetry, such as exists in an astigmatic or comatic image. It also presents the practical difficulty of locating the centers of the apertures or disks in two dimensions with precision.

A different integrative method that allows a considerable increase in the available energy is one in which the integration is limited in one direction only. This is done by measuring the flux contained in successive widths about the center of the point image (fig. 5). The normalized flux contained in a region as a function of the width of the region (fig. 6) is the function that is here proposed as the most useful and practical characteristic function describing the point image. It will hereafter be called the point-image characteristic $K(w)$, where w represents the width of the region. For example, if the integration is limited in the x-direction, then

$$K(w_x)=\int_{-w_x/2}^{w_x/2}\int_{-\infty}^{\infty}\varphi(z,y)\,dy\,dx. \tag{7}$$

It should be noticed that this function does not actually involve a reduction of three dimensions to two, because the curve obtained is a function of the direction in which the integration is taken. This makes the data somewhat less convenient than is true of the previous method, but this is not objectionable because of the additional information obtained. The new method will detect a lack of radial symmetry. However, the inconvenience involved is not too great because most images are bilaterally symmetrical, or nearly so, and two mutually perpendicular orientations are all that are necessary to characterize the image.

The point-image characteristic has other virtues beyond the relatively large amount of available energy and the ability to detect lack of radial symmetry. These arise out of the unidirectional limitation of the integration.

For example, the point-image characteristic is closely connected with the fine-line image. The fine-line image is the image of a line of infinite length but infinitesimal width. The flux density is constant along the length of the line image and varies in a direction perpendicular

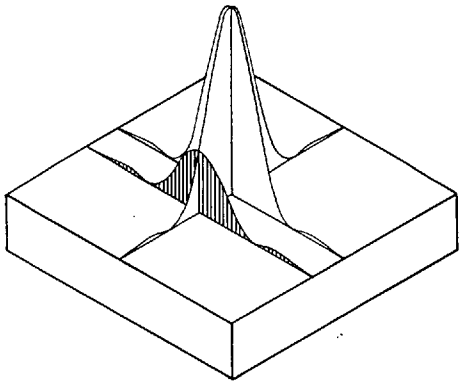


FIGURE 7. *Generation of fine-line image.*

The flux density at a point in a fine-line image is proportional to the area of the section of the point image orientated in the direction of the line.

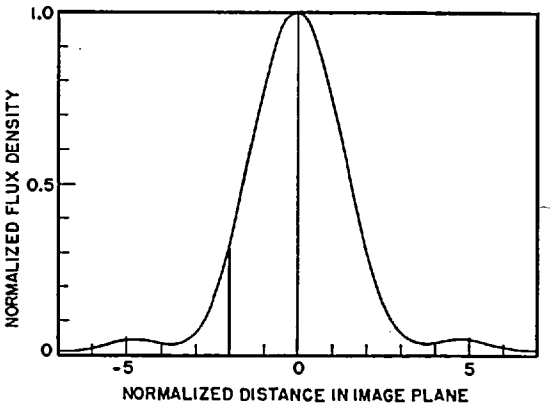


FIGURE 8. *Cross section of fine-line image.*

The vertical line corresponds to the area of the section shown in figure 7.

to the length. It is obtained by integrating the point image in one direction only. For example, the image function for a fine line, the length of which is in the y-direction, is given by

$$L(x)=\int_{-\infty}^{\infty}\varphi(x,y)\,dy. \tag{8}$$

This is illustrated in figures 7 and 8.

The point-image characteristic can be obtained from the fine-line image by

$$K(w_x)=\int_{-w_x/2}^{w_x/2}L(x)\,dx. \tag{9}$$

This is illustrated in figure 9.

Up to this point we have been considering the object to be a point or a fine line and the detector aperture to be a variable slit, the operating mechanism that produces the variation in w. But, because of the interchangeability of the object and aperture, an illuminated variable slit could be used as an object and a fine slit centered on the variable slit image as the detector aperture. This method of obtaining the point-image characteristic is illustrated in figure 10.

The point-image characteristic is also closely connected with the knife-edge image. The latter is related to the point image as follows:

$$S(x')=\int_{-\infty}^{x'}\int_{-\infty}^{\infty}\varphi(x,y)\,dy\,dx. \tag{10}$$

This is shown in figures 11 and 12.

The relationship between the knife-edge image and the fine-line image is given by

$$S(x')=\int_{-\infty}^{x'}L(x)dx. \tag{11}$$

The point-image characteristic is obtainable from $S(x)$ by observing the values of $S(x)$ at $x=-w_x/2$ and $x=+w_x/2$. Then

$$K(w_x)=S(w_x/2)-S(-w_x/2)=\int_{-\infty}^{w_x/2}L(x)dx-\int_{-\infty}^{-w_x/2}L(x)dx=\int_{-w_x/2}^{w_x/2}L(x)dx. \tag{12}$$

These relationships are illustrated in figures 13 and 14.

In summary, it may be pointed out that the point-image characteristic provides a simple yet practical and informative way of evaluating an imaging system. It may be obtained from a variety of test objects, namely, a point, a fine line, a variable-width line, or a knife edge.

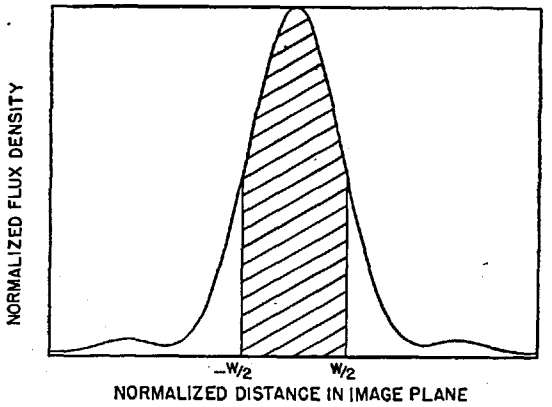


FIGURE 9. *Determination of point-image characteristic from fine-line image profile.*

The cross-hatched area corresponds to the solid in figure 5.

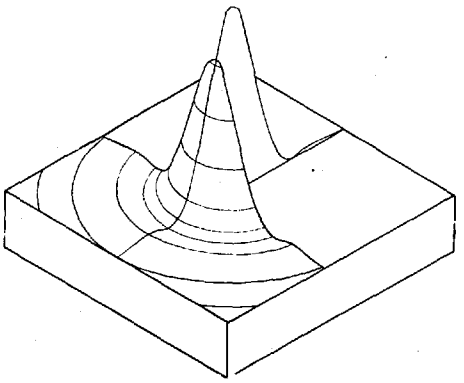


FIGURE 11. *Generation of knife-edge image from point image.*

The flux density at a point in the knife-edge image is proportional to the volume of the solid indicated where the limiting plane has the same orientation as the knife edge and passes through the point in question.

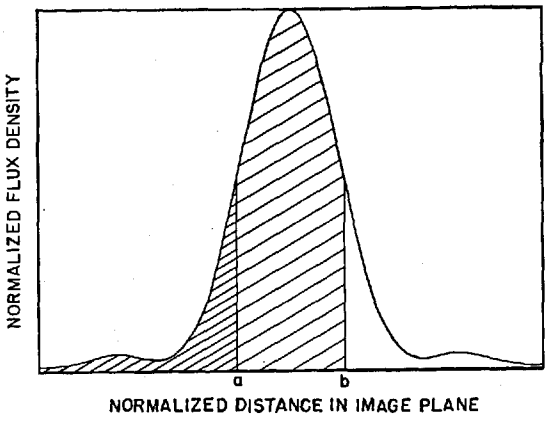


FIGURE 13. *Generation of knife-edge image from fine-line image.*

The flux density at a point in the knife-edge image is proportional to the area under the fine-line image to the left of the corresponding abscissa. The area under the fine-line image between any two abscissa values is proportional to the difference between the corresponding ordinate values of the knife-edge image.

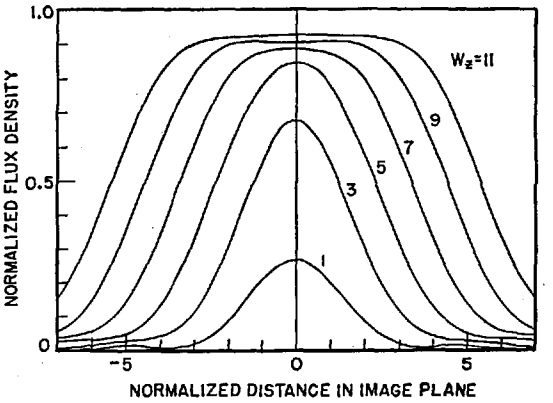


FIGURE 10. *Cross sections of images of finite-width object lines.*

The reduced object line widths are 1, 3, 5, 7, 9, and 11 x-units. Plotting the central flux density as a function of the reduced object line width results in the point-image characteristic curve shown in figure 6.

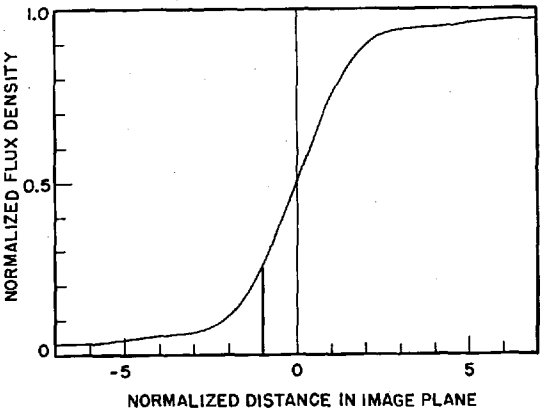


FIGURE 12. *Cross section of knife-edge image.*

The vertical line corresponds to the solid shown in figure 11.

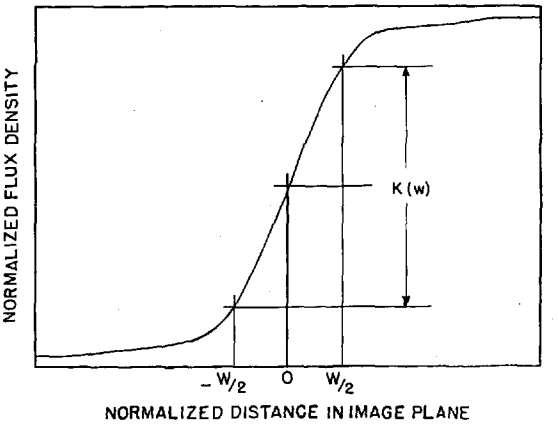


FIGURE 14. *Determination of point-image characteristic from knife-edge image.*

If the abscissa values a and b in figure 13 are made equal to $-w/2$ and $+w/2$ as in this figure, then the ordinate difference shown here is proportional to the corresponding area in figure 13, which itself is proportional to the corresponding value of the point-image characteristic, as indicated in figure 9.

Furthermore, regardless of which way it is obtained, it can be interpreted in terms of any of the objects. This is discussed further in a later section.

3. Evaluation by Fourier Analysis

A general object $O(x,y)$ can be analyzed into a continuum of sinusoidal spatial waves, differing from each other in direction, frequency, amplitude, and phase. The characteristics of these waves are given by the Fourier transform of the object,

$$T_o(\omega_x,\omega_y)=\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} O(x,y)\ \exp\ [-i(\omega_x x+\omega_y y)]dxdy, \tag{13}$$

where $T_o(\omega_x,\omega_y)$ is complex, containing both amplitude and phase factors, and ω_x and ω_y are directional frequency components of the component waves.

A component wave itself is represented by

$$W_c=T_o(\omega_x,\omega_y)\ \exp\ [i(\omega_x x+\omega_y y)]. \tag{14}$$

Consider this to be an object. Then by applying eq (2),

$$I_W(x',y')=\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\varphi(x,y)T_o(\omega_x,\omega_y)\ \exp\ [i\{\omega_x(x'-x)+\omega_y(y'-y)\}]dxdy$$

$$=T_o(\omega_x,\omega_y)\ \exp\ [i(\omega_x x'+\omega_y y')]\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\varphi(x,y)\ \exp\ [-i(\omega_x x+\omega_y y)]dxdy. \tag{15}$$

The integral is the Fourier transform $\Phi(\omega_x,\omega_y)$ of $\varphi(x,y)$, and therefore

$$I_W(x',y')=\Phi(\omega_x,\omega_y)T_o(\omega_x,\omega_y)\ \exp\ [i(\omega_x x'+\omega_y y')]. \tag{16}$$

The function modified by Φ is seen to be simply the component object wave. Each component image wave then will be the product of Φ and the corresponding component object wave. Therefore,

$$T_I(\omega_x,\omega_y)=\Phi(\omega_x,\omega_y)T_o(\omega_x,\omega_y), \tag{17}$$

where T_I is the Fourier transform of the image.

If $I(x,y)$ can be considered to be the object of a second imaging precess, then

$$T_{I_2}(\omega_x,\omega_y)=\Phi_2(\omega_x,\omega_y)T_{I_1}(\omega_x,\omega_y)=\Phi_2\Phi_1 T_o. \tag{18}$$

This can be extended to as many imaging processes as desired.

The transform

$$\Phi(\omega_x,\omega_y)=\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\varphi(x,y)\ \exp\ [-i\omega_x x+\omega_y y)]dxdy \tag{19}$$

is seen to be characteristic of the imaging process. Let us examine it in more detail.

The transform is the double integral of the product of two functions, one being the point image $\varphi(x,y)$ and the other a two-dimensional wave $exp[-i(\omega_x x+\omega_y y)]$.

This two-dimensional wave is sinusoidal in one direction and constant in the perpendicular direction, like a corrugated roof. The direction of the sinusoidal variation is inclined to the x-axis by an angle θ, where $\tan\theta=\omega_y/\omega_x$, and the frequency is given by $\omega_\theta=\sqrt{\omega_x^2+\omega_y^2}$ (see fig. 15). The transform can then be considered to be $\Phi(\theta,\omega_\theta)$.

For $\theta=0$, $\omega_y=0$ and the transform becomes

$$\Phi(\omega_x)=\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\varphi(x,y)\exp[-i(\omega_x x)]dydx=\int_{-\infty}^{\infty}L(x)\exp[-i(\omega_x x)]dx, \tag{20}$$

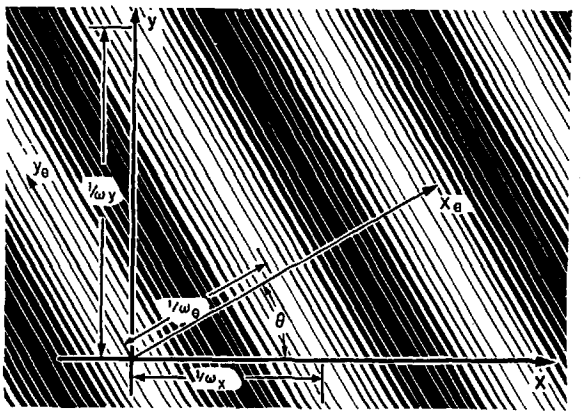


FIGURE 15. *Transformation of coordinates.*

If the direction of the pattern is inclined to the original x-axis by an angle θ, then, in the new coordinate system in line with the pattern, $\theta=\sqrt{\omega_x^2+\omega_y^2}$ and $\tan\theta=\omega_y/\omega_x$.

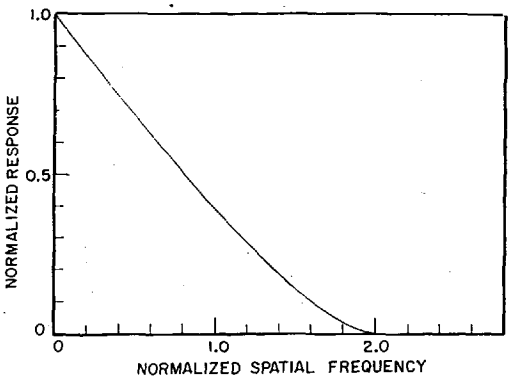


FIGURE 16. *Sine-wave response curve.*

This is the positive side of the cross-section of the Fourier transform of the point image shown in figure 1. It is also the Fourier transform of the fine-line image shown in figure 8.

where $L(x)=\int_{-\infty}^{\infty}\varphi(x,y)dy$ is the fine-line image (fig. 16).

If the x, y coordinates are rotated in x_θ, y_θ through the angle θ (fig. 15), then eq (20) can be generalized into

$$\Phi(\theta,\omega_\theta)=\int_{-\infty}^{\infty}L(x_\theta)\exp[-i(\omega_\theta x_\theta)]dx_\theta. \tag{21}$$

For any θ, the transform Φ of eq (21) is the cross section in the θ-direction of the solid representing the transform $\Phi(\omega_x, \omega_y)$. This is shown in figure 17.

Equation (21) serves as the basis for the experimental determination of Φ, since for each θ it is the transform of the fine-line image oriented with its length perpendicular to the θ-direction. The object can be one in which the flux density varies sinusoidally in one direction, and the detector aperture can be a fine slit.

The object can be represented by

$$O(x_\theta)=A+B\ \exp[i\omega_\theta x_\theta],\quad A\geq B. \tag{22}$$

The constant term is necessary because all flux densities must be positive. The image is given by

$$I(x_\theta)=A+\Phi(\omega_\theta)B\ \exp[i\omega_\theta x_\theta]. \tag{23}$$

Let $M_o=B/A$ be defined as the modulation in the object. The corresponding modulation in the image is $M_I=\Phi B/A$. The transform Φ is obtained by

$$\Phi=\frac{M_I}{M_o}. \tag{24}$$

The use of the modulation automatically compensates for any factor that changes the signal output of the testing device proportionally, such as the transmission factor of the lens, change in brightness of the source, change in gain of the detector amplifier, etc.

It must be remembered that Φ is actually complex, involving both an amplitude and a phase factor. However, if the origin of $\varphi(x,y)$ is properly selected, the phase shift is small. If $\varphi(x,y)$ is symmetrical about its origin, then there is no phase shift involved. In most cases, for evaluation purposes, the phase-shift factor can be neglected.

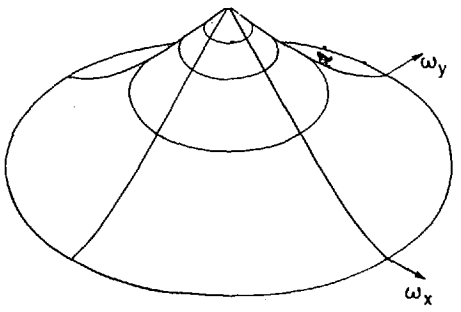


FIGURE 17. *Three-dimensional model of Fourier transform of point image.*

Each radial cross section is the sine-wave response in that direction. Note that the space is frequency space.

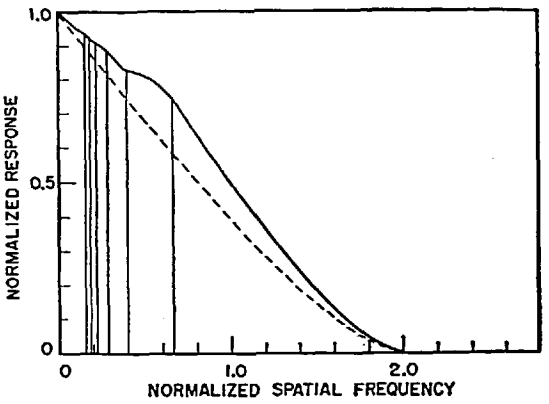


FIGURE 18. *Square-wave response curve.*

The dashed line is the sine-wave response curve from which this was derived. The vertical lines separate regions in which the square-wave image contains different numbers of harmonic components, each region having one more component than its neighbor to the right. In the region farthest to the right, the square-wave image contains only the fundamental.

4. Coordination

If $\Phi(\omega)$ is given, it is possible to predict a similar response curve in which the test object consists of alternate dark and light stripes, a spatial square wave. The square-wave test object can be analyzed into its component harmonics, each of which is attenuated by the value of Φ corresponding to its frequency, and the image is obtained by adding together these attenuated harmonics. For the response curve we are only interested in the peak-to-peak values that are obtained from the values at the centers of the lines and spaces. The square-wave response curve then is given by

$$\psi(\omega)=\frac{4}{\pi}\left[\Phi(\omega)-\frac{1}{3}\Phi(3\omega)+\frac{1}{5}\Phi(5\omega)-\cdots\right]. \tag{25}$$

It should be noticed that there will be only a finite number of terms in the sum because there is a limiting value of ω beyond which Φ remains at zero. This limit exists because of the finite dimensions of the aperture of the system; the larger the aperture the higher the limiting frequency.

Because of the finite range in ω, it is possible to obtain $\Phi(\omega)$ from $\psi(\omega)$ by the inverse process. Here it is necessary to start at the limiting value of ω and work backward. From this limiting value ω_c back to $\omega_c/3$ the sine-wave response is given by

$$\Phi(\omega)=\pi/4\ \psi(\omega) \tag{26}$$

because Φ and ψ are both zero for the odd multiples of ω for ω greater than $\omega_c/3$.

From $\omega_c/3$ to $\omega_c/5$,

$$\Phi(\omega)=\pi/4\ \psi(\omega)+\tfrac{1}{3}\Phi(3\omega), \tag{27}$$

where $\Phi(3\omega)$ may not be zero, and has already been found.

From $\omega_c/5$ to $\omega_c/7$,

$$\Phi(\omega)=\pi/4\ \psi(\omega)+\tfrac{1}{3}\Phi(3\omega)-\tfrac{1}{5}\Phi(5\omega), \tag{28}$$

and so forth.

The relationship between the square-wave response and the sine-wave response is illustrated in figure 18.

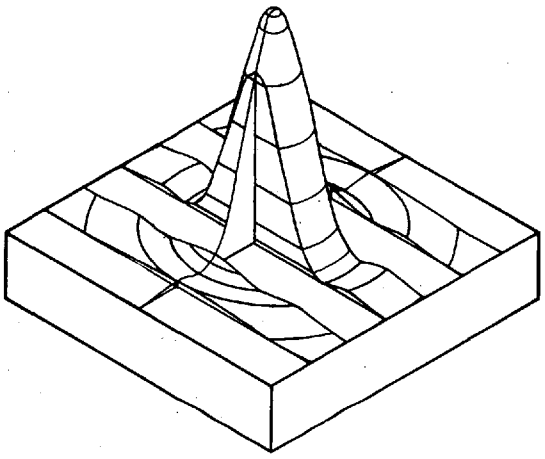


FIGURE 19. *Generation of square-wave response from point image.*

The sum of the volumes of the slabs is proportional to the flux density at the center of a bright line in the image. The sum of the volumes of the slabs that have been removed corresponds to the dark line. The difference between the two is proportional to the square-wave response.

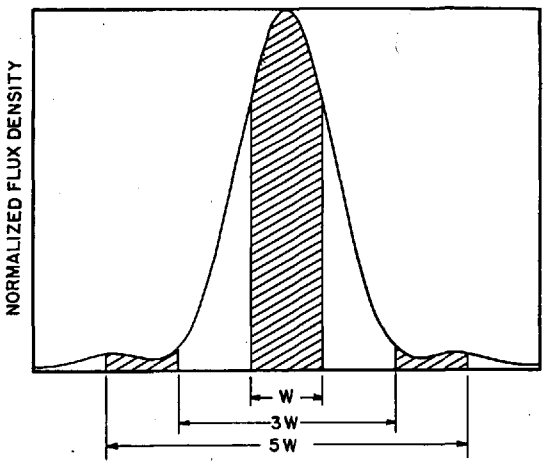


FIGURE 20. *Generation of square-wave response from fine-line image.*

The cross-hatched areas correspond to the slabs in figure 19. The marked widths indicate how the square-wave response can be obtained from the point-image characteristics.

The square-wave response may also be obtained from the point-image characteristic. If we consider the imaging of a square-wave test object as a point-image integration, then the flux density at the center of the image of a bright line in the pattern is proportional to the sum of the volumes of the slabs illustrated in figure 19, or the areas of the stripes in figure 20. The flux density at the center of the image of a dark line is proportional to the sum of the volumes of the slabs (or areas of the stripes) which lie between those illustrated. The sum of these two flux densities is proportional to the total flux in the point or fine-line image, and the square-wave response is given by

$$\Psi(\omega)=E\left(\frac{1}{2\omega}\right)-\left[1-E\left(\frac{1}{2\omega}\right)\right]=2E\left(\frac{1}{2\omega}\right)-1, \tag{29}$$

where $E(1/2\omega)$ is the flux density at the center of the image of a bright line and $1/2\omega=w$ is the width of a line, bright or dark, in the pattern.

It can be seen from figure 20 that

$$E\left(\frac{1}{2\omega}\right)=E(w)=K(w)+[K(5w)-K(3w)]+[K(9w)-K(7w)]+\ldots, \tag{30}$$

where $K(w)$ is the point-image characteristic, so

$$\Psi(\omega)=2\{K(w)+[K(5w)-K(3w)]+\ldots\}-1. \tag{31}$$

It is now apparent that $\Phi(\omega)$, the sine-wave response, may be obtained from $K(w)$, the point-image characteristic. This is done by determining $\psi(\omega)$, the square-wave response, from $K(w)$ and then applying the procedure indicated in eq (26), (27), and (28) to obtain the sine-wave response.[2]

To obtain the point-image characteristic from the sine-wave response is more direct. Consider the test object to be a variable-width line. The transform of the image of this object is obtained by multiplying the transform of the object by the sine-wave response of the system.

[2] This calculation for the case of the unaberrated image was made by the author before he was aware of Steel's equation (see appendix). It was done for 30 different spatial frequencies, using the tabulated values for the knife-edge image given by Struve (see appendix) as well as his approximation for values beyond those tabulated. The agreement between this calculated sine-wave response and Steel's equation is well within the error of calculation.

The transform of this object is of the form

$$T_o = w \frac{\sin \pi\omega w}{\pi\omega w} \tag{32}$$

shown in figure 21, and the transform of the image is given by $T_I = \Phi T_o$. The image itself would be given by the inverse transform

$$I(x_\theta) = \int_{-\infty}^{\infty} \Phi T_o \exp\left[i(\omega_\theta x_\theta)\right] d\omega_\theta, \tag{33}$$

but we are interested only in the value at the center of the line image where $x_\theta = 0$, so

$$K(w_\theta) = I(O) = \int_{-\infty}^{\infty} \Phi T_o d\omega_\theta = 2w \int_0^{\infty} \Phi(\omega_\theta) \frac{\sin \pi\omega w}{\pi\omega w} d\omega_\theta. \tag{34}$$

This equation shows the direct manner in which $K(w)$ may be obtained from $\Phi(\omega)$. For each w selected, the sine-wave response is multiplied by the proper calculated function of the form $\sin x/x$, and the product is integrated graphically or numerically. The result is then multiplied by $2w$, and the value obtained is the desired $K(w)$.

5. Application

It has been shown that an imaging system can be evaluated by means of a description of the point image or, alternatively, a description of its Fourier transform. Each of them has its advantages.

The Fourier approach allows us to combine several systems or to analyze a system into its components, under certain conditions. The principal condition is that each intermediate image formed by each component must be equivalent to a real luminous object having the same flux-density distribution in the image plane. Thus the combining or analyzing process can be applied to a photographic process, a television system, or a system in which each intermediate image is formed on a diffusing surface. It cannot be applied to a telescopic system because the light emerging from the primary image still contains the wave-front deformations produced by the objective and is subject to aberration correction by the ocular.

This does not mean that the Fourier approach cannot be applied to the entire telescopic system. It is just that the whole telescope cannot be precisely evaluated by simply evaluating the components.

Using the Fourier approach also enables one to apply the concepts developed in information and communication theory to imaging processes.

Another virtue of the Fourier approach is that the transform determined for any orientation of the sine-wave pattern in the image plane is a true section of the solid representing the entire transform in the image plane. A similar property does not hold for the point-image characteristic.

It should be noticed also that the sine-wave response of a lens has a finite boundary determined by the aperture of the system, whereas the point-image characteristic is unbounded.

The point-image characteristic, on the other hand, directly provides two basic types of information about the performance of the system, apart from the distribution of flux in the point image. These are the contrast at which an isolated object will be imaged as a function of the object size, and the gradient in the image of the edge of an extended object. The first holds because the point-image characteristic can be obtained by measuring the flux densities at the centers of finite-width line images. The second holds because the characteristic can also be obtained from the image of a knife-edge object.

However, if one knows either the point-image characteristic or the sine-wave response, he can calculate the other, as has been shown. It would probably be preferable to have a research instrument that would determine the sine-wave response, because the transformation to the point-image characteristic is more direct and more suitable for calculating machines than the reverse transformation.

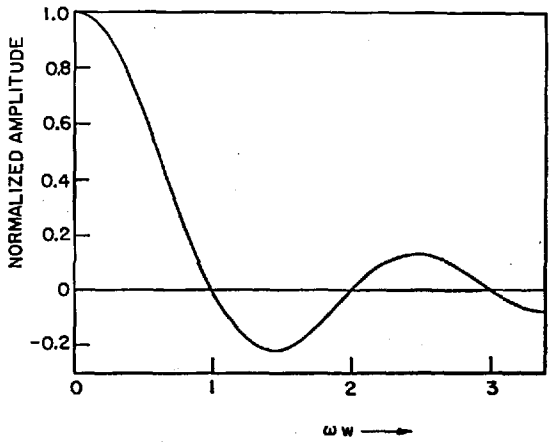


FIGURE 21. *Transform of finite-width line object.*

ω is the frequency of the component wave, and w is the reduced object-line width.

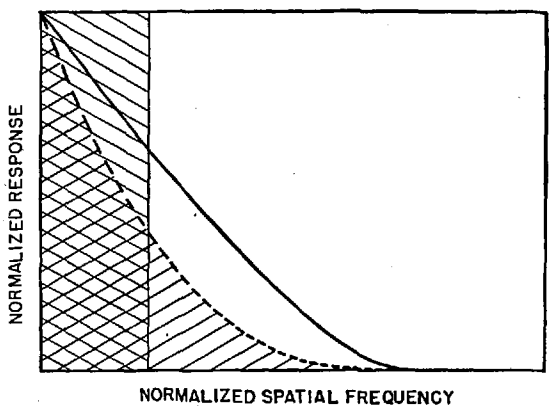


FIGURE 22. *Schade's equivalent bandwidth.*

The solid curve is the sine-wave response curve; the dashed curve is the square of the solid curve; and the rectangular area is equal to the area under the dashed curve. This rectangular area is measured by a single number, its limiting frequency, thereby providing a single number to describe the sine-wave-response curve.

Another problem that comes up is one that is involved in routine testing. It would be desirable to reduce the evaluating curve to a single number with as little loss of significant information as possible, and it would be desirable that this reduction be done automatically in the test procedure.

With respect to the sine-wave response curve, Schade has suggested a reduction in which the curve is squared, ordinate by ordinate, and then integrated. The resulting number is equal to the limiting frequency of a rectangular "response" curve having the same area as the squared sine-wave response curve. This establishes an equivalent bandwidth, shown in figure 22.

The mechanism that would permit this determination directly would be similar to the sine-wave response mechanism, except that a "noise" pattern instead of a sinusoidal pattern would be used. The ideal noise pattern contains all frequencies at equal amplitude but with random phase relations. The fluctuations in the photoelectric output produced by this pattern are fed through a squaring circuit and then integrated. The resulting steady current, indicated on a meter, is proportional to Schade's equivalent bandwidth. One trouble with this system is the difficulty involved in producing an acceptable noise pattern.

Equation (34) indicates another approach to the problem of representing the sine-wave response curve with a single number. The object transform T_0 in this equation can be considered to be a weighting factor for the sine-wave response curve, and the integral to represent the equivalent bandwidth Ω. Then this equivalent bandwidth can be determined by the use of a variable slit or a knife edge, for

$$\Omega=\frac{K(w)}{2w}. \tag{35}$$

For this to be single-valued, w must be fixed, and two convenient possible values appear evident. One is to set $w=1/\omega_c$, where ω_c is the limiting frequency for a theoretically perfect lens having the same aperture as the lens under test. Then the weighting function goes to its first zero at ω_c. This method is illustrated in figure 23.

The other convenient value for w is zero, for the limit of $K(w)/w$ as w goes to zero is the slope at the center of the knife-edge image, as can be seen with the aid of figure 14. As indicated in eq (35), this equivalent bandwidth is given by one-half the slope at the center of the knife-edge image. The weighting function approaches unity for small values of ω as w approaches zero, so this equivalent bandwidth is given by the area under the sine-wave response curve itself. This is illustrated in figure 24. It might be emphasized that this latter relationship is an important and fundamental condition. The slope at the center of a knife-edge image is exactly proportional to the area under the sine-wave response curve.

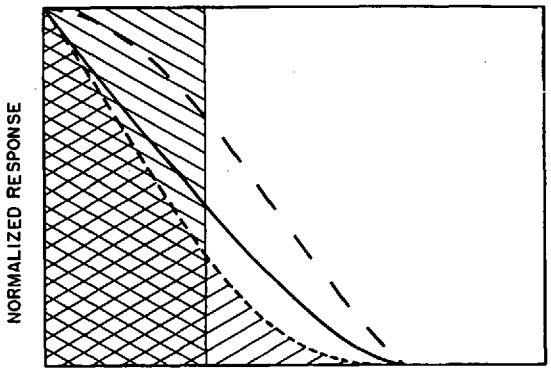


FIGURE 23. *Another equivalent bandwidth.*

The solid curve is the sine-wave response curve; the long-dashed curve is the transform of the finite-width-line object; the short-dashed curve is the product of the other two; and the rectangular area is equal to the area under the short-dashed curve. This provides a different single characteristic number, which can be obtained more easily than Schade's equivalent bandwidth.

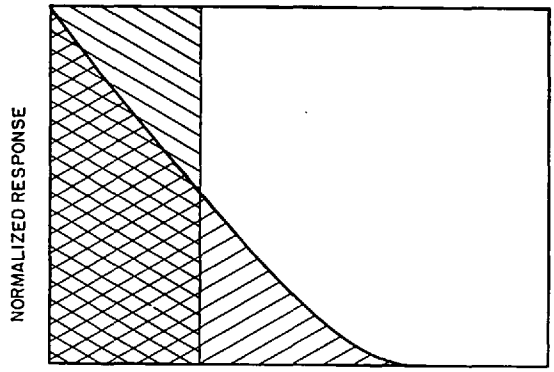


FIGURE 24. *A third equivalent bandwidth.*

This bandwidth is measured by the area under the sine-wave-response curve itself. The area is exactly proportional to the slope at the center of a knife-edge image.

The idea of measuring an equivalent bandwidth may be a good way of reducing the sine-wave response curve to a single number, but this process eliminates one of the advantages of the sine-wave approach, and that is the ability to combine directly a sequence of imaging processes. For most telescopic systems this may be unimportant, but if such a combination is desirable, then perhaps the sine-wave response curve can be characterized well enough by some particular frequency. If the response to some such frequency were established as a measure of quality, then the measure of the quality of the combination is simply the product of the measures of quality of the components.

To summarize, the Fourier approach seems to be more desirable for research and detailed testing, but the determination of the point-image characteristic lends itself to rapid routine testing.

6. Appendix

6.1. Images Produced by an Aberration-Free System With a Circular Aperture in Monochromatic Light

The extended objects are assumed to be illuminated noncoherently. In the following expressions, the unit of displacement in the image is:

$$z=\frac{2\pi a x}{\lambda d}, \tag{36}$$

where a is the radius of the circular aperture, x is the distance in the image plane from the center of the image, d is the distance from the image plane to the aperture, and λ is the wavelength of light.

a. Point Image

This is well known, and its section is given by

$$\varphi(z)=4\left(\frac{J_1(z)}{z}\right)^2. \tag{37}$$

It is illustrated in figures 1 and 2.

b. Fine-Line Image

An expression for the cross section of the fine-line image, implicit in the original work of Struve [1], is explicitly given by Steel [2] as

$$L(z)=\frac{3\pi}{8}\frac{H_1(2z)}{(2z)^2}, \tag{38}$$

where H_1 is what is known as the Struve function, and has been tabulated.[3] The cross section of the fine-line image is illustrated in figure 8.

c. Knife-Edge Image

Struve [1] developed this in a series expansion and tabulated it for z up to 15. If the edge is oriented so that the gradient for the image is positive, then the image is given by

$$S(z)=\frac{1}{2}+\frac{2}{\pi^2}\left\{\frac{3}{1}\frac{2^2z}{1^2\cdot3^2}-\frac{5}{3}\frac{2^4z^3}{1^2\cdot3^2\cdot5^2}+\frac{7}{5}\frac{2^6z^5}{1^2\cdot3^2\cdot5^2\cdot7^2}-\cdots\right\}. \tag{39}$$

Struve also gave a simple approximation, which is in error by less than 0.1 percent for $Z>7$ and 1.0 percent for $Z>3$.

$$S(z)\cong1-\frac{2}{\pi^2z}. \tag{40}$$

The knife-edge image is illustrated in figure 12.

d. Finite-Width Line Image

If the width of the Gaussian image is w_z, then

$$\overline{L(z)}=S\left(z+\frac{w_z}{2}\right)-S\left(z-\frac{w_z}{2}\right). \tag{41}$$

This family of images is illustrated in figure 10.

e. Point-Image Characteristic

This is given by

$$K(w_z)=S\left(\frac{w_z}{2}\right)-S\left(-\frac{w_z}{2}\right)=2S\left(\frac{w_z}{2}\right)-1. \tag{42}$$

Combining this with eq (39) gives

$$K(w_z)=\frac{4}{\pi^2}\left\{\frac{3}{1}\frac{2w_z}{1^2\cdot3^2}-\frac{5}{3}\frac{(2w_z)^3}{1^2\cdot3^2\cdot5^2}+\frac{7}{5}\frac{(2w_z)^5}{1^2\cdot3^2\cdot5^2\cdot7^2}-\cdots\right\}. \tag{43}$$

This is illustrated in figure 6.

f. Sine-Wave Response

Steel [2] gives the following expression for the sine-wave response:

$$\Phi(\omega)\begin{cases}=\dfrac{2}{\pi}\left(\arccos\dfrac{\omega}{2}-\dfrac{\omega}{2}\sqrt{1-\dfrac{\omega^2}{4}}\right), & 0<\omega<2\\ =0, & \omega>2,\end{cases} \tag{44}$$

where $\omega=\pi/w_z$, and w_z is the wavelength of the sinusoid. Notice that there is an absolute limit to the frequency of the pattern that a lens can form, this limit being where the wavelength of the pattern in z units is equal to π.

The frequency, ν, in the image plane is given by

$$\nu=\frac{\omega}{2N\lambda}, \tag{45}$$

where $N=d/2a$ and λ is the wavelength of light. If λ is in millimeters, ν is in cycles per millimeter.

Figures 16 and 17 show the sine-wave response.

---

[3] E. Jahnke and F. Emde, Tables of functions with formulas and curves, 4th ed. (Dover Publications, New York, N. Y., 1945).

7. References

The data for the curves in the illustrations were calculated from the information contained in the following references:

[1] H. Struve, Beitrag zur theorie der diffraction an fernrohren, Ann. Physik Chemie **17,** 1008 (1882).

[2] W. H. Steel, Étude des effets combinés des aberrations et d'une obturation centrale de la pupille sur le contraste des images optiques. Premiére Partie, Rev. opt. **32,** 4 (1953).

Steel provides information on the point, disk, fine-line, knife-edge, sine wave, and square wave images. Material on the sine wave or Fourier transform approach is contained in the following references.

[3] O. H. Schade, A new system of measuring the specifying image definition, NBS Circular 526 (April 29, 1954).

[4] O. H. Schade, Image gradation, graininess, and sharpness in television and motion picture systems. Part II, J. Soc. Motion Picture Television Engrs. **58,** 181 (1952).

[5] P. Elias, D. S. Grey, and D. Z. Robinson, Fourier treatment of optical processes, J. Opt. Soc. Am. **42,** 127 (1952).

[6] E. W. H. Selwyn, The photographic and visual resolving power of lenses. I. Visual resolving power, Phot. J. **88B,** 6 (1948).

Schade is the first to build a practical machine for obtaining a sine-wave analysis of an optical system. Selwyn applies the Fourier treatment to the prediction of resolving power.

Work on the photoelectric examination of point images through circular apertures is reported in the following paper:

[7] R. E. Hopkins, Measurements of the energy distribution of optical images, NBS Circular 526 (April 29, 1954).

Interest in the knife edge is shown in the following references:

[8] W. Weinstein, Criteria of image-forming quality in photographic objectives, Phot. J. **91B,** 138 (1951).

[9] R. V. Shack, A proposed approach to image evaluation, NBS Circular 526 (April 29, 1954).

Shack also discusses the variable slit. Hopkins is now using knife edges in his work.

WASHINGTON, June 1, 1955.

Reprinted from JOURNAL OF THE OPTICAL SOCIETY OF AMERICA, Vol. 46, No. 9, 755–757, September, 1956
Printed in U. S. A.

Outline of Practical Characteristics of an Image-Forming System*

ROLAND V. SHACK
National Bureau of Standards, Washington, D. C.
(Received September 6, 1955)

(Presented at the Symposium on Formation and Evaluation of Images, University of Rochester, June 15–18, 1955.)

The relationships between certain fundamental characteristic functions relating to an optical image-forming system are indicated, some being observable experimentally and others being unobservable. Of the observable characteristics two are especially useful in the analysis of the image-forming properties of the system. One of these represents the image of a point object and the other the Fourier transform of this point image. The two are considered for their practicality as bases for objective laboratory techniques in image evaluation. The transform of the point image is practical as it stands whereas the point image itself is not. A function directly related to the point image and to a number of other simple objects is suggested as a practical representative function. The practical functions are correlated and the relative merits of the two approaches are considered.

1. INTRODUCTION

DURING the past few years the National Bureau of Standards, as well as a number of other laboratories, has been concerned with the problem of developing an objective procedure for evaluating the quality of images formed by optical instruments. One general approach has been to try to find some physical parameter which can be measured objectively and then to try to correlate this parameter with some already existing quality criterion. Although it is true that, if successful, such an empirical approach will eliminate the subjective element in the determination of image quality, it is made under the assumption that the chosen criterion is in itself a satisfactory index of image quality. There is reason to doubt the validity of this assumption.

A more fundamental approach is to analyze the image-forming process as a phenomenon, the aim being to characterize the process in as general and yet informative way as possible, consistent with practical instrumentation. New criteria of image quality would of course be expected to be developed. The present paper is intended to be an outline of such an approach.†

2. OBSERVABLE AND UNOBSERVABLE CHARACTERISTICS

Figure 1 illustrates the relationships existing between some of the more significant or characteristic functions of an optical image-forming system. The actual functions are functions of two spatial dimensions, but to avoid complexity in the figures, they are drawn as functions of one dimension. These curves can be considered to represent cross sections of the more complete functions.

Perhaps the most elementary characteristic function is that represented by Fig. 1(a). This curve represents the variation in optical path length between an object point and its Gaussian image point plotted across the pupil of the system. It can also be interpreted as the wave-front deformation across the pupil.

From this function can be derived the complex amplitude of the wave function in the pupil [Fig. 1(b)], another characteristic function. The Fourier transform of this complex amplitude in the pupil gives the complex amplitude of the wave function in the image plane [Fig. 1(c)]. The product of this third characteristic function with its complex conjugate yields a function which has been traditionally called the intensity distribution in the image of a point [Fig. 1(f)].

Of the functions so far considered, only the last is an observable characteristic. That is, it can be directly ascertained by experimental means. The others cannot, and must be considered to be unobservable characteristics.

However, the point image is not the only observable characteristic. The product of the complex amplitude in the pupil [Fig. 1(b)] with its complex conjugate

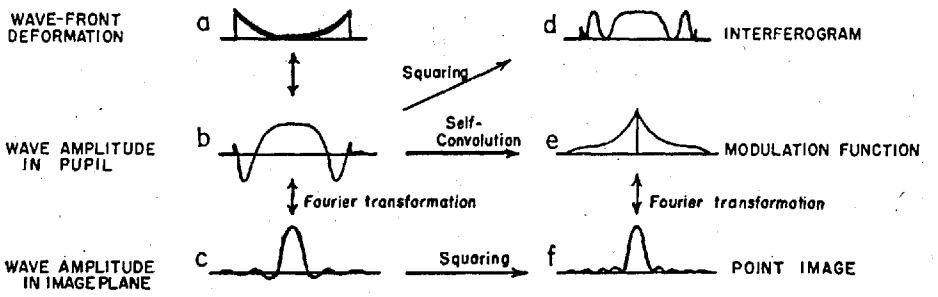


FIG. 1. Characteristic functions of an optical image-forming system.

* This work was done in connection with a project sponsored by Army Ordnance, Frankford Arsenal.
† Most of the material in this paper is somewhat more thoroughly dealt with in another paper by the author, "Characteristics of An Image-Forming System," to be published in the *Journal of Research of the National Bureau of Standards.*

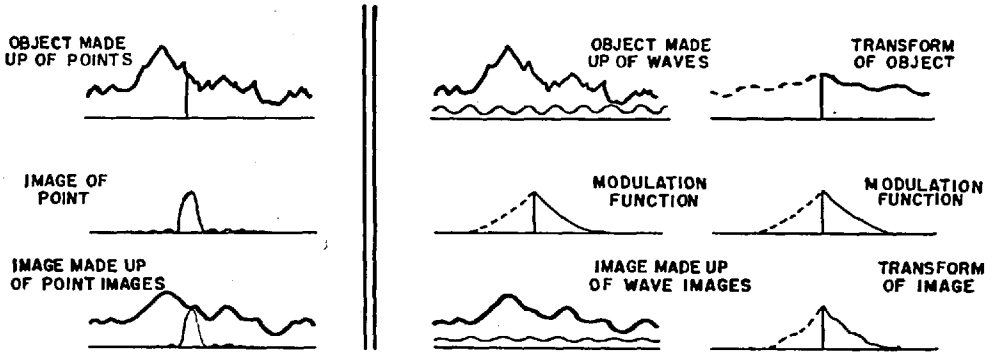


FIG. 2. Object-image relationships.

yields the function represented by the curve of Fig. 1(d), the kind of curve obtainable by means of an interferometer. This last function is also observable, and from it can be inferred the nature of the wave-front deformation [Fig. 1(a)], although not unambiguously.

A third observable function, only recently investigated, is that represented by Fig. 1(e). This is the Fourier transform of the point image [Fig. 1(f)]. It can also be obtained by convolving the complex amplitude in the pupil [Fig. 1(b)] with its complex conjugate.

We therefore have six interrelated functions characteristic of an optical image-forming system. Three of these (on the left of Fig. 1) are unobservable, cannot be directly obtained experimentally. The other three are observable. The interferogram [Fig. 1(d)] is valuable primarily because of its indication of the nature of the wave front in the pupil, but it is very difficult to infer from it the manner in which an image will be formed by the system. However, both of the remaining observable functions are concerned with object-image relationships.‡

3. OBJECT-IMAGE RELATIONSHIPS

That both are equally applicable to the problem of relating the image to the object is indicated in Fig. 2. Any object can be looked at as a summation of points and its image as the summation of the corresponding point images. The point image function is the characteristic modifying function.

On the other hand, the object can be considered to be the summation of a set of sinusoidal waves distributed in the object plane. These component waves, differing from each other in amplitude, frequency, phase, and orientation, are spatially distributed waves. That is, they are spatial, not temporal, sinusoidal variations in brightness throughout the object plane. The image consists of the summation of the images of these component waves.

Now an image-forming system, if linear, does not affect the frequencies, orientations, or sinusoidal character of the component waves. It can only affect their amplitudes and phase relationships. The function which describes these modifications as a function of the spatial frequencies and orientations of the component waves is that represented by the curve in Fig. 1(e), the Fourier transform of the point image. This function has many names in the literature, e.g., sine-wave response, system function, modulation function, transmission factor, contrast reduction factor, etc.

It might be noted, as indicated in Fig. 2, that, since an object can be considered to be made up of component sinusoidal waves, it can also be represented by a function in which the amplitudes and phase relationships of the component waves are plotted against their frequencies and orientations. This function is the Fourier transform of the object function. Multiplying this by the modulation function yields the Fourier transform of the image. An inverse transformation will yield the image function itself.

Now an image-forming process is useless unless the image is picked up by a detector, which may be the eye, a photographic emulsion, or a photoelectric detector. The detector itself modifies the image and must be considered to be a secondary image-forming process. Its inclusion in the over-all process is easily done with the approach involving the modulation function because the modulation function of an entire system is the product of the modulation functions of the components.

The approach involving the point image, however, requires a convolution for every component image-forming system, rather than a simple product.

4. EXPERIMENTAL DETERMINATION

Although any given optical system may be intended for use with one type of detector (i.e., the eye, a photographic emulsion, or a photoelectric detector), it is not necessary that a laboratory performance test use the same detector, provided that the effect of the detector on the image is predictable. The choise of detector for laboratory use should be made on the basis of convenience, controllability, and objectivity of the data obtained. The photoelectric detector in most cases will best meet the requirements.

However to avoid integrating the fine detail in the

‡ Strictly speaking, both are significant only when the objects with which the system is to be used can be considered to be noncoherent extended luminous sources.

image, the photoelectric detector must be used behind a small aperture, and provision must be made for producing a relative displacement between the aperture and the image in order to explore various parts of the image.

The required smallness of the detector aperture places a premium on the amount of energy available, but the development of the photomultiplier tube has made this technique feasible. The shape of the detector aperture can be rigidly controlled and the results obtained should be highly reproducible. However, the photomultiplier will generate noise which will interfere with weak signals.

It should be pointed out that the electrical signal obtained from the detector cannot distinguish between the effect of the object and the effect of the detector aperture. The test object and the detector aperture are interchangeable. For example, the signal obtained by using a luminous disk for a test object and a pinhole for the detector aperture will be the same as the signal obtained by using a pinhole for a test object and a circular hole for the detector aperture, provided that the Gaussian geometry is the same. Also there are a number of ways of producing the relative motion between the image and the detector aperture. The detector aperture can be moved across a stationary image, or the image can be moved across a stationary detector aperture by moving the object or by rotating or rocking some element in the optical train.

The modulation function of an optical system can be obtained experimentally by using an object whose transform is known and a detector aperture whose modulation function is known, and determining the transform of the output signal. The modulation function of the optical system is given by the ratio of the latter to the product of the former. There are a number of ways of doing this, but the maximum energy in the signal for any given frequency will be obtained by using a sine-wave or a square-wave pattern for one of the functions.

The point image is not so easily obtained. The effect of the detector aperture cannot be divided out as in the case of the modulation function, and so the detector aperture must be made as minute as possible. This reduces drastically the energy available. To make matters worse, a large part of the energy in the point image may be distributed in the flare about the central peak, the area of the flare region being so large that the intensity is very low at any given point. The effect of this flare is noticeable with extended objects, although it may seem to be negligible in the image of a point.

A test object which is more easily handled and which shows the effect of the flare is the well-known knife-edge. If the data obtained with a knife-edge object is recorded as a curve in which the intensity difference between two points located at equal distances on either side of the geometric image is plotted against the distance between the points, it can be shown that this curve can be obtained from or interpreted in terms of a variety of objects, namely, a point, a fine line, a variable slit, and, of course, a knife-edge. Such a curve would be a useful characteristic, related to the point image, for determining image quality, and in the rest of this paper will be called the point-image characteristic.

5. COORDINATION

If the modulation function is known, this point-image characteristic is easily obtained because the Fourier transform of a variable slit object is simple and well known. When multiplied by the modulation function it yields the transform of the image, and because only the value at the center of the image is needed, the desired value for a given width is obtained by a direct numerical or graphic integration of the product of the two curves, the oscillatory component of the inverse transformation dropping out.

An interesting result of this relationship is that the area under the curve representing the modulation function itself is exactly proportional to the slope at the center of the knife-edge image. This is also the Strehl definition for the image of a fine line.

The modulation function can also be obtained from the point-image characteristic. For a given frequency the response of the system to a square-wave bar pattern can be predicted from the point-image characteristic by a simple addition and subtraction of selected ordinates. Once a square-wave response curve is obtained, the sine-wave response curve (modulation function) can be obtained by subtracting out the harmonic components.

6. MERITS

The modulation function has a number of features to recommend it for image evaluation. It is a bounded function, relatively simple in form, relatively easy to measure, and easily manipulated. It allows the convenient analysis of cascaded systems. Its use opens the possibility of applying to optical image formation many of the concepts of communication theory and allied fields.

The point-image characteristic has the advantage of presenting graphically the image characteristics of a number of important types of objects. It is also easy to obtain. However, it is an indirect measure of the properties of a point image and is relatively difficult to manipulate.

The Use of Edge Gradients in Determining Modulation-Transfer Functions

FRANK SCOTT, RODERIC M. SCOTT, and ROLAND V. SHACK

Reprinted from

PHOTOGRAPHIC SCIENCE AND ENGINEERING

Vol. 7, No. 6, pp. 345–349, November–December 1963

PHOTOGRAPHIC SCIENCE AND ENGINEERING
Volume 7, Number 6, November–December 1963

The Use of Edge Gradients in Determining Modulation-Transfer Functions

FRANK SCOTT, RODERIC M. SCOTT, AND ROLAND V. SHACK,
The Perkin-Elmer Corporation, Norwalk, Conn.

A method is described for obtaining the modulation-transfer function from an edge in an image, making the procedure especially useful for evaluation of images not containing targets. Microdensitometric data obtained from an image-edge is treated to yield the square-wave response from which the sine-wave response, or the modulation-transfer function, is determined. The method only involves the taking of finite sums and differences. In addition to a discussion of the method, a detailed example is given of two typical applications.

The increasing use of the modulation-transfer function[1] in the design and performance evaluation of optical and photo-optical systems has been accompanied by an increasing variety of methods of its measurement. One method for obtaining the modulation-transfer function, based on a principle described by one of the authors,[2] is discussed.

The procedure to be described results in a good estimate of the modulation-transfer function of systems producing images which do not contain targets normally used for measurement purposes. Instead, data obtained from an edge in the image is treated to yield the square-wave response from which the sine-wave response, or the modulation-transfer function, is determined. This method, which theoretically is exact, only involves the taking of finite sums and differences. The accuracy of the procedure is limited primarily by the practical aspects of microdensitometry and graphical techniques. The theoretical basis and the mathematical treatments involved in modulation-transfer functions have been described[3–7] and will not be discussed in this paper.

Description of Method

If an edge in an image is assumed to have a step-function brightness distribution, like a knife-edge

Received 18 July 1963.

1. E. Inglestam, *Phot. Sci. Eng.*, **5**: 282 (1961).
2. R. V. Shack, *J. Res. Natl. Bur. Std.*, **56**: 245 (1956).
3. G. C. Higgins and F. H. Perrin, *Phot. Sci. Eng.*, **2**: 66 (1958).
4. F. H. Perrin, *J. SMPTE*, **69**: 151 (1960).
5. F. H. Perrin, *J. SMPTE*, **69**: 239 (1960).
6. R. M. Scott, *Phot. Sci. Eng.*, **3**: 201 (1959).
7. R. L. Lamberts, G. C. Higgins, and R. N. Wolfe, *J. Opt. Soc. Am.*, **48**: 487 (1958).

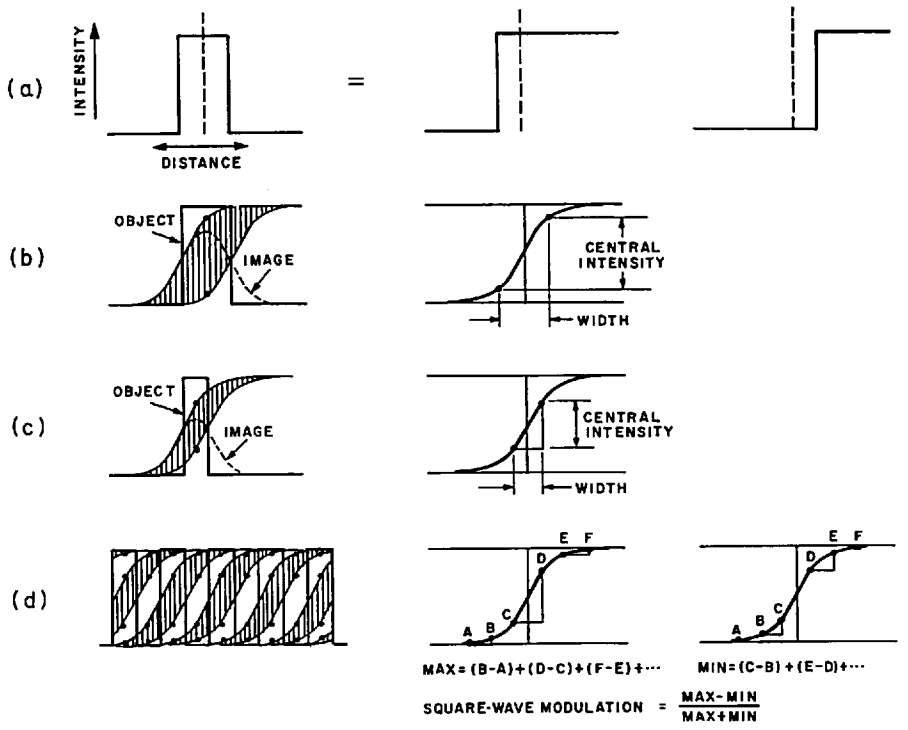


Fig. 1. Square-wave modulation from knife-edge image.

image, the image of any structure made up of sharp edges (Fig. 1a) can be calculated by the corresponding addition and subtraction of the images of each edge (Fig. 1b). Specifically, the intensity of the image of a bar of finite width can be calculated by the difference of two displaced edge-images or by the ordinate difference between two points on the knife-edge image curve plotted against the midpoint between the abscissas of the points, the separation of the points being equal to the width of the bar (Fig. 1, b and c). The central intensity of the image is obtained when the midpoint is at the center of the sharp-edge image curve.

If the object consists of bars sufficiently close together for the successive knife-edge images to overlap appreciably (Fig. 1d), the intensity at

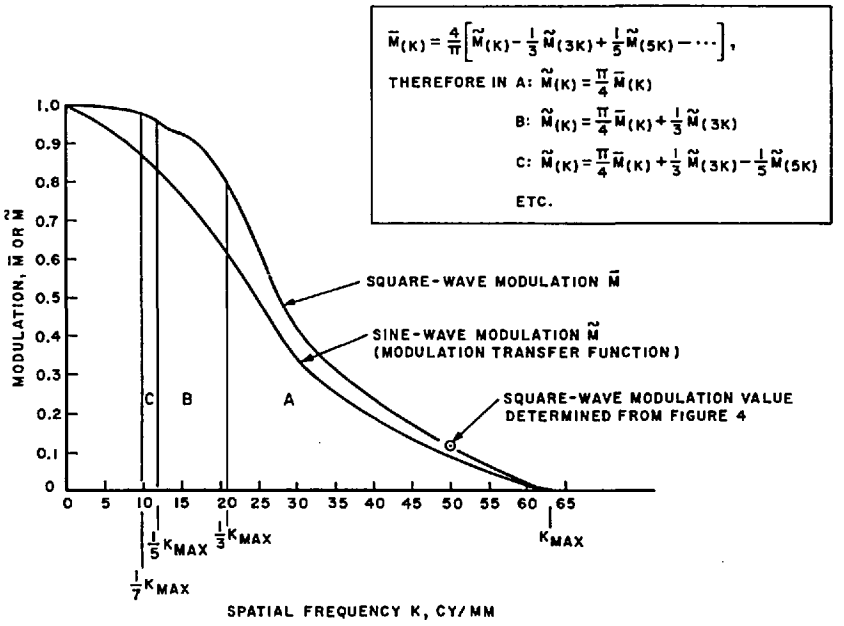


Fig. 2. Sine-wave modulation from square-wave modulation.

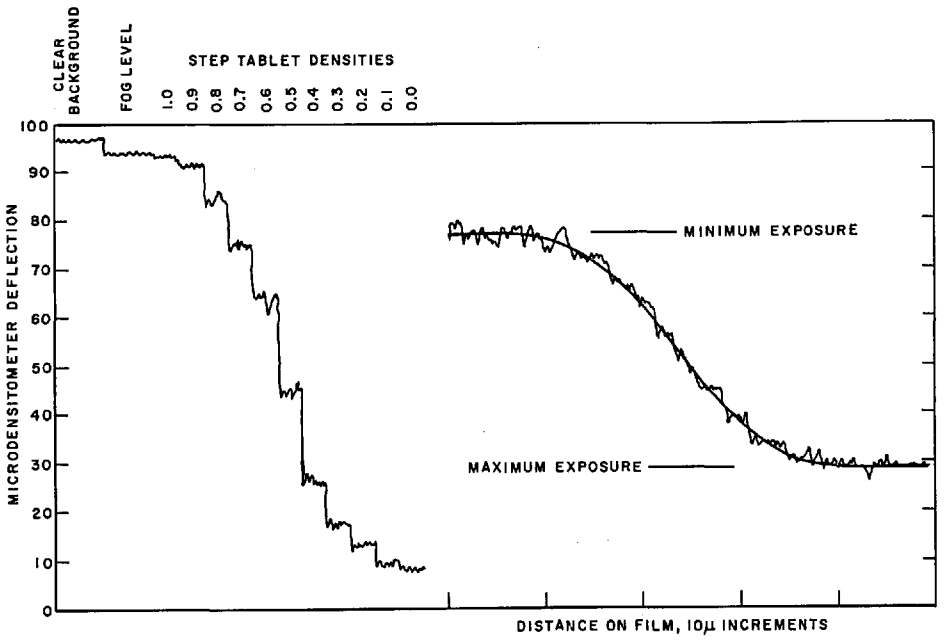


Fig. 3. Microdensitometer trace of sensitometric exposure (*left*) and image edge (*right*).

any point is the sum of the contributions from all the bars in the vicinity, or alternatively is the sum of the ordinate differences between pairs of points along the knife-edge image curve representing the bars. Thus, if the object is a square wave (bars and spaces all equal), the intensity at the center of a bar or space[8] is obtained from the ordinate values of points equally spaced along the knife-edge image curve and straddling the center. The square-wave modulation is obtained from these two intensities by the usual formula. This procedure can be repeated for different frequencies in order to build up the square-wave modulation curve (Fig. 2).

Because the transfer function has a finite upper limit in spatial frequency, any square-wave image has only a finite number of harmonic components. In fact, for all spatial frequencies from ⅓ maximum to the maximum itself, the square-wave image is itself a sine-wave, because only the fundamental can have a nonzero value. Consequently it is possible to derive the sine-wave modulation from the square-wave modulation by starting at the high-frequency end and working backwards introducing harmonic components when appropriate.

Example of Method

The determination of the modulation-transfer function of an aerial camera system is taken as an example of the use of the method. The method is, however, applicable to other optical and photographic systems or components as is briefly discussed later.

The first and most important step is to provide a sensitometric exposure on the film to be evaluated. This is done best in the camera at almost the same time as the exposure. No absolute intensity calibration is required, but care should be taken that the relative exposures are accurately known. A uniformly illuminated density step tablet is satisfactory if the duration of the exposure and the spectral quality of the illumination simulate the scene exposure. The exposures may be made at different times if the interval to processing is long enough so that any time effects between exposures and processing are minimized. From this sensitometric exposure a modified form of a characteristic (H & D) curve will be developed and the accuracy of the method depends on the accuracy of this curve.

The steps in this method, to proceed from an edge in a picture to the transfer function of the system, are as follows:

1. Select an edge in the scene which is straight for many resolution elements and is known to have a step-function brightness distribution. Examples are shadows of straight edges of buildings on smooth surfaces and the ridge of a peaked roof with different illumination on the two sides. The edge must separate two areas of uniform density which are large enough to be well resolved.

2. Trace the sensitometric density steps and the edge with the same slit and settings of the microdensitometer. The slit must be long enough (and thus the edge) to give a good trace with a minimum of grain noise. The slit must be narrow enough so that its transfer function $T(k)$ does not obscure the function of the system being evaluated.

$$T(k) = \frac{\sin \pi\, wk}{\pi\, wk}$$

where

w = effective width of slit

k = spatial frequency

Figure 3 illustrates a typical tracing.

[8] R. L. Lamberts, *J. Opt. Soc. Am.*, 48: 490 (1958).

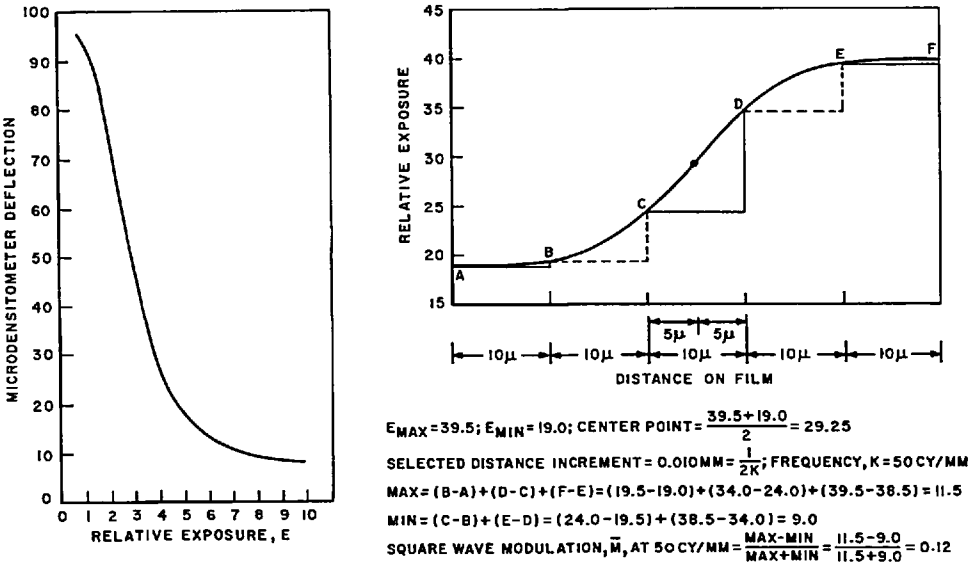


Fig. 4. Curves of sensitometric relative exposure vs. microdensitometer deflection (*left*) and relative exposure vs. distance on film of edge image (*right*).

3. Convert the deflections of the densitometer into relative exposures by plotting deflection vs. relative exposure from the sensitometric images. The left side of Fig. 4 is typical.

4. Draw a smooth curve through the edge trace (care at this point will be repaid later), convert the deflection axis to relative exposure using the curve made in Step 3, and draw the curve of relative exposure vs. displacement as shown on right side of Fig. 4. If the trace shows excessive deflections due to granularity, scan several sections of the edge and graphically average the traces.

5. Find the center of the central intensity. This point on the curve is the average exposure; that is,

$$\text{center point} = (E_{max} + E_{min})/2$$

6. Select a distance increment which is equal to $\frac{1}{2}k$ where k is a spatial frequency at which the transfer function is expected to have a value greater than zero. Center the selected distance increment on the center point and divide the remainder of the curve abscissa into the selected distance increments, as shown in Fig. 4. Determine the ordinate difference of each distance increment. Add the differences of every other increment (solid lines in Fig. 4), and consider this to be the MAX value. The MAX set must include the center increment. Add the differences of the remaining increments (dotted lines in Fig. 4), and consider this the MIN value.

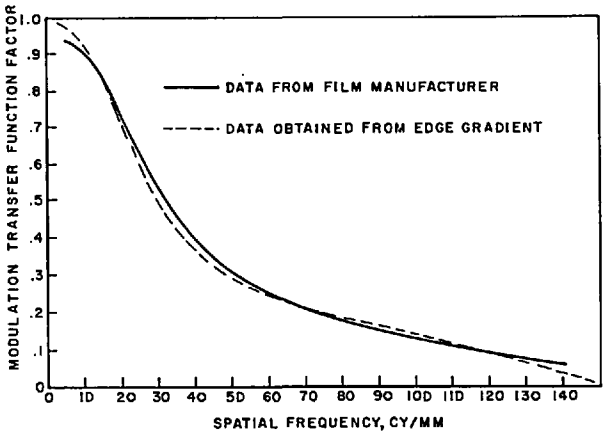


Fig. 5. Measurement of modulation-transfer function of Kodak Plus-X Aerecon Film, Type No. 8401, developed in D-19 developer for 8 min at 68° F.

7. Determine the square-wave modulation value of the selected frequency k, using the MAX and MIN values obtained in Step 6.

$$\text{Square-wave modulation} = (\text{MAX} - \text{MIN})/(\text{MAX} + \text{MIN})$$

As a check, MAX + MIN should equal $E_{max} - E_{min}$.

8. Repeat Steps 6 and 7 several times in each case selecting a different distance increment (spatial frequency) which will enable the plotting of the square-wave modulation curve as shown in Fig. 2.

9. Change the square-wave modulation curve to a sine-wave modulation curve. This is accomplished by treating sections A, B, C, etc. of the curve in Fig. 2 separately. In section A, consisting of the spatial frequency range of $\frac{1}{3}k_{max}$ to k_{max}, multiply the square-wave modulation $\bar{M}$ by $\pi/4$ to obtain the sine-wave modulation $\tilde{M}$.

In Section A,

$$M_k = (\pi/4)\bar{M}_k$$

In Section B ($\frac{1}{3}k_{max}$ to $\frac{1}{3}k_{max}$), change the square-wave modulation to sine-wave modulation by

$$M_k = (\pi/4)\bar{M}_k + \tfrac{1}{3}M_{3k}$$

Likewise, in Section C,

$$\tilde{M}_k = (\pi/4)\bar{M}_k + \tfrac{1}{3}\,\tilde{M}_{3k} - \tfrac{1}{5}\,\tilde{M}_{5k}$$

Correction for harmonic components above the 5th is usually not necessary as interpolation of the curve to a modulation value of 1.0 is adequate.

10. Remove the transfer function of the microdensitometer objective and slit which can be determined by scanning a knife-edge and applying the procedure described in this paper.

It is to be noted that this method gives the transfer function of the aerial camera system without a knowledge of the actual intensities at the target, the haze[9–10] in the atmosphere, or the wide-angle scattering in the system. This is because the brightness in the scene was assumed to be as measured photometrically in the image on the film. Thus, these effects are not determined by this method without additional knowledge of the target. Nevertheless, in those cases where only unstandardized targets are available, the method yields considerable useful information on aerial camera system performance.

Other Applications

The procedure can be applied to determine the modulation-transfer function of photo-optical devices other than aerial camera systems. For example, as mentioned in the application discussed above, the modulation-transfer function of a microdensitometer can be determined by this method.

Another example of application is the measurement of film modulation-transfer functions. Figure 5 shows the results obtained by this method and that obtained by the manufacturer.[11] The particular procedure followed in this case was:

1. An edge-image was produced on a high-resolution film.
2. The edge-image was contact-printed on the film being evaluated. The contrast of the edge-image produced in Step 1 and the exposure incident on the film in Step 2 were such that the minimum density produced on the film in Step 2 was above gross fog.
3. The film was processed along with a sensitometrically exposed film sample.
4. The processed film was scanned on a microdensitometer and the resulting edge-trace changed to an exposure–distance plot using the microdensitometer trace of the sensitometric exposure.
5. The edge of Step 1 was scanned on the microdensitometer. After application of the procedure descri ed in this paper, the transfer function of the microdensitometer plus the exposing edge was determined.
6. From the exposure–distance plot of Step 4 a modulation-transfer function was produced which when divided by the modulation-transfer function produced in Step 5 yielded the film modulation-transfer function shown in Fig. 5.

Conclusions

The procedure described above gives the real part of the optical-transfer function[1] which, in the case of symmetrical line-spread functions, is equal to the modulation-transfer function. When unsymmetrical edge-traces are encountered, the imaginary part of the optical-transfer function can be determined by repeating the process using distance increments not centered on the central intensity. The repeated procedure should employ distance increments centered on points A, B, C, D, . . . in Fig. 4 and yields the Fourier sine transform of the line-spread function. As a check, this function should have its origin at zero. The modulation-transfer function is obtained by combining the first-obtained transfer function (which mathematically is the Fourier cosine transform) with the function obtained by the repeated process by taking the square root of the sum-of-the-squares.

The method for determining the modulation-transfer function discussed in this paper should not be employed when more conventional methods and suitable targets[10,12] are available yielding more accurate data. The procedure described does, however, offer a fairly rapid and accurate determination of modulation-transfer functions and is particularly useful with images not containing sine-wave or other types of targets.

Acknowledgment

The authors wish to thank Dr. R. E. Hufnagel for suggestions and Miss L. Bozak for invaluable assistance in the development of the procedure.

9. M. D. Rosenau, Jr., *Phot. Sci. Eng.*, **6**: 265 (1962).
10. M. D. Rosenau, Jr., F. Scott, and W. F. Thiessen, Jr., *Phot. Sci. Eng.*, **7**: 92 (1963).
11. *Manual of Physical Properties of Aerial and Special Materials*, Section 19, Eastman Kodak Co., Rochester, N.Y., 1963.
12. R. L. Lamberts, *Appl. Opt.* **2**: 273 (1963).

Reprinted from APPLIED OPTICS, Vol. 3, page 1171, October 1964

The Influence of Image Motion and Shutter Operation on the Photographic Transfer Function

Roland V. Shack

In Part I a general two-dimensional expression for the transfer function of the photographic exposure image is derived, including as factors the optical transfer function, the transfer function arising from scattering of light in the emulsion, the shutter function, and the motion of the image. The effect of shutter operation in the absence of image motion is discussed and it is shown that a focal-plane shutter acts as an independent agent with its own effective transfer function, whereas a between-the-lens shutter is not independent but acts as an apodizing agent. In Part II, the effect of image motion is discussed, with the assumption that the optical transfer function is independent of the shutter action. The concept of an equivalent spread function for image motion is developed, and uniform linear image motion and simple harmonic image motion are discussed in detail, the latter especially for the case where there are at the most only a few periods in the exposure time. An approximation for small degradations is then obtained in which all possible combinations of uniform linear motion and simple harmonic motion are contained. Throughout the paper the functions discussed are functions of two-dimensional space.

Part I

Introduction

The spatial distribution of the effective exposure obtained in creating a photographic image is not, in general, identical to that of the instantaneous optical image being recorded. Apart from the blurring resulting from the diffusion of light in the emulsion, effectively treated elsewhere,[1] there are, in general, time-dependent variations in the optical image taking place during the exposure time. Two classes of phenomena which are considered in the present paper are: (1) the modification of the image structure resulting from the action of the shutter, which causes the shape and size of the pupil to vary as it opens and closes, and (2) motion of the image.

For the purposes of this paper we do not wish to concern ourselves with the photographic process as a whole but only with the formation of the latent image and not even with all of that. The latent image consists of a distribution of discrete grains of silver halide made developable by the action of light. This discrete structure can be related to the practically continuous structure of the optical image by postulating a continuous statistical latent image, the value of which at any point represents the probability that a developable silver halide grain will occur at that point after exposure. This probability will depend on a variety of factors, for example, the distribution of grain sensitivities and the population density of the grains, but the major factor which determines the spatial structure of the image is the amount of light received and integrated over the exposure time.

Thus the statistical latent image is related to the optical image producing it by the process of time integration. The time-integrated optical image is itself a kind of hypothetical image, which is in units of energy distribution rather than power distribution over the image surface. It is commonly called the exposure image.

The exposure image is an intermediate stage between the optical image and the statistical latent image. The statistical latent image can be obtained from the exposure image through a nonlinear transfer, taking into account the grain distribution factors, etc. However, if reciprocity can be assumed to hold, the exposure image can be obtained from the optical image by a linear process of integration. This means that the formation of the exposure image is the logical termination of the linear input stage of the photographic process, after which nonlinear procedures must be employed. Throughout the linear stage, Fourier methods can be applied and the image-forming capability of that part of the photographic system can be characterized by its transfer function, or, alternatively, its spread function.

The author was at Imperial College, London, England when this work was done. He is now with the Steward Observatory, University of Arizona, Tucson, Ariz.

Received 26 December 1963. Revised manuscript received 19 May 1964.

This work forms part of a thesis in preparation for a Ph.D. degree at the University of London.

Throughout the paper the transfer functions and spread functions discussed are functions of two spatial dimensions. To avoid unnecessary complication in the symbology, the same spatial coordinates u, v are used to represent object and image space. It is understood that the necessary, and simple, geometrical transformation has taken place. Also, to avoid using a great many different symbols, the author has chosen to use the tilde over those functions and variables which are in the spatial frequency domain, but maintaining the symbol which is used in the spatial domain. Thus u and v represent spatial coordinates whereas $\tilde{u}$ and $\tilde{v}$ represent the corresponding spatial frequency coordinates, and $\tilde{F}$ is the Fourier transform of F. The tilde, which is seldom used in optical symbology, suggests a sine wave.

To simplify the appearance of several of the mathematical expressions, $\mathfrak{S}(x)$ will be used to represent the frequently occurring $(\sin x)/x$.

In the following section a general expression for the transfer function of the exposure image is derived. Following that is a discussion of the possible effects of a focal-plane and a between-the-lens shutter on the transfer function in the absence of image motion. Part II will discuss the effect of image motion. The concept of the image-motion spread function is developed and uniform linear and simple harmonic motion are discussed in detail, the latter especially for exposure times which are of the same order of magnitude as the period of vibration. An approximation appropriate for small degradations is obtained, and combined motions are considered.

The Transfer Function of the Exposure Image

In this section we shall derive a general expression for the transfer function of the exposure image. In order to do this we must examine the process of formation of the exposure image, starting with the object.

The Fourier Description of the Object

Let $N(u,v)$ represent the radiance distribution over the object plane, limited to a finite field of area A. It will be advantageous to factor out the mean value: thus

$$N(u,v) = \bar{N}n(u,v), \tag{1}$$

where

$$\bar{N} = A^{-1}\iint_A N(u,v)du\,dv = A^{-1}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} N(u,v)du\,dv. \tag{2}$$

The total radiant intensity of the object is given by $A\bar{N}$. The purpose of the factoring is the separation of the object function into two factors, one without spatial structure but carrying the dimensional units, and the other providing the structural information in a dimensionless form.

The spatial Fourier transform of the object, which we shall call its structure spectrum, is given by

$$\tilde{N}(\tilde{u},\tilde{v}) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} N(u,v)\,\exp[-2\pi i(\tilde{u}u + \tilde{v}v)]du\,dv. \tag{3}$$

Here, too, we wish to separate the function into a dimensional factor and a structural factor. In order to do this, we define the normalized structure spectrum by

$$\begin{aligned}\tilde{n}(\tilde{u},\tilde{v}) &\equiv \frac{\tilde{N}(\tilde{u},\tilde{v})}{\tilde{N}(0,0)}\\ &= \frac{\bar{N}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} n(u,v)\exp[-2\pi i(\tilde{u}u + \tilde{v}v)]du\,dv}{\bar{N}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} n(u,v)du\,dv}\\ &= A^{-1}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} n(u,v)\exp[-2\pi i(\tilde{u}u + \tilde{v}v)]du\,dv,\end{aligned} \tag{4}$$

which is seen to be the normalized Fourier transform of the structural factor of the object distribution. The dimensional structure spectrum of Eq. (3) is then given by

$$\tilde{N}(\tilde{u},\tilde{v}) = A\bar{N}\tilde{n}(\tilde{u},\tilde{v}), \tag{5}$$

from which it is seen that the dimensional factor of the structure spectrum is simply the total radiant intensity of the object.

The Fourier Description of the Optical Image

In the same fashion, assuming the effective size of the spread function is much smaller than A, the irradiance in the image plane is given by

$$H(u,v) = \bar{H}h(u,v), \tag{6}$$

and the image structure spectrum by

$$\tilde{H}(\tilde{u},\tilde{v}) = A\bar{H}\tilde{h}(\tilde{u},\tilde{v}). \tag{7}$$

The average irradiance in the image is related to the average radiance of the object by

$$\bar{H} = K\bar{N}, \tag{8}$$

where K represents the usual photometric factor, and the normalized image structure spectrum is related to the corresponding object spectrum by

$$\tilde{h}(\tilde{u},\tilde{v}) = \tilde{\Phi}_O(\tilde{u},\tilde{v})\tilde{n}(\tilde{u},\tilde{v}), \tag{9}$$

where $\tilde{\Phi}_O$ is the transfer function of the optical system.

The Fourier Description of the Exposure Image

The exposure image is obtained by a time integration of the optical image in the emulsion. Thus,

$$\begin{aligned}E(u,v) &= \int_{-\infty}^{\infty} S(t)H(u,v)dt\\ &= \int_{-\infty}^{\infty} S(t)\left\{\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\tilde{H}\exp[2\pi i(\tilde{u}u + \tilde{v}v)]d\tilde{u}\,d\tilde{v}\right\}dt,\end{aligned} \tag{10}$$

where $S(t)$ is the shutter function in the usual sense. If we assume that the image is moving,

$$H(u,v;t) = H[u - \underline{u}(t), v - \underline{v}(t)], \tag{11}$$

where $\underline{u}(t)$ and $\underline{v}(t)$ are parametric expressions describing the path of motion. Then, on interchanging the order of integration, Eq. (10) becomes

$$E(u,v) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\tilde{E}(\tilde{u},\tilde{v})\exp[2\pi i(\tilde{u}u + \tilde{v}v)]d\tilde{u}\,d\tilde{v}, \tag{12}$$

where the structure spectrum of the exposure image is

$$\tilde{E}(\tilde{u},\tilde{v}) = \int_{-\infty}^{\infty} S\tilde{H} \exp[-2\pi i(\tilde{u}\underline{u} + \tilde{v}\underline{v})]dt. \quad (13)$$

Factoring $\tilde{H}$ gives us

$$\tilde{E}(\tilde{u},\tilde{v}) = AK\bar{N}t_e\left\{t_e^{-1}\int_{-\infty}^{\infty} S\tilde{\Phi}_O \exp[-2\pi i(\tilde{u}\underline{u} + \tilde{v}\underline{v})]dt\right\}\tilde{n}(\tilde{u},\tilde{v}), \quad (14)$$

where $t_e = \int_{-\infty}^{\infty} S dt$ is the effective exposure time. Comparing Eq. (14) with Eqs. (5) and (7) shows that $\tilde{E}$ can be factored in the same way as $\tilde{N}$ and $\tilde{H}$; that is,

$$\tilde{E}(\tilde{u},\tilde{v}) = A\bar{E}\tilde{e}(\tilde{u},\tilde{v}), \quad (15)$$

where

$$\bar{E} = K\bar{N}t_e, \quad (16)$$

which is the usual sensitometric expression, and

$$\tilde{e}(\tilde{u},\tilde{v}) = \tilde{\Phi}_E(\tilde{u},\tilde{v})\tilde{n}(\tilde{u},\tilde{v}), \quad (17)$$

where

$$\tilde{\Phi}_E(\tilde{u},\tilde{v}) = t_e^{-1}\int_{-\infty}^{\infty} S\tilde{\Phi}_O \exp[-2\pi i(\tilde{u}\underline{u} + \tilde{v}\underline{v})]dt \quad (18)$$

is the effective transfer function for the exposure image.

The optical transfer function $\tilde{\Phi}_O$ in the integrand of Eq. (18) is the optical transfer function in the emulsion, i.e., it includes the effect of light scatter in the emulsion. Thus

$$\tilde{\Phi}_O = \tilde{\Phi}_L\tilde{\Phi}_F, \quad (19)$$

where $\tilde{\Phi}_L$ is the transfer function for the optical system *per se*, and $\tilde{\Phi}_F$ is the transfer function for the scattering in the emulsion. The latter is independent of time and can be taken outside the integral. The former, as we shall see in the next section, is not, in general, independent of time and must be kept inside. The exposure-image transfer function then becomes

$$\tilde{\Phi}_E = \tilde{\Phi}_F t_e^{-1}\int_{-\infty}^{\infty} S\tilde{\Phi}_L \exp[-2\pi i(\tilde{u}\underline{u} + \tilde{v}\underline{v})]dt. \quad (20)$$

The Effect of the Shutter

In this section we shall assume the absence of image motion. The transfer function for the exposure image then becomes

$$\tilde{\Phi}_E = \tilde{\Phi}_F t_e^{-1}\int_{-\infty}^{\infty} S\tilde{\Phi}_L dt. \quad (21)$$

If we can also assume that $\tilde{\Phi}_L$ is independent of time, then

$$\tilde{\Phi}_E = \tilde{\Phi}_L\tilde{\Phi}_F t_e^{-1}\int_{-\infty}^{\infty} S dt = \tilde{\Phi}_L\tilde{\Phi}_F, \quad (22)$$

and the shutter has no effect.

Strictly speaking, this condition holds only when the opening and closing times of the shutter are negligibly small with respect to the effective exposure time, or when the shutter action is obtained by varying the transparency of the pupil uniformly, as with an ideal polarizing shutter. In practice, however, most shutters are mechanical devices which vary the shape of the pupil during their opening and closing times, thereby altering the pupil function and the optical transfer function $\tilde{\Phi}_L$ which depends on it.

The two most common types of shutter are the focal-plane and the between-the-lens shutter. We shall consider both in our investigation.

The Focal-Plane Shutter

The influence of a narrow focal-plane shutter on the image has been investigated by Bechtel[2] and by Asakura[3] by calculating the effect on the spread function of diffraction by the slit. Approaching the problem from the point of view of transfer function theory, however, leads to a very simple and elegant solution, as follows.

For mechanical reasons the so-called focal-plane shutter is never actually in the focal plane, but is a considerable distance forward of it. We shall assume for our purposes that the shutter is sufficiently removed from the focal plane for its effect to be virtually identical to that of an equivalent shutter in the pupil. Then the shutter function can be incorporated as a time-dependent modifier of the pupil function from which the optical transfer function can be calculated. Thus

$$\tilde{\Phi}_L = \tilde{\Phi}_L(\tilde{u},\tilde{v};t), \qquad \tilde{\Phi}_L(0,0;t) = S(t),$$

and

$$\tilde{\Phi}_E = \tilde{\Phi}_F t_e^{-1}\int_{-\infty}^{\infty} \tilde{\Phi}_L dt. \quad (23)$$

The modified optical transfer function can be expressed as the autocorrelation integral of the modified pupil function.

If the unmodified pupil function is represented by $f(x,y)$ and the pupil shutter function by $s(x + \dot{x}t)$ where the shutter velocity $\dot{x}$ is constant and the shutter is moving in the negative direction of x, then the transfer function for the exposure image is given by

$$\begin{aligned}\tilde{\Phi}_E &= \tilde{\Phi}_F t_e^{-1}\int_{-\infty}^{\infty}\left\{A^{-1}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}[s(x + \dot{x}t + {}^1/_2\tilde{u})\right.\\ &\quad f(x + {}^1/_2\tilde{u}, y + {}^1/_2\tilde{v})][S^*(x + \dot{x}t - {}^1/_2\tilde{u})\\ &\quad \left. f^*(x - {}^1/_2\tilde{u}, y - {}^1/_2\tilde{v})]dx\,dy\right\}dt\\ &= \tilde{\Phi}_F A^{-1}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\left[t_e^{-1}\int_{-\infty}^{\infty} s(x + \dot{x}t + {}^1/_2\tilde{u})\right.\\ &\quad \left. s^*(x + \dot{x}t - {}^1/_2\tilde{u})dt\right] f(x + {}^1/_2\tilde{u}, y + {}^1/_2\tilde{v})\\ &\quad f^*(x - {}^1/_2\tilde{u}, y - {}^1/_2\tilde{v})dx\,dy, \quad (24)\end{aligned}$$

where A represents the area of the pupil and the asterisks indicate complex conjugates. The time integral in the brackets is in the form of an autocorrelation integral in which both components have been shifted by an amount x, but equal shifts in the same direction do not alter the value of the integral which is therefore independent of x. The integral is also independent of y and therefore may be taken outside the larger integral:

$$\tilde{\Phi}_E = \tilde{\Phi}_F\left[t_e^{-1}\int_{-\infty}^{\infty} o(\dot{x}t + {}^1/_2\tilde{u})s^*(\dot{x}t - {}^1/_2\tilde{u})dt\right]$$
$$\left[A^{-1}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(x + {}^1/_2\tilde{u}, y + {}^1/_2\tilde{v})f^*(x - {}^1/_2\tilde{u}, y - {}^1/_2\tilde{v})dx\,dy\right]. \quad (25)$$

The expression within the second square bracket is simply the transfer function for the optical system with the shutter fully open, so the expression within the first square bracket must describe the equivalent transfer function for the focal plane shutter, and

$$\tilde{\Phi}_E = \tilde{\Phi}_L\tilde{\Phi}_F\tilde{\Phi}_S. \quad (26)$$

By changing the variable of integration, letting $x' = \dot{x}t$, the form of the integral for the shutter transfer function is given by

$$\tilde{\Phi}_S = \frac{\int_{-\infty}^{\infty} s(x' + {}^1/_2\tilde{u})s^*(x' - {}^1/_2\tilde{u})dx'}{\int_{-\infty}^{\infty} s(x')s^*(x')dx'}, \quad (27)$$

which is the form of a normalized spatial autocorrelation function strictly analogous to the expression for the optical transfer function. It is, however, a function of $\tilde{u}$ alone, being independent of $\tilde{v}$.

It is interesting to note the rather unexpected result that the equivalent shutter transfer function for a focal plane shutter is quite independent of the shape and size of the pupil and the aberrations of the optical system. Unfortunately, this independence only holds in the absence of image motion.

The derivation assumes the most general complex form for the pupil shutter function, subject to the restrictions that it be independent of y, convergent in x, and constant in its velocity. However, the most common form of shutter is a simple slit. If we let w represent the ratio of the shutter width projected into the pupil to the radius of the pupil aperture, then, for a simple slit shutter

$$\tilde{\Phi}_S = 1 - \frac{|\tilde{u}|}{w}, \quad 0 < |\tilde{u}| < w$$
$$= 0, \quad |\tilde{u}| > w. \quad (28)$$

This is a roof-shaped function of triangular cross section, as shown in Fig. 1.

The Between-the-Lens Shutter

The between-the-lens shutter was also investigated by Bechtel by calculating the spread function. The following illustrates the transfer function approach.

Unfortunately, the between-the-lens shutter which operates approximately radially in the pupil does not lend itself to the elegant reduction of the previous section. It is not possible in general to separate the shutter effect from the aberrations or the diaphragm setting, and the problem is usually complicated by the shape of the shutter aperture, which usually resembles a bent star or pinwheel.

In general Eq. (21) must be used where the variation of $\tilde{\Phi}_L$ with time must be determined beforehand. The shutter function S is given by the ratio of the area of the clear aperture at time to the area of the aperture when the shutter is fully open. Both $\tilde{\Phi}_L$ and S will depend on the diaphragm setting as well.

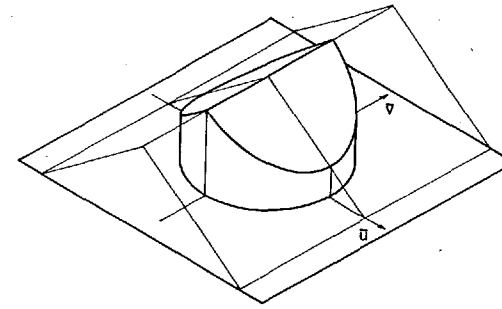


Fig. 1. Transfer function for focal-plane shutter. The cylinder in the figure represents the boundary of the frequency domain limited by optical diffraction.

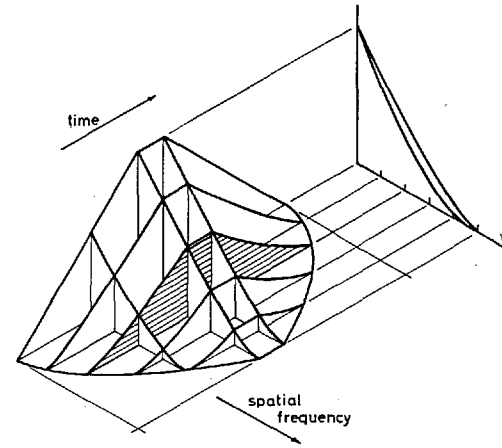


Fig. 2. Influence of a between-the-lens shutter on the transfer function. This figure is a graphical interpretation of Eq. (21). The optical system is assumed to be free of aberrations.

In order to get a general idea of the effect of a between-the-lens shutter, a simplified model consisting of a circular aperture, the radius of which varies with time, will be examined. It will be assumed that its area increases linearly during its opening and that it closes linearly at the same rate.

Figure 2 is a graphical interpretation of Eq. (21) for this simplified model in the absence of aberrations. Note that as the spatial frequency of a component increases there is a corresponding delay before a contribution to that component can pass through the spatial filter, and a corresponding reduction in the effective transfer function. This occurs because the smaller aperture of the partly open shutter results in a lower maximum spatial frequency which can be transferred.

Figure 3 is a corresponding illustration for the same model with a defect of focus of ${}^1/_2\lambda$. Here the same spatial frequency limit applies as in Fig. 2, but the amplitudes are reduced. The amount of reduction, however, increases rapidly with the size of the shutter aperture, so that the amplitude with the shutter partly open may exceed that obtained with the shutter fully

open, and the resulting transfer function value may be greater than that obtained with a "perfect" shutter which opens and closes instantaneously.

Figure 4 shows for the same model the difference between the two extreme shutter functions, the rectangular "perfect" shutter function and the triangular "50% efficient" shutter function. In Fig. 4(a) we see the difference in shape of the transfer functions and in Fig. 4(b) the manner in which the transfer function values vary with focus for a particular frequency.

Taking the defect of focus as an aberration, it is clear that, although a finite shutter effect degrades an aberration-free system, it can improve an aberrated system. The improvement should be even more striking for higher order aberrations.

The character of the curves in Fig. 4 implies that the ripples of the corresponding spread functions must be considerably reduced for the triangular shutter function relative to those of the rectangular shutter function; that is, a kind of apodization has taken place. In fact, the simplified model of shutter we employed does result in noncoherent apodization for the exposure image, and the effect of a between-the-lens shutter can be thought of in terms of such an apodization. The manner and extent of the apodization will depend heavily on the shape of the actual shutter aperture, however, and the concept of apodization in the usual sense only applies in the absence of image motion.

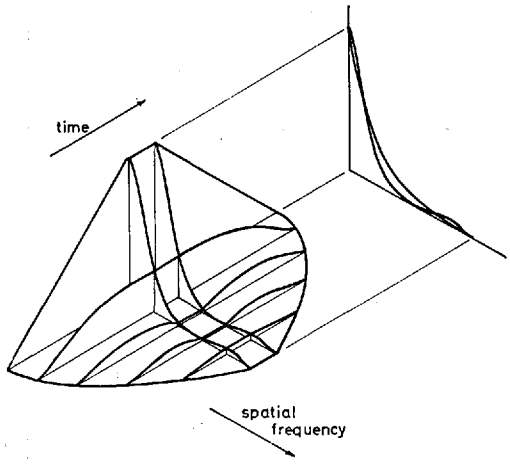
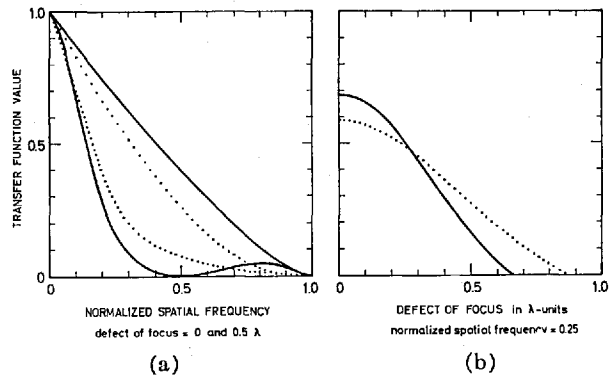


Fig. 3. Same as Fig. 2 except that the optical system is defocused by $\frac{1}{2}\lambda$.



(a) (b)

Fig. 4. Influence of a between-the-lens shutter on the transfer function. The solid curves are for the rectangular shutter function and the dotted curves for the trianglar shutter function.

Part II

Introduction

Earlier investigations of the effect of image motion on the quality of the image have mostly been concerned with predicting the reduction in resolving power resulting from image motion. Scott[4] reported expressions for simple forms of one-dimensional image motion. Rosenau[5] reported an expression for parabolic image motion. Chang[6] has investigated the effect of simple harmonic motion on resolving power when the exposure time is a fraction of a cycle of the motion. Paris[7] has attempted to determine the two-dimensional transfer function for combined linear and simple harmonic image motion. Paris,[7] and Hendeberg and Welander[8] have considered the modification of the transfer function for linear motion by a low-efficiency shutter, neglecting the effects described in Part I of this paper.

In Part II we shall also neglect the effects described in Part I, i.e., we shall assume that the optical transfer function is independent of time. As should be clear from Part I, this assumption is strictly valid only in comparatively rare cases. However, it is a reasonably good approximation if the effective exposure time is appreciably greater than the time it takes for the shutter to open and close, or if we are restricted to low spatial frequencies and small aberrations.

The Transfer Function and Spread Function for Image Motion

Under this assumption the expression for the transfer function for the exposure image becomes

$$\tilde{\Phi}_E = \tilde{\Phi}_L \tilde{\Phi}_F \tilde{\Phi}_M, \tag{29}$$

where the effective transfer function for image motion is

$$\tilde{\Phi}_M(\tilde{u},\tilde{v}) = t_e^{-1} \int_{-\infty}^{\infty} S \exp[-2\pi i(\tilde{u}\underline{u} + \tilde{v}\underline{v})]dt. \tag{30}$$

Remembering that $\underline{u}$ and $\underline{v}$ are parametric expressions describing the path of image motion, we shall henceforth omit the bar.

In order to evaluate this integral at some spatial frequency $(\tilde{u},\tilde{v})$ we can simplify our problem by rotating the reference axes through an angle ψ so that the new abscissa $\tilde{u}'$ is in the direction of $(\tilde{u},\tilde{v})$. Then

$$\begin{aligned} \tilde{u}' &= \tilde{u}\cos\psi + \tilde{v}\sin\psi, \qquad \tilde{v}' = 0 \\ u' &= u\cos\psi + v\sin\psi, \end{aligned} \tag{31}$$

and

$$\tilde{\Phi}_M(\tilde{u}',\psi) = t_e^{-1} \int_{-\infty}^{\infty} S \exp(-2\pi i\tilde{u}'u')dt. \tag{32}$$

If we change the variable of integration to u',

$$\tilde{\Phi}_M = \frac{\int_{-\infty}^{\infty} [S/\dot{u}'] \exp(-2\pi i \tilde{u}' u') du'}{\int_{-\infty}^{\infty} [S/\dot{u}'] du'}, \tag{33}$$

which has the form of a normalized Fourier transform, and, consequently, the quantity in the square brackets represents an equivalent image-motion spread function. However, the form of this quantity depends on orientation, suggesting that it corresponds to a line-spread function rather than a point-spread function.

An equivalent point-spread function does exist, but it differs radically in character from the more familiar optical or emulsion spread functions. The latter functions can be represented by continuous, approximately bell-shaped surfaces overlaying the u,v plane, whereas the former is represented by a ribbonlike cylinder, the directrix of which is the path traced by the parametric variables $u(t)$ and $v(t)$ and the generatrix of which is a line segment perpendicular to and extending up from the u,v plane and having a length given by $S|(\dot{u}^2 + \dot{v}^2)^{-1/2}|$, that is, the shutter function divided by the magnitude of the velocity of the image motion along the path (see Fig. 5).

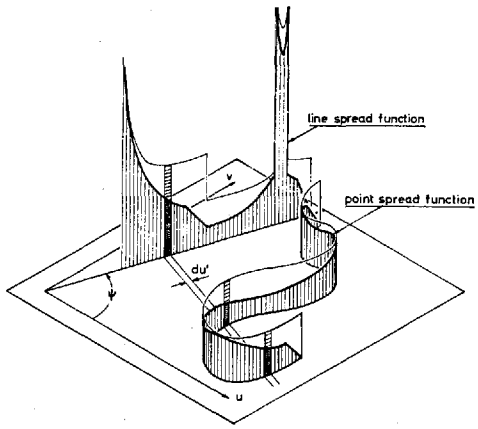


Fig. 5. Image motion spread functions.

The line-spread function in the ψ direction is found by integrating this point-spread function in the perpendicular direction. That is, an element of area in the line-spread function of width du' is equal to the area of that portion of the ribbon representing the point-spread function intercepted by the interval du', this area being given by

$$dA = \frac{du'}{\cos(\psi - \phi)} \cdot \frac{S}{(\dot{u}^2 + \dot{v}^2)^{1/2}}, \tag{34}$$

where ϕ is the angle between the u axis and the tangent to the curve of the directrix. But

$$\begin{aligned} \dot{u}' &= \dot{u}\cos\psi + \dot{v}\sin\psi \\ &= (\dot{u}^2 + \dot{v}^2)^{1/2}\cos(\psi - \phi), \end{aligned} \tag{35}$$

so the element of area of the line-spread function is represented by

$$dA = [S/\dot{u}']du', \tag{36}$$

in agreement with Eq. (33).

It should be noted that the directrix may be a multivalued function of u', in which case the line-spread function is the sum of all the intercepted elements (see Fig. 5), and has an infinite discontinuity at each value of u' corresponding to a bend and possibly a finite discontinuity at each value corresponding to a terminus. But the integral of the line-spread function exists (it is equal to the total area of the ribbon representing the point-spread function, which in fact is equal to the effective exposure time t_e) and there can be only two termini and a finite number of bends, so the validity of the Fourier transform expressed in Eq. (33) is not jeopardized.

We shall now use Eqs. (32) and (33) to investigate a few special cases of image motion which are of practical importance.

Uniform Linear Motion

If $\dot{u}$ and $\dot{v}$ are constant with respect to time, then we have uniform linear motion having a velocity of magnitude

$$\dot{l} = (\dot{u}_l^2 + \dot{v}_l^2)^{1/2}, \tag{37}$$

in a direction

$$\phi_l = \arctan(\dot{v}_l/\dot{u}_l). \tag{38}$$

In Eq. (33)

$$\dot{u}_l = \dot{l}\cos(\psi - \phi_l) \tag{39}$$

is a constant with respect to u', so $\tilde{\Phi}_M$ becomes simply the normalized Fourier transform of the shutter function. If S is assumed to be unity during the exposure, then

$$\begin{aligned} \tilde{\Phi}_M &= \mathfrak{S}(\pi\tilde{u}'\dot{u}_l' t_e) \\ &= \mathfrak{S}[\pi\tilde{u}' l\cos(\psi - \phi_l)], \end{aligned} \tag{40}$$

where $l = \dot{l}t_e$ is the distance the image travels in the exposure time. The corresponding point-spread function is a rectangle of length l and height $\dot{l}^{-1}$. The two functions are shown in Fig. 6.

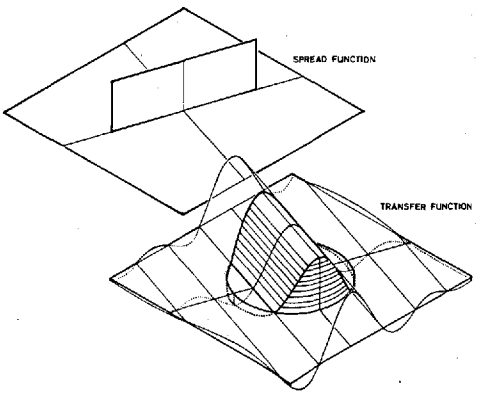


Fig. 6. Spread function and transfer function for uniform linear motion and a rectangular shutter function. The cylinder represents the boundary of the frequency domain limited by optical diffraction.

If the shutter function is a trapezoid with a rise time t_r, then

$$\tilde{\Phi}_M = \mathfrak{S}(\pi\tilde{u}'\hat{u}_i't_e)\mathfrak{S}(\pi\tilde{u}'\hat{u}_i't_r), \tag{41}$$

because a trapezoid is the convolution of two rectangles. If the shutter function is S-shaped during the rise time, that is, the convolution of a rectangle and a circle (focal-plane shutter),

$$\tilde{\Phi}_M = \mathfrak{S}(\pi\tilde{u}'\hat{u}_i't_e)\left[\frac{2J_1(\pi\tilde{u}'\hat{u}_i't_r)}{(\pi\tilde{u}'\hat{u}_i't_r)}\right]. \tag{42}$$

Thus for uniform linear motion, the image-motion transfer function can be factored into a term dependent only on the effective exposure time regardless of the so-called shutter efficiency, and a term dependent only on the rise time. The first is the Fourier transform of a rectangular function and the second the Fourier transform of the derivative of the shutter function during the rise time.

This result is valid only if the optical transfer function $\tilde{\Phi}_L$ is independent of time and no other form of image motion is present.

Simple Harmonic Motion

Next to uniform linear motion the most important elementary kind of image motion is that resulting from vibration in the camera system. In its simplest form this is simple harmonic motion, described by

$$u = u_a \sin(2\pi\nu t + \alpha), \qquad v = v_a \sin(2\pi\nu t + \alpha), \tag{43}$$

where u_a and v_a are the component amplitudes of the motion (assumed constant), ν is the frequency of the vibration in cycles per second, and α is a phase term defining the position of the image in its path at time $t = 0$. The path of the image motion is a straight line making an angle

$$\phi_a = \arctan(v_a/u_a) \tag{44}$$

with the u axis, and the amplitude is

$$a = (u_a^2 + v_a^2)^{1/2}. \tag{45}$$

On rotation of coordinates we obtain

$$u' = u_a' \sin(2\pi\nu t + \alpha), \tag{46}$$

where

$$u_a' = a\cos(\psi - \phi_a). \tag{47}$$

Substituting into Eq. (32) we obtain

$$\tilde{\Phi}_M = t_e^{-1}\int_{-\infty}^{\infty} S\exp[-2\pi i\tilde{u}'u_a'\sin(2\pi\nu t + \alpha)]dt. \tag{48}$$

It is convenient to normalize the variable by $\hat{t} = t/t_e$; thus

$$\tilde{\Phi}_M = \int_{-\infty}^{\infty} S\exp[-2\pi i\tilde{u}'u_a'\sin(2\pi m\hat{t} + \alpha)]d\hat{t}, \tag{49}$$

where $m = \nu t_e$ is the number of cycles of vibration occurring in the effective exposure time. If now we assume that the shutter function is constant during the exposure time (rectangular shutter function), then we obtain

$$\tilde{\Phi}_M = \int_{-1/2}^{1/2} \exp[-2\pi i\tilde{u}'u_a'\sin(2\pi m\hat{t} + \alpha)]d\hat{t}, \tag{50}$$

where the origin of $\hat{t}$ is taken in the center of the exposure time.

It can be shown that, for integral values of m, Eq. (50) reduces to the Bessel function

$$\tilde{\Phi}_M = J_0(2\pi\tilde{u}'u_a'), \tag{51}$$

which is independent both of α and of the order of m. Also, for continuous values of m,

$$\lim_{m\to\infty}\tilde{\Phi}_M = J_0(2\pi\tilde{u}'u_a'). \tag{52}$$

The latter expression, given by Scott,[4] is of limited practical value, however, because in most cases not more than a few cycles of vibration will occur in the exposure time, and in this region the general dependence on m and on α (an uncontrolled parameter) may make $\tilde{\Phi}_M$ in any particular instance deviate considerably from that predicted by Eqs. (51) or (52).

Figure 7 shows the manner in which $\tilde{\Phi}_M$ varies with α and with m for m ranging from 0 to 2. In this range, at least, it is clear that $\tilde{\Phi}_M$ depends strongly on both α and m. Figure 8 shows this dependence more clearly, at least for the main lobe of $\tilde{\Phi}_M$.

Several features are apparent. The general trend is for $\tilde{\Phi}_M$ to approach the limiting value specified by Eq. (52), but, for small values of m, $\tilde{\Phi}_M$ rises as m decreases, approaching unity as m approaches zero. The reason

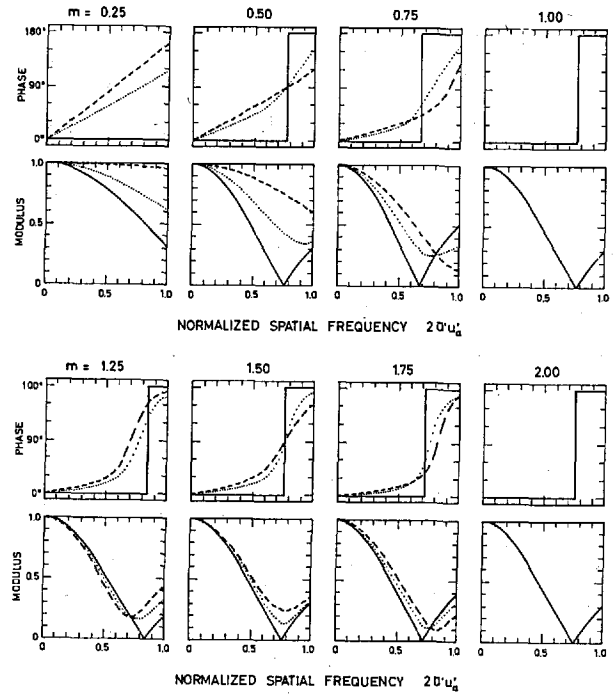


Fig. 7. Transfer functions for simple harmonic image motion. The values of m indicate the number of cycles in the effective exposure time. The shutter function is rectangular. The solid curve is for $\alpha = 0$, the dotted curve for $\alpha = {}^1/_4\pi$, and the dashed curve for $\alpha = {}^1/_2\pi$.

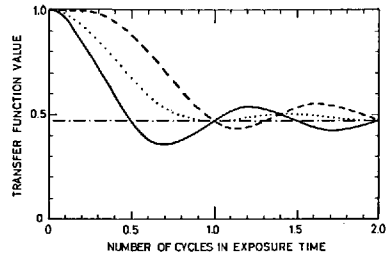


Fig. 8. Variation of the simple harmonic image motion transfer function with m and α, for a normalized spatial frequency value of 0.5. The shutter function is rectangular. The solid curve is for $\alpha = 0$, the dotted curve for $\alpha = {}^1/_{4\pi}$, and the dashed curve for $\alpha = {}^1/_{2\pi}$.

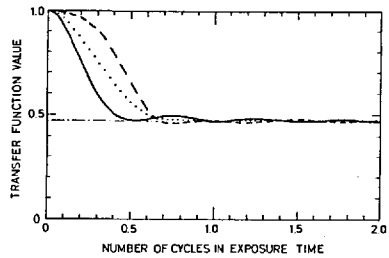


Fig. 9. Same as Fig. 8 except that the shutter function is triangular. The solid curve is for $\alpha = 0$, the dotted curve for $\alpha = {}^1/_{4\pi}$, and the dashed curve is for $\alpha = {}^1/_{2\pi}$.

for the rise is, of course, the fact that, if the exposure time is less than one period, the distance the image travels during the exposure diminishes with m, and the blur is less. At integral values of m we have the nodes predicted by Eq. (51). Between $m = 1$ and $m = 2$ is another node, of which more will be said later. Between the nodes, $\tilde{\Phi}_M$ can take on a range of values dependent on α.

The dependence on α is quite important. Although the amplitude and frequency of the vibration might conceivably be specified in advance, the reference phase α is unpredictable in any given case, all values of α being equally likely. Thus the best that can be done in describing or predicting $\tilde{\Phi}_M$ is to specify its probability distribution. In the present case this distribution has an upper and a lower limit, and it is more probable that a given value of $\tilde{\Phi}_M$ will be near one or the other of these limits than that it will lie midway between.

If instead of a rectangular shutter function, a triangular one is assumed, the same nodes occur at integral values of m [Eq. (51)], and the same limiting function applies [Eq. (52)], but the function converges much more rapidly to the limiting value. Figure 9 shows the result obtained for the same parameters as Fig. 8. It can be seen that, except for values of m less than about 0.7, the value specified by the limiting function is quite a reasonable prediction for $\tilde{\Phi}_M$ for any value of α, at least for the main lobe of the transfer function.

Combined Motions

In the general problem of combined motions,

$$u = u_1 + u_2 + \ldots, \qquad v = v_1 + v_2 + \ldots, \tag{53}$$

from which we obtain

$$u' = u_1' + u_2' + \ldots, \tag{54}$$

and

$$\dot{u}' = \dot{u}_1' + \dot{u}_2' + \ldots. \tag{55}$$

It should be noted that each pair of functions in (53) is associated with a particular set of parameters, but because the orientation in general differs from pair to pair, the distribution of parameters in Eqs. (54) and (55) will vary with ψ.

The equivalent line-spread function is given by

$$S/\dot{u}' = S/(\dot{u}_1' + \dot{u}_2' + \ldots), \tag{56}$$

from which it can be seen that the resultant spread function is not a linear combination, product, or convolution of the separate component spread functions, and so the equivalent image-motion transfer function cannot be a simple combination of the component transfer functions. It, of course, is given by

$$\tilde{\Phi}_M = t_e^{-1} \int_{-\infty}^{\infty} S \exp[-2\pi i \tilde{u}'(u_1' + u_2' + \ldots)]dt \tag{57}$$

The simplest case of combined motions is the sum of two or more uniform linear motions. Clearly the resultant is itself a uniform linear motion, and no further discussion is necessary.

The next most simple case of combined motions is that resulting from two or more vibrations of the same frequency. Then

$$u = u_a \sin(2\pi\nu t + \alpha), \qquad v = v_a \sin(2\pi\nu t + \alpha + \beta), \tag{58}$$

from which we obtain

$$u' = u_a' \sin(2\pi\nu t + \alpha'), \tag{59}$$

where

$$\alpha' = \alpha + \beta', \qquad \tan\beta' = \frac{(v_a/u_a)\sin\beta \tan\psi}{1 + (v_a/u_a)\cos\beta \tan\psi}, \tag{60}$$

and

$$u_a' = (u_a^2 \cos^2\psi + v_a^2 \sin^2\psi + 2u_a v_a \sin\psi \cos\psi \cos\beta)^{1/2}. \tag{61}$$

Equation (59) has the same form as Eq. (46), and thus the preceding analysis for a single source of vibration applies here as well for each direction, but in general the variation of $\tilde{\Phi}_M$ with ψ is different.

Of course, if $\beta = 0$, the two components of motion are in phase and the resultant motion is indistinguishable from that resulting from a single source of vibration, and the previous analysis applies exactly.

On the other hand, if $\beta = {}^1/_{2\pi}$ and $u_a = v_a = a$, the path of motion is a circle and

$$u' = a \sin(2\pi\nu t + \alpha + \psi), \tag{62}$$

where the amplitude is independent of orientation but the reference phase is a linear function of ψ.

Figure 10 illustrates these two extreme cases along with a general intermediate case for $m = 1$. The three different point-spread functions have identical line-

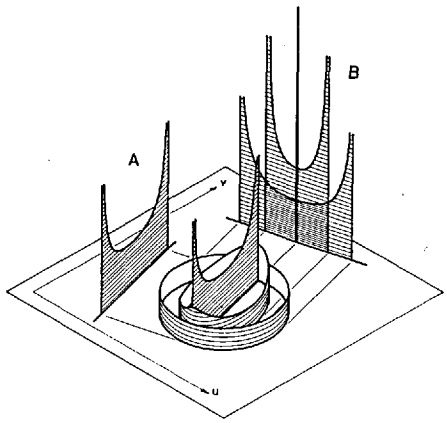


Fig. 10. Spread functions for combined simple harmonic motions of the same frequency. Exactly 1 cycle in the exposure time is assumed. The shutter function is rectangular.

spread functions in the direction of A but different ones in the direction of B.

The more general problems of the combination of linear and simple harmonic motions and combined vibratory motions of different frequencies do not have any corresponding simplicity. Equations (56) and (57) must be applied directly.

Approximation for Small Degradations

One of the principal uses of transfer function theory is in the establishment of performance tolerances, that is, the determination of the magnitude of a source of degradation when the degradation is assumed to be just tolerable. For this purpose we can use an approximation which is valid for small values of the total degradation, the approximation being the use of the first three terms of the power series expansion for the exponential expression in the integrand of Eq. (32).

First, however, let us establish as our coordinate reference for the spread function its center of gravity. Thus, before shifting our coordinates, we let

$$u' = \bar{u}' + u'', \tag{63}$$

and substitute in Eq. (32).

Shifting our coordinates by making $\bar{u}' = 0$, we obtain

$$\tilde{\Phi}_M = t_e^{-1}\int_{-\infty}^{\infty} S \exp(-2\pi i \tilde{u}' u'')dt. \tag{64}$$

The effect of shifting the coordinates is only the introduction of a phase term which is a linear function of the spatial frequency.

We now make our approximation, obtaining

$$\tilde{\Phi}_M \approx t_e^{-1}\int_{-\infty}^{\infty} S[1 - 2\pi i\tilde{u}'\bar{u}'' - 2\pi^2\tilde{u}'^2u''^2]dt$$

$$= 1 - 2\pi^2\tilde{u}'^2\left(t_e^{-1}\int_{-\infty}^{\infty} Su''^2\,dt\right). \tag{65}$$

This expression is free of imaginary terms, thus indicating that, to the extent that the approximation is valid, asymmetry of the spread function is of no importance, and any central section of the transfer function is a parabola. It is always pessimistic, actual values of the function being somewhat higher.

A transfer function which corresponds to no degradation of the image is simply equal to unity for all spatial frequencies. Thus the second term of Eq. (65) is conveniently considered as the loss L_M associated with the transfer function:

$$\tilde{\Phi}_M = 1 - L_M, \tag{66}$$

where

$$L_M = 2\pi\tilde{u}'^2\left(t_e^{-1}\int_{-\infty}^{\infty} Su''^2dt\right). \tag{67}$$

By changing the variable of integration to u', as in Eq. (33), Eq. (67) becomes

$$L_M = 2\pi^2\tilde{u}'^2\left[\int_{-\infty}^{\infty}(S/\dot{u}')u''^2du'\right]. \tag{68}$$

and it becomes apparent that the quantity in square brackets is the second moment of the spread function.

If we assume a rectangular shutter function and the image motion is uniform linear motion, that is, $u' = \dot{u}_l' t$ then $\bar{u}' = 0$ and

$$L_M = {}^1/_6\pi^2\tilde{u}'^2\dot{u}_l'^2. \tag{69}$$

If the image motion is simple harmonic, that is

$$u' = u_a' \sin(2\pi\nu t + \alpha),$$

then

$$\bar{u}' = u_a' \sin\alpha\mathfrak{S}(\pi m), \tag{70}$$

and

$$L_M = \pi^2\tilde{u}'^2u_a'^2\{[1 - \mathfrak{S}^2(\pi m)] + (\cos 2\alpha)\,[\mathfrak{S}^2(\pi m) - \mathfrak{S}(2\pi m)]\}. \tag{71}$$

The quantity inside the curly brackets expresses the dependence on m, the number of cycles of vibration in the effective exposure time, and α, the reference phase. It is plotted in Fig. 11, which on comparison with Fig. 8, shows that the dependence is quite adequately accounted for.

The quantity in the first set of square brackets gives the average dependence on m. The quantity in the second set of brackets defines the amplitude by which

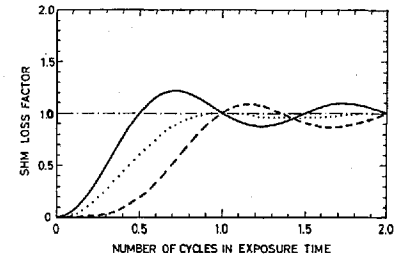


Fig. 11. Dependence of the transfer function loss factor on m and α [see Eq. (71)]. The solid curve is for $\alpha = 0$, the dotted curve for $\alpha = {}^1/_4\pi$, and the dashed curve for $\alpha = {}^1/_2\pi$. This figure should be compared with Fig. 8.

the loss fluctuates with α about the average value given by the first quantity.

The total expression in the curly brackets goes to zero for all α as m approaches zero and to unity as m approaches infinity. It also goes to unity for integral values of m. The intermediate node previously mentioned occurs when the quantity in the second set of square brackets is zero, that is, when $\tan\pi m = \pi m$. Other intermediate nodes will occur at larger values of m which satisfy this relationship.

If the image motion is the sum of uniform linear and simple harmonic motion, that is,

$$u' = \dot{u}_l' t + u_a \sin(2\pi\nu t + \alpha),$$

then

$$L_M = {}^1/_6\pi^2\tilde{u}'^2 u_l'^2 + \pi^2\tilde{u}'^2 u_a'^2\{[1 - \mathfrak{S}^2(\pi m)] + (\cos 2\alpha)[\mathfrak{S}^2(\pi m) - \mathfrak{S}(2\pi m)]\} + 2\pi^2\tilde{u}'^2 u_l' u_a'\{\cos\alpha[(\pi m)^{-1}(\mathfrak{S}(\pi m) - \cos\pi m)]\}. \quad (72)$$

Here it is apparent that the first term is the loss which would result from the linear motion alone, the second is the loss which would result from the simple harmonic motion alone, and the third the loss resulting from the interaction between the two.

The interaction loss also depends on m and α, as shown by the expression inside the curly brackets. This expression is shown in Fig. 12, where it can be seen from the nodes present that the interaction term vanishes when the quantity inside the square brackets is zero, that is, when $\tan\pi m = \pi m$, the same condition which applies to the intermediate nodes of the pure vibration dependence. Thus, at these values of m the image motion transfer function is independent of the reference phase α. The interaction term, however, does not, in general, vanish at integral values of m.

It is instructive to observe how the transfer function varies with m and α as the proportion between the linear component and the vibration component is varied, keeping the total loss constant (for $m \to \infty$). Figure 13 shows this for a total loss of 0.2 with the proportion ranging from pure linear motion to pure vibratory motion.

The most striking feature is the dominant effect of the interaction term for even small amounts of vibration added to the linear image motion. This of course reflects the fact that when the vibratory component travels in the same direction as the linear component the blur is increased, whereas when it travels in the opposite direction the blur is reduced or even cancelled.

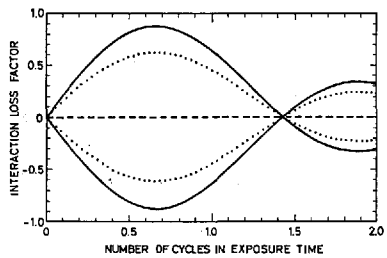


Fig. 12. Interaction loss factor for uniform linear image motion combined with simple harmonic motion [see Eq. (72)]. The factor 2 is incorporated in the figure. The solid curve is for $\alpha = 0$ or π, the dotted curve for $\alpha = {}^1/_4\pi$ or ${}^3/_4\pi$, and the dashed curve for $\alpha = {}^1/_2\pi$.

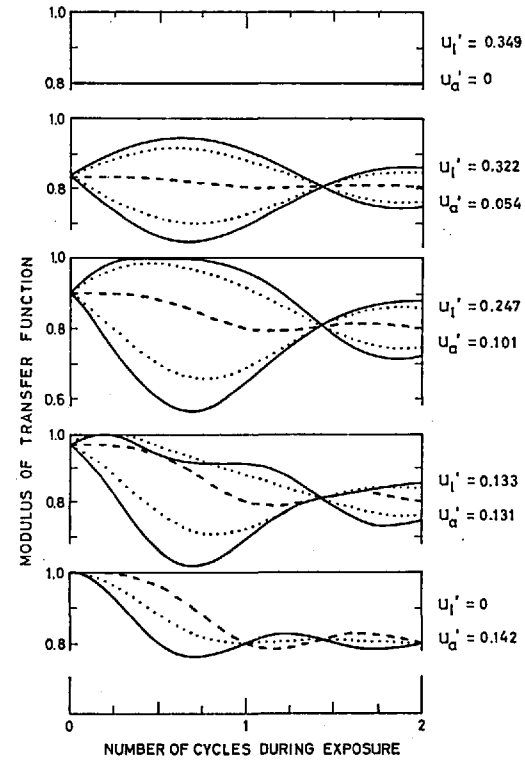


Fig. 13. Variation of the transfer function for combined uniform linear image motion and simple harmonic motion as the proportion of the two components is varied from pure linear motion to pure simple harmonic motion. The solid curve is for $\alpha = 0$ or π, the dotted curve for $\alpha = {}^1/_4\pi$ or ${}^3/_4\pi$, and the dashed curve for $\alpha = {}^1/_2\pi$.

As was stated previously, α is a parameter over which there is normally no control, all values of α being equally likely. Thus, even if all other parameters are specified, no specific value of the transfer function can be predicted unless the value of m corresponds to one of the nodes as seen in Fig. 13. Instead, the transfer function may take on a range of values which is best described by a probability distribution, and a tolerance level can be established on the basis of the statistical performance required. In the present case a fair approximation to the correct procedure can be made by using absolute values for the quantities in square brackets in Eq. (72) and choosing a value of α to conform to the statistical performance required. Thus, no photograph will be worse than the value obtained by setting $\alpha = 0$, and if the peak-to-peak amplitude of the vibration does not exceed the distance of linear motion, then 75% of the photographs taken will be better than the value obtained by setting $\alpha = {}^1/_4\pi$ and 50% for $\alpha = {}^1/_2\pi$.

If the image motion is the sum of two simple harmonic motions of differing frequency, that is,

$$u' = u_{a1}' \sin(2\pi\nu_1 t + \alpha_1) + u_{a2}' \sin(2\pi\nu_2 t + \alpha_2),$$

then

$$\begin{aligned} L_M = {} & \pi^2\tilde{u}'^2 u_{a1}'^2 \{[1 - \mathfrak{S}^2(\pi m_1)] \\ & + (\cos 2\alpha_1)[\mathfrak{S}^2(\pi m_1) - \mathfrak{S}(2\pi m_1)]\} + \\ & \pi^2\tilde{u}'^2 u_{a2}'^2 \{[1 - \mathfrak{S}^2(\pi m_2)] + \\ & (\cos 2\alpha_2)[\mathfrak{S}^2(\pi m_2) - \mathfrak{S}(2\pi m_2)]\} \\ & + 2\pi^2\tilde{u}'^2 \tilde{u}_{a1}' u_{a2}' \\ & \{\cos(\alpha_1 - \alpha_2)[\mathfrak{S}(\pi m_1 - \pi m_2) - \mathfrak{S}(\pi m_1)\mathfrak{S}(\pi m_2)] + \\ & \cos(\alpha_1 + \alpha_2)[\mathfrak{S}(\pi m_1)\mathfrak{S}(\pi m_2) - \mathfrak{S}(\pi m_1 + \pi m_2)]\}. \end{aligned} \quad (73)$$

Here again the first two terms are the losses which would result from each component of image motion by itself and the third term is the loss resulting from interaction between the two.

This example is more difficult to describe than the previous examples because we now have six parameters, two of which are independent random variables. In particular the probability distribution of L_M becomes much more complicated.

In order to get a general idea of the properties of Eq. (73) we must look more closely at the interaction term. This term itself contains all six parameters, but they are conveniently grouped. The expressions in the square brackets describe the dependence on both m_1 and m_2. Figure 14 shows the manner in which the quantity in the first square brackets varies and Fig. 15,

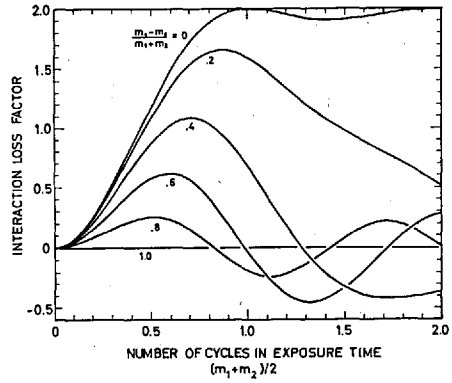


Fig. 14. Interaction loss factor for combined simple harmonic image motions of different frequencies. This is for the expression in the first square brackets of the interaction term in Eq. (73). The dependence on α_1 and α_2 is not shown.

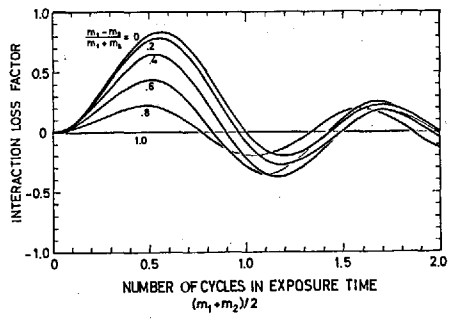


Fig. 15. Interaction loss factor for combined simple harmonic image motions of different frequencies. This is for the expression in the second square brackets of the interaction term in Eq. (73). The dependence on α_1 and α_2 is not shown.

the second. The parameters m_1 and m_2 have been transformed to $(m_1 - m_2)/(m_1 + m_2)$ and $\frac{1}{2}(m_1 + m_2)$ in the figures.

It is clear that the magnitudes of the expressions generally increase as the difference between the frequencies decreases. This means that the interaction between the two components reaches its maximum when the two frequencies become the same. If the two frequencies are the same, then the loss is a maximum when the motions are in phase with each other, that is, $\alpha_1 = \alpha_2 = \alpha$, in which case the loss can be expressed as in Eq. (71), where $u_a' = u_{a1}' + u_{a2}'$. For all other frequency and phase combinations the loss will be diminished. Thus, if a tolerance is to be established for a combination of two different frequencies, then a safe maximum limit for the loss can be set by assuming that both frequencies are equal to their average and that they are in phase with each other. This procedure is, of course, a short-cut and may be excessively conservative if the frequency difference is large. The correct procedure is to calculate the probability distributions from which the tolerance can be determined according to the statistical performance required.

If the image motion is the sum of any number of uniform linear motions and simple harmonic motions, then the loss can be calculated with the terms contained in Eqs. (72) and (73). First all linear motions can be combined into a resultant linear motion and the vibrations for each frequency combined into a single resultant term. Then the loss will consist of a term for the linear motion as given by the first term in Eq. (72), a term for each vibration as given by the second term in Eq. (72) or either of the first two terms in Eq. (73), an interaction term between the linear motion and each of the vibrations as given by the third term in Eq. (72), and an interaction term for every pair of vibrations, as given by the third term in Eq. (73).

The general procedure followed in this section can easily be extended to other kinds of image motion, if desired. The two kinds which have been used in this section are, however, the ones of greatest practical interest.

The author would like to acknowledge the advice and encouragement given by H. H. Hopkins, especially for suggesting the derivation of the equivalent transfer function for a focal-plane shutter, and also the help of his wife, Pamela Shack, in doing some of the more tedious calculations and preparing the manuscript.

References

1. H. Frieser, Phot. Sci. Eng. **4**, 324 (1960).
2. M. E. Bechtel, Cornell Aeronautical Laboratory, Rept. No. VF-1260-P-2, Contract No. AF33(616)-8870 (1959).
3. T. Asakura, J. Appl. Phys. Japan **30**, 797 (1961).
4. R. M. Scott, Phot. Sci. Eng. **3**, 201 (1959).
5. M. D. Rosenau, Photogrammetric Eng. 421 (June 1961).
6. T. T. Chang, Cornell Aeronautical Laboratory, Rept. No. VF-1260-P-2, Contract No. AF33(616)-8870 (1959).
7. D. P. Paris, Phot. Sci. Eng. **7**, 233 (1963).
8. L. O. Hendeberg and W. E. Welander, Appl. Opt. **2**, 379 (1963).