OPTICAL SCIENCES

THE UNIVERSITY OF ARIZONA
TUCSON, ARIZONA

TECHNICAL REPORT NO. 24

HOW WELL CAN A LENS SYSTEM TRANSMIT INFORMATION?

B. ROY FRIEDEN

March 15, 1968

OPTICAL SCIENCES

THE UNIVERSITY OF ARIZONA
TUCSON, ARIZONA

TECHNICAL REPORT NO. 24

HOW WELL CAN A LENS SYSTEM TRANSMIT INFORMATION?

B. ROY FRIEDEN

March 15, 1968

ABSTRACT

A lens system may be judged by its ability to relay information from object to image. A pertinent criterion of optical quality is h, the change in entropy between corresponding sampling points in the object and image planes. Since h is a unique function of the optical pupil, for a given bandpass 2Ω of the object, through the proper choice of a pupil function it is possible to maximize h at a given Ω. Physically, the optimum pupil function is an absorption coating applied to a diffraction-limited lens system. A numerical procedure is established for determining, with arbitrary accuracy, the optimum absorption coating, the resulting transfer function, and the maximum h, all at a given Ω. These quantities are determined, both for the one-dimensional pupil and the circular pupil, in the approximation that the optimum pupil function may be represented as a Fourier—(Bessel) series of five terms. The computed values of h_{max}, at a sequence of Ω values, are estimated to be correct to 0.2% for the 1-D pupil, and to 0.5% for the circular pupil. The optimum pupil functions are apodizers at small Ω and superresolvers at large Ω. Finally, we use the computed curve of $h_{max}^{(\Omega)}$ to relate the concept of "information transfer" to that of "classical resolving power": we show that a binary object (as defined) cannot radiate information to the image when the spacing between object sampling points is less than 0.87 times the Rayleigh resolution length.

CONTENTS

BACKGROUND

In 1947, Duffieux[1] discovered that a lens system is basically a low-pass, linear filter. Since then, the phenomenon of optical image formation has become increasingly viewed as a special case of linear systems analysis.[2] For example, such optical concepts as the system point response, transfer function, and convolution theorem[2] derive from linear systems analysis. Another such concept is "information," which was originally defined by Shannon[3] and has since been applied to the image-forming phenomenon.[4] In the absence of noise, as we shall assume, the amount of information in the message "event B has occurred" is measured by the "entropy" H, defined as

$$H = -\log(\text{probability of occurrence of B}) \qquad (1)$$

If the logarithm is of base 2, H is given the unit of "bit."

The total information transfer ΔH between object and image is defined as the sum of the incremental differences h in information content between H-values at corresponding sampling points in the object and image planes. It is a remarkable fact that all these increments h are equal, for any given object and lens. Furthermore, h is a unique function of the object bandpass 2Ω and the contrast transfer function[5] $\tau(\omega)$. (These properties of h are developed below.) Hence, h does not depend upon the object distribution, except through its cutoff frequency Ω. The total information transfer between object and image is therefore completely specified by one number, $h(\Omega)$, which is a property of the optics alone.

Using definition (1), with B now an intensity level at any sampling point in the image, we see that h determines the total possible number of distinguishable images. Hence, h is a measure of the system's ability to distinguish nearly identical objects. Intuitively, this information transfer must then correlate with the resolution capabilities of the system. This will be shown to be the case.

For the preceding reasons, the $h(\Omega)$ variation for an optical system is a valid figure of merit for the optical performance. This paper establishes a numerical method of finding the optical system (as specified by its pupil function) that maximizes h for a specified frequency bandwidth 2Ω of the object. The best 5-term pupil function is calculated for a subdivision of Ω values. The resulting curve of maximum h against Ω may be used as a standard for evaluating the performance of given lens systems. Also, the curve is shown to determine the maximum possible density of binary information (as defined) which is attainable in the object such that information reaches the image. Both one-dimensional and circular pupils are treated, and in that order.

In the Appendix, we use the calculus of variations to arrive at a formal solution to the problem, in the form of an integral equation connecting the optimum pupil function with the optimum transfer function. Although this integral equation is not used to generate a solution, we use it to check the approximate solutions due to the numerical scheme mentioned above.

ONE DIMENSIONAL CASE

The fundamental relation between information transfer and contrast transfer

We apply the linear systems approach of Goldman[6] to the optical case. Specifically, we seek the optical analog to his result[7] for the transfer of information per degree of freedom (sampling point) due to a linear process. Hence, the number of independent degrees of freedom in the optical object must first be established.

Degrees of freedom in optical object--The optical signal is the object, o(x), where x is a coordinate of length. The spectrum $\mathcal{O}(\omega)$ is defined as

$$\mathcal{O}(\omega) = (2\pi)^{-\frac{1}{2}} \int_{-\infty}^{\infty} dx\, o(x) e^{-j\omega x} \;, \quad j = (-1)^{\frac{1}{2}} \;, \tag{2a}$$

where ω is a spatial frequency. We restrict consideration to objects having a finite spatial extent 2X, i.e.

$$o(x) = 0 \quad \text{for } |x| > X \;. \tag{2b}$$

Substitution of (2b) into (2a) leads to the desired result,

$$\mathcal{O}(\omega) = \sum_{n=-\infty}^{\infty} \mathcal{O}(n\pi/X)\,\mathrm{sinc}(X\omega - n\pi) \tag{2c}$$

where $\mathrm{sinc}\,\phi = \sin(\phi)/\phi$

for general ϕ. This shows that $\mathcal{O}(\omega)$ is determined by $\mathcal{O}$ at the finite spacing π/X of frequencies $n\pi/X$. Values $n\pi/X$ locate the independent degrees of freedom in the object frequency space. Within an object bandwidth of 2Ω (i.e., all $-\Omega \le \omega \le \Omega$) there are then

$$2\pi^{-1}\Omega X \text{ degrees of freedom} \quad . \tag{3}$$

The information transfer--Using (3) in Goldman's result,[7] we find

$$\Delta H(\Omega) = \sum_{n=0}^{2\pi^{-1}\Omega X} \log |\tau(n\pi/X - \Omega)|^2 \tag{4a}$$

where

$$\Omega = \begin{cases} \text{object bandwidth if } \Omega < \Omega_0, \text{ or} \\ \text{largest multiple of } \pi/X < \Omega_0 \text{ if } \Omega \geq \Omega_0 \end{cases} \quad . \tag{4b}$$

Quantity $\Delta H(\Omega)$ is defined as the total entropy change, from object to image, over all frequencies $n\pi/X$ for which the *image* conveys non-zero information. Also, Ω_0 is the optical cutoff frequency, i.e.

$$\tau(\omega) = 0 \text{ for all } |\omega| \geq \Omega_0 \quad . \tag{5}$$

The reason for the bottom constraint in (4b) now becomes clear. Since[2] image $I(\omega) = \tau(\omega)O(\omega)$, by Eq. (5) $I(\omega) = 0$ for all $|\omega| \geq \Omega_0$. Hence, for all frequencies $n\pi/X \geq \Omega_0$, $I(n\pi/X) = 0$ with probability unity. Therefore, by definition (1) frequencies $n\pi/X \geq \Omega_0$ convey zero information, and by the above definition of ΔH such frequencies must be excluded from (4a).

Realistically assuming that the number of degrees of freedom

$$2\pi^{-1}\Omega X >> 1 \quad , \tag{6}$$

summation (4a) may be replaced by the integral

$$\Delta H(\Omega) = \pi^{-1} X \int_{-\Omega}^{\Omega} \log |\tau(\omega)|^2 d\omega \tag{7}$$

with negligible error. The change in information *per degree of freedom*, $h(\Omega)$, may then be obtained by dividing Eq. (7) by quantity (3):

$$h(\Omega) = (2\Omega)^{-1} \int_{-\Omega}^{\Omega} \log |\tau(\omega)|^2 d\omega \quad . \tag{8}$$

This is the quantity which we maximize below. Because $\tau(\omega) = \tau(\omega/\Omega_o)$, Eq. (8) shows that $h = h(\Omega/\Omega_o)$. The notation $h(\Omega)$ will be frequently used, however, for brevity.

It is important to note that, no matter how large the object bandwidth may be, Eq. (8) remains finite. This is due to the bottom constraint in (4b).

Properties of the information transfer function

We notice some important properties of $h(\Omega)$ in Eq. (8). First, because it depends on the modulus of $\tau(\omega)$, h is insensitive to spatial phase displacements over the image. Thus, information is not lost due to, e.g., strong spurious resolution. Of course, as far as the unaided human observer is concerned, the information is (in this case) hopelessly scrambled. However, Eq. (8) states that it is all there, and can be extracted provided the information is processed in the proper way.

Next, we note that because $|\tau| \leq 1$,

$$h < 0 \quad . \tag{9}$$

Therefore, information can only be *lost* in traversal through the optics. Or, pictorial information can only proceed from a state of "disorder" in the object, to one of "order" (i.e., gray blur) in the image. Our maximuzed $h(\Omega)$ must therefore be negative, and closer to zero than any other $h(\Omega)$.

Information loss h is observed in (8) to be independent of the object structure. All dependence is solely upon the contrast transfer function $\tau(\omega)$. Function $h(\Omega)$ is therefore characteristic of the lens system *alone;* hence a continuous plot of $h(\Omega)$ determines the system's ability to transmit the information in *any* object. This is the pertinent quality criterion if information relay is of prime importance. If $h(\Omega)$ is used as a quality criterion, it is important to know the curve of maximum $h(\Omega)$, and this we shall establish.

Finally, we note that $h(\Omega)$ is not generally maximized by the use of diffraction-limited, uncoated optics. For example, when Ω is small, there are an infinite number of pupils, known as "apodizers,"[8] for which $|\tau(\omega)|$ exceeds the diffraction-limited $|\tau(\omega)|$ at all $|\omega| \leq \Omega$. Eq. (8) shows that h for these apodizers exceeds the diffraction-limited h.

Fundamental properties of the best pupil

We show that the best pupil is both real and even. This greatly simplifies its calculation.

Let $U(\beta)$ represent a general pupil function, where β is the "reduced" pupil coordinate $2\pi\lambda^{-1}t$, λ is the wavelength of light and t is the coordinate of length in the pupil. Let the optimum pupil function, contrast transfer function, and information transfer function be denoted as U_{max}, τ_{max} and h_{max}, respectively, for a given ratio Ω/Ω_o. We sometimes use the notation $U_{max}(\beta;\Omega/\Omega_o)$, $\tau_{max}(\omega;\Omega/\Omega_o)$ and $h_{max}(\Omega/\Omega_o)$ to emphasize that these quantities depend upon Ω through the ratio Ω/Ω_o.

We have already concluded that for Ω/Ω_o small, U_{max} corresponds to an apodizing pupil. We next consider the opposite case, that of Ω close to Ω_o.

As Ω approaches Ω_o, the largest values of ω in integral (8) also approach Ω_o. Now $\tau(\omega) \to 0$ as $\omega \to \Omega_o$. Therefore $\log |\tau(\omega)|^2$ is negative and very large in magnitude for these values of ω. Hence, for Ω close to Ω_o the overwhelming contribution to h derives from values of $\tau(\omega)$ at large ω. In order to maximize h toward zero, the best pupil will therefore emphasize $\tau(\omega)$ at large ω, and this must result in small values for $\tau(\omega)$ at small ω. This is the known property of a "superresolver."[9]

Having deduced the qualitative trend in τ_{max} and U_{max} as $\Omega \to 0$ and as $\Omega \to \Omega_o$, we next consider general Ω and show that h_{max} results from a purely real pupil. This result greatly simplifies the ensuing calculation of $h_{max}(\Omega/\Omega_o)$.

The best pupil is real--Proof--Transfer function $\tau(\omega)$ is related to the general pupil function $U(\beta)$ by[2]

$$\tau(\omega) = \int_{\omega-\beta_o}^{\beta_o} d\beta U(\beta)U^*(\beta - \omega) \Bigg/ \int_{-\beta_o}^{\beta_o} d\beta U(\beta)U^*(\beta) \ , \tag{10}$$

where * denotes the complex conjugate

Now $U(\beta)$, being a general complex number, may be expressed as

$$U(\beta) = U_{re}(\beta)e^{j\phi(\beta)} \tag{11a}$$

where

$$U_{re}(\beta) = |U(\beta)| \tag{11b}$$

and $\phi(\beta)$ is the phase. Substituting (11a) into (10), and multiplying (10) by its complex conjugate, leads to

$$|\tau(\omega)|^2 = \left| \int_{\omega-\beta_o}^{\beta_o} d\beta U_{re}(\beta)U_{re}(\beta - \omega)e^{j\Phi(\beta,\omega)} \right| \Bigg/ \left(\int_{-\beta_o}^{\beta_o} d\beta [U_{re}(\beta)]^2 \right)^2 \tag{12a}$$

where

$$\Phi(\beta,\omega) = \phi(\beta) - \phi(\beta - \omega) \quad . \tag{12b}$$

Now Schwarz's inequality is

$$\left| \int_a^b d\beta f(\beta) g^*(\beta) d\beta \right|^2 \leq \int_a^b d\beta |f(\beta)|^2 \int_a^b d\beta |g(\beta)|^2 \tag{13}$$

for any functions f,g. We let

$$f(\beta) = [U_{re}(\beta) U_{re}(\beta - \omega)]^{\frac{1}{2}}, \tag{14a}$$

$$g^*(\beta) = f(\beta) e^{j\Phi(\beta,\omega)} \tag{14b}$$

and substitute into (12a). Then by (13),

$$|\tau(\omega)|^2 \leq [\int_{\omega-\beta_o}^{\beta_o} d\beta U_{re}(\beta) U_{re}(\beta - \omega)]^2 \Bigg/ \left(\int_{-\beta_o}^{\beta_o} d\beta [U_{re}(\beta)]^2 \right)^2 . \tag{15}$$

The right-hand side may be recognized as the square of $\tau_{re}(\omega)$, the transfer function due to pupil $U_{re}(\beta)$. We have therefore shown that

$$|\tau(\omega)|^2 \leq [\tau_{re}(\omega)]^2 \ , \tag{16}$$

or that *at every* ω *the* $|\tau(\omega)|$ *due to a generally complex pupil* $U(\beta)$ *is exceeded by the* $|\tau(\omega)|$ *resulting from the real pupil* $|U(\beta)|$. Since a real pupil maximizes $|\tau(\omega)|$ at all ω, it likewise maximizes the integral (8) for $h(\Omega)$.

Since U_{max} is real, U_{max} represents an absorption coating in the pupil of a diffraction-limited lens system.

The best pupil is even--This will be established a posteriori, i.e., after the solutions U_{max}, τ_{max} are known. In the Appendix we derive an integral equation which must be satisfied by the optimum pupil. We substitute into this equation the independently computed solutions U_{max}, τ_{max} which are based on the assumption that

$$U_{max}(-\beta) = U_{max}(\beta) \quad . \tag{17}$$

It is shown that the integral equation is well satisfied by the even solutions.

It is a fortunate coincidence that the even pupil has other desirable attributes.[10] As compared to a general, real pupil $U(\beta)$, the even pupil $[U(\beta) + U(-\beta)]$ produces (a) a larger Strehl intensity ratio, and (b) a narrower central core in the point spread function, with (c) smaller secondary maxima.

Because U_{max} is both real and even, τ_{max} is both real and even in ω. Hence, the problem has been simplified to finding the $U(\beta)$ that results in a $\tau(\omega)$ which maximizes

$$h(\Omega) = 2\Omega^{-1} \int_0^{\Omega} \log\tau(\omega)d\omega \quad . \tag{18}$$

<u>Method of solution</u>

The calculus of variations is often the first method tried in an extremum problem such as ours, since it frequently leads to an analytic solution.[11] In our case, however, it leads to integral equation (j), of the Appendix, which involves both $\tau(\omega)$ and $U(\beta)$ as unknown functions. We could not enforce numerical solutions in Eq. (j) because the relation between $\tau(\omega)$ and $U(\beta)$, Eq. (10), only tends to further complicate (j). Consequently, (j) was used only as a check on solutions which were found in the following way.[12]

We expand $U(\beta)$ as a Fourier cosine series with unknown coefficients a_m

$$U(\beta) \equiv U^{(M+1)}(\beta) = \sum_{m=0}^{M} a_m \cos m\pi\beta/\beta_o \tag{19}$$

where β_o is the pupil half-width. The a_m that simultaneously satisfy

$$\partial h/\partial a_m = 0, \quad m = 0,1,\ldots, M \tag{20}$$

determine $U_{max}^{(M+1)}(\beta)$, the (M + 1)-term approximation to $U_{max}(\beta)$.

We denote the maximized h due to (20) as $h_{max}^{(M+1)}$, the corresponding $\tau(\omega)$ as $\tau_{max}^{(M+1)}(\omega)$. Although constraints (20) guarantee that h is an extreme value, and not necessarily a maximum, comparison of $h_{max}^{(M+1)}$ with the uncoated, diffraction-limited value h_o establishes that h is indeed maximized by the procedure. This is shown below.

Eqs. (19) and (20) imply that $h_{max}^{(\infty)} = h_{max}$, which might be termed the "absolute maximum." Therefore, for M finite, $h_{max}^{(M+1)} < h_{max}$. However, we find below that $h_{max}^{(5)} \simeq h_{max}$ to better than 0.2%.

Substitution of Eq. (19) into Eq. (10) results in

$$\tau(\omega) \equiv \tau^{(M+1)}(\omega) = N(\omega)/N(0) \tag{21a}$$

where

$$N(\omega) = \sum_{\ell,m} a_\ell a_m f_{\ell m}(\omega) \tag{21b}$$

and the

$$f_{\ell m}(\omega) = \int_{\omega-\beta_o}^{\beta_o} d\beta \cos(m\pi\beta/\beta_o)\cos[\ell\pi(\beta-\omega)/\beta_o] \tag{22}$$

The integral in (22) may be evaluated in closed form. All summations vary from 0 to M independently. Substitution of Eqs. (21), (22) into (18) then leads to an expression for $h(a_o, a_1, \ldots, a_M)$. In order to find the a_m that enforce conditions (20) we use the following technique. Replace the M + 1 conditions (20) by the "merit function"

$$W(a_o, a_1, \ldots, a_M) \equiv \sum_{m=0}^{M} (\partial h/\partial a_m)^2 = 0 \quad . \tag{23}$$

If W = 0, constraints (20) must be satisfied. We reduce W to zero by a recursive technique based on Newton's method.[13] Let superscript (k) for a quantity represent its value during iteration k. Newton's method consists of assuming $W^{(k)}$ to be so close to zero that infinitesimally small changes $\Delta a_m^{(k+1)}$ in $a_m^{(k)}$ can make $W^{(k+1)} = 0$. That is, that

$$W^{(k+1)} - W^{(k)} \equiv -W^{(k)} \simeq \sum_{m=0}^{M} (\partial W/\partial a_m)^{(k)} \Delta a_m^{(k+1)} \tag{24a}$$

where

$$a_m^{(k+1)} \equiv a_m^{(k)} + \Delta a_m^{(k+1)} \tag{24b}$$

At each iteration k, Eq. (24a) represents one linear equation in M unknowns $\Delta a_m^{(k+1)}$. It therefore has an infinite number of solutions. We remove this ambiguity and, at the same time, strengthen the validity of the approximate equality in Eq. (24a). This is accomplished by demanding the $\Delta a_m^{(k+1)}$ to be minimally small. Hence, let the $\Delta a_m^{(k+1)}$ satisfying (24a) also obey

$$\sum_{m=0}^{M} [\Delta a_m^{(k+1)}]^2 = \text{minimum} \quad . \tag{24c}$$

Fortunately, the unique solution to Eqs. (24) is known, and has the very simple form

$$\Delta a_m^{(k+1)} = - \frac{W^{(k)}}{\sum_{n=0}^{M} [(\partial W/\partial a_n)^{(k)}]^2} \left(\frac{\partial W}{\partial a_m}\right)^{(k)} \tag{25}$$

Solution (25) is established by use of Lagrange's method of undetermined multipliers.[14] All quantities on the right-hand side of (25) may be related to the $a_m^{(k)}$, by use of Eqs. (18), (22) and (23).

For calculation of the right-hand side of (25), the following formulae were programmed on an IBM 7072 computer. The superscript (k) is dropped because the formulae all apply to the k*th* iteration.

$$W = \sum_{m=0}^{M} (\partial h/\partial a_m)^2 \ , \tag{26a}$$

where

$$\partial h/\partial a_m = 4\Omega^{-1} \sum_{i=1}^{N} [H_i/N(\omega_i)] \sum_{n=0}^{M} a_n f_{mn}(\omega_i) - 4\beta_o \varepsilon_m a_m/\Omega N(0) \ ; \tag{26b}$$

and

$$\partial W/\partial a_m = 2 \sum_{n=0}^{M} (\partial h/\partial a_n)(\partial^2 h/\partial a_n \partial a_m) \ , \tag{26c}$$

where

$$\frac{\partial^2 h}{\partial a_n \partial a_m} = 4\Omega^{-1} \sum_{i=1}^{N} \frac{H_i}{N(\omega_i)} [f_{nm}(\omega_i) -$$

$$\frac{2}{N(\omega_i)} \left(\sum_{j=0}^{M} a_j f_{jm}(\omega_i) \right) \left(\sum_{k=0}^{M} a_k f_{kn}(\omega_i) \right)]$$

$$-4N(0)^{-1} \beta_o \varepsilon_n \delta_{nm} + 8N(0)^{-2} \beta_o^{2} \varepsilon_n \varepsilon_m a_n a_m \quad . \tag{26d}$$

In the above, the notation

$$\varepsilon_o = 2, \quad \varepsilon_m = 1 \quad \text{for } m = 1, 2, \ldots$$
$$\delta_{nm} = 0 \quad \text{for } n \neq m, \quad \delta_{nm} = 1 \quad \text{for } n = m \tag{26e}$$

is used. Also, the $\sum_{i=1}^{N} H_i$ is the Gauss-quadrature[15] approximation to $\int_0^\Omega d\omega$. The H_i, ω_i are the weights and abscissae for the N-point Gauss quadrature over interval $0 \leq \omega \leq \Omega$.

In order to evaluate $h^{(M+1)}(\Omega)$ at each iteration k, the Gauss-quadrature formula

$$h^{(M+1)}(\Omega) = 2\Omega^{-1} \sum_{i=1}^{N} H_i \log \tau^{(M+1)}(\omega_i) \tag{27}$$

was used. This relation derives from Eq. (18).

Starting the calculation with an arbitrarily chosen set of a_m (e.g., $a_o = 1$, $a_j = 0$, $j = 1, 2, \ldots, M$), Eqs. (22), (25), (26) and (27) were recursively evaluated until $h^{(M+1)}$ changed by less than 0.001% from one iteration to the next.

Computer output

Once a solution a_m, m = 0,..., M, was found for a given ratio Ω/Ω_o, the following quantities were printed out: $h_{max}^{(M+1)}$, the a_m, normalized coefficients a_m' proportional to the a_m and such that the largest value of $U_{max}^{(M+1)}$ is unity, the normalized pupil function $U_{max}^{(M+1)}(\beta)$, $\tau_{max}^{(M+1)}(\omega)$, the Strehl[2] intensity ratio S, and the total energy relative to the energy through the uncoated aperture, E. The last two quantities were computed by the formulae

$$S = a_o'^2 \tag{28a}$$

$$E = a_o'^2 + 2^{-1} \sum_{m=1}^{M} a_m'^2 \quad . \tag{28b}$$

These formulae derive from use of series (19) in the Fraunhofer approximation[2] formula for the point amplitude distribution. The simplicity in the form of Eqs. (28) is due to the orthogonality of the cosine functions in Eq. (19).

Numerical results

The case M = 4, i.e., a 5-term pupil coating, is of interest because the computed values $h_{max}^{(5)}(\Omega/\Omega_o)$ are less than $h_{max}(\Omega/\Omega_o)$ by no more than 0.2%, for $0.1 \le \Omega/\Omega_o \le 0.9$. This was established by comparison among $h^{(5)}$, $h^{(10)}$ and $h^{(15)}$ for test values of Ω/Ω_o in the indicated range. Hence, in most practical cases, the solutions $U_{max}^{(5)}(\beta)$, $\tau_{max}^{(5)}(\omega)$, $h_{max}^{(5)}$ will suffice.

In Fig. 1 we show five pairs of curves, the $U_{max}^{(5)}(\beta)$ and $\tau_{max}^{(5)}(\omega)$ for relative bandpass values of Ω/Ω_o = 0.2, 0.4, 0.6, 0.8 and 0.98. For comparison, the corresponding curves $U_o(\beta) = 1$, $\tau_o(\omega)$ due to uncoated, diffraction limited optics, are also shown, where[2]

$$\tau_o(\omega) = \begin{cases} 1 - |\omega|/\Omega_o, & |\omega| \le \Omega_o \\ 0, & |\omega| > \Omega_o \end{cases} \tag{29}$$

We note that for small Ω/Ω_o, $\tau^{(5)}_{max}(\omega)$ is emphasized at low ω, behavior typical of an apodizer,[8] while for large Ω/Ω_o, $\tau^{(5)}_{max}(\omega)$ is emphasized at high ω, which is typical of a superresolver.[9]

Every curve $U^{(5)}_{max}(\beta)$ in Fig. 1 should appear to have zero slope both at the origin and at the pupil edge. The Fourier series representation (19) for $U^{(M+1)}_{max}(\beta)$ has this property independent of the a_m.

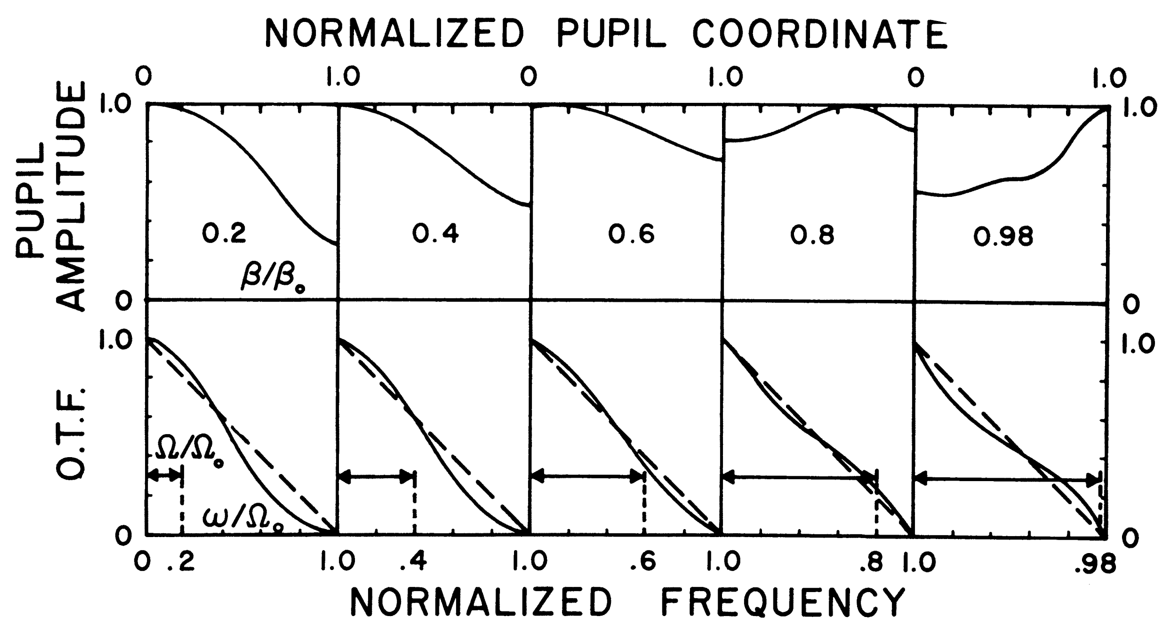

Fig. 1. One-dimensional case. Optimum 5-term pupil functions $U^{(5)}_{max}(\beta)$ *(top row), and resulting optical transfer functions* $\tau^{(5)}_{max}(\omega)$ *(bottom row, solid) for a subdivision of bandpass values* Ω/Ω_o*. Comparison of* $\tau^{(5)}_{max}(\omega)$ *with transfer function* $\tau_o(\omega)$ *(dashed curves) due to uncoated, diffraction-limited optics, shows that low values of* Ω/Ω_o *emphasize the low-frequency band of* $\tau^{(5)}_{max}(\omega)$*; high values emphasize the high-frequency band. These are the known properties of apodizers and superresolvers, respectively.*

Table 1 lists the coefficients a'_m for a subdivision of values $\Omega/\Omega_o = 0.1p$, $p = 1,2,\ldots,9$. These determine the normalized pupil functions $U^{(5)}_{max}(\beta;\Omega/\Omega_o)$ through Eq. (19). It appears that the subdivision of Ω/Ω_o values is fine enough to allow approximate determination of the a'_m for any Ω/Ω_o in the *continuum* between values 0.1 and 0.5.

Table 1. Fourier coefficients for determination of the optimum five-term pupil coating at given Ω/Ω_o

Ω/Ω_o	a'_o	a'_1	a'_2	a'_3	a'_4
0.1	0.676028	0.386682	-0.078226	0.032734	-0.016217
0.2	0.712258	0.338922	-0.064711	0.022647	-0.009115
0.3	0.748043	0.297808	-0.046348	0.010488	-0.009990
0.4	0.774973	0.244634	-0.022603	0.014912	-0.011917
0.5	0.792682	0.184117	-0.001445	0.025059	-0.000412
0.6	0.890320	0.130003	-0.021582	0.001442	-0.014290
0.7	0.931532	0.039009	-0.063136	-0.004379	0.000151
0.8	0.906596	-0.067098	-0.049395	0.033155	-0.003577
0.9	0.842173	-0.153790	0.027137	0.008971	-0.018282

Fig. 2 plots the computed $h^{(5)}_{max}(\Omega/\Omega_o)$ and compares this with the curve $h_o(\Omega/\Omega_o)$ due to diffraction-limited, uncoated optics. The latter is obtained by substitution of Eq. (29) into Eq. (18) for h. An integration produces the result

$$h_o(\omega') = -2\omega'^{-1}(1-\omega')\log(1-\omega') - 2\omega', \qquad (30a)$$

where

$$\omega' \equiv \Omega/\Omega_o. \tag{30b}$$

We see that for $0 \le \Omega/\Omega_o \le 0.5$ there is considerable relative gain in relayed information due to the optimum coatings; in fact, for $0 < \Omega/\Omega_o \le 0.2$ this gain is at least 100%.

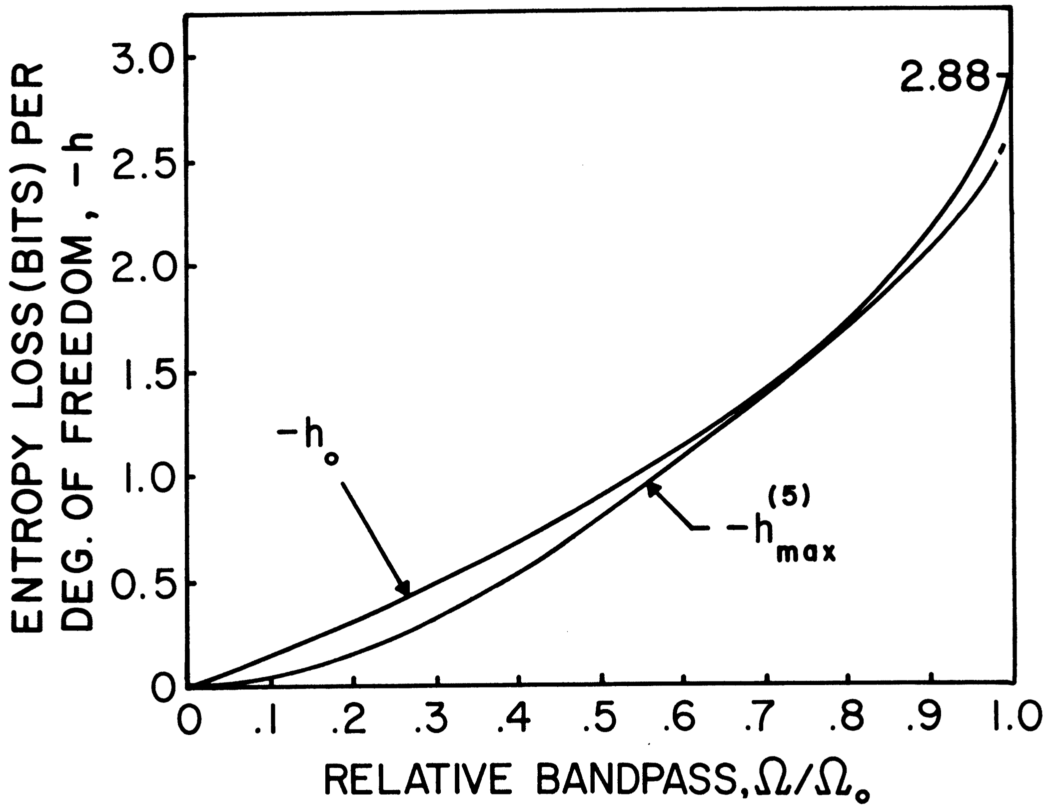

Fig. 2. In answer to the question posed by the title, we show the curve of minimum information loss per degree of freedom, $-h^{(5)}_{max}(\Omega/\Omega_o)$. *This is compared with* $-h_o(\Omega/\Omega_o)$, *information loss per degree of freedom due to uncoated, diffraction-limited optics. For* Ω/Ω_o *in the range 0.1 through 0.9,* $h^{(5)}_{max} = h_{max}$, *the absolute optimum value, to better than 0.2% accuracy. The advantage in information relay of* h_{max} *over* h_o *is seen to be mainly for* $\Omega/\Omega_o \le 0.5$.

A real coating, such as U_{max}, will decrease both the total radiant energy passing through the pupil, and the maximum of intensity in the point spread function, from the respective values for U_o. The parameters E and S are plotted against Ω/Ω_o in Fig. 3. These two curves indicate that coatings $U_{max}^{(5)}$ generally allow more than 50% of the incident light energy to reach the image space; and, for $\Omega/\Omega_o > 0.2$, a Strehl intensity ratio better than 50%. Generally speaking, these figures do not indicate a serious light loss.

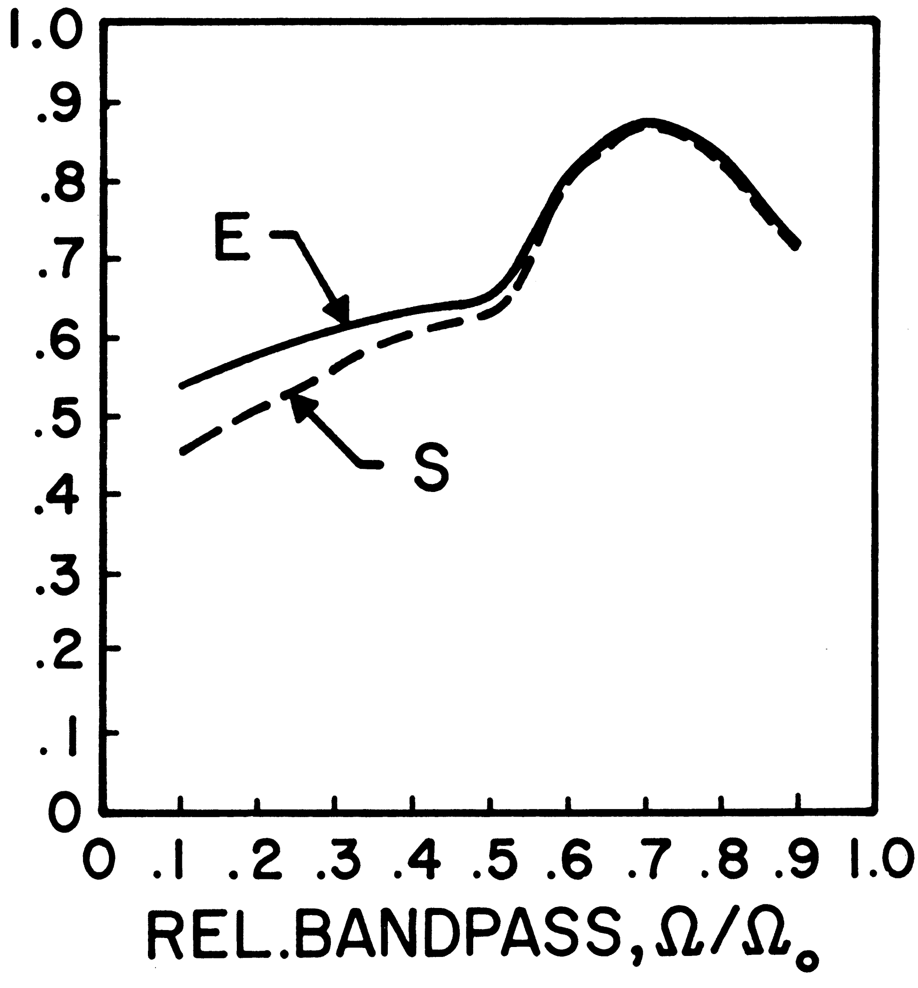

Fig. 3. Total relative energy E and Strehl intensity ratio S resulting from use of coatings $U_{max}^{(5)}(\beta)$ at $0.1 \le \Omega/\Omega_o \le 0.9$. Parameters E and S are observed to be smallest for $\Omega/\Omega_o \lesssim 0.5$, where the advantage of $h_{max}^{(5)}$ over h_o is the greatest (see Fig. 2). The curves of E and S indicate a moderately small loss of light owing to the optimum coatings.

TWO-DIMENSIONAL CASE

To keep notation to a minimum, we use the same symbol for a 2-D quantity as for the corresponding 1-D quantity. The argument of a quantity is made 2-D, where required, by use of subscripts 1 and 2. For example, $\mathcal{O}(\omega) \to \mathcal{O}(\omega_1, \omega_2)$.

Those derivations which are analogous to 1-D derivations are abbreviated.

Loss of information from object to image

For a general object $o(x_1, x_2)$ confined to a rectangular area $|x_1| \leq X_1$, $|x_2| \leq X_2$ the degrees of freedom in object frequency space (ω_1, ω_2) are located at values $\omega_1 = m\pi/X_1$, $\omega_2 = n\pi/X_2$. Then within the rectangular passband $|\omega_1| \leq \Omega_1$, $|\omega_2| \leq \Omega_2$ there are

$$4\pi^{-2}\Omega_1 X_1 \Omega_2 X_2 \text{ degrees of freedom} \equiv P \quad . \tag{31}$$

The 2-D analog to Eq. (4a) results. There is a total information relay of

$$\Delta H(\Omega_1, \Omega_2) = \sum_{m=0}^{P} \sum_{n=0}^{P} \log \left|\tau\left(\frac{m\pi}{X_1} - \Omega_1, \frac{n\pi}{X_2} - \Omega_2\right)\right|^2 \tag{32}$$

from object to image, where (Ω_1, Ω_2) are constrained as in Eq. (4b). Assuming $2\pi^{-1}\Omega_1 X_1$ and $2\pi^{-1}\Omega_2 X_2$ to be sufficiently large numbers, Eq. (32) may be approximated by the integral

$$\Delta H(\Omega_1, \Omega_2) = \pi^{-2} X_1 X_2 \int_{-\Omega_1}^{\Omega_1} \int_{-\Omega_2}^{\Omega_2} \log |\tau(\omega_1, \omega_2)|^2 \, d\omega_1 d\omega_2 \quad . \tag{33}$$

Combining Eq. (33) with (31), the change in information per degree of freedom, $h(\Omega_1, \Omega_2)$ obeys

$$h(\Omega_1,\Omega_2) = (4\Omega_1\ \Omega_2)^{-1} \int_{-\Omega_1}^{\Omega_1} \int_{-\Omega_2}^{\Omega_2} \log\ |\tau(\omega_1,\omega_2)|^2\ d\omega_1 d\omega_2\ . \tag{34}$$

This is independent of the object structure. The comments following Eq. (8) apply equally well to Eq. (34).

We next consider general properties of the pupil function $U(\beta_1,\ \beta_2)$ which maximizes $h(\Omega_1,\ \Omega_2)$.

Fundamental properties of best 2-D pupil

The best pupil is real--The connection between h and U is through Eq. (29) and[2]

$$\tau(\omega_1,\omega_2) = \iint_{-\infty}^{\infty} U(\beta_1,\beta_2)U^*(\beta_1 - \omega_1,\beta_2 - \omega_2)d\beta_1 d\beta_2 \Big/ \iint_{-\infty}^{\infty} |U(\beta_1,\beta_2)|^2\ d\beta_1 d\beta_2 \tag{35}$$

The infinite limits are required because we have not yet decided what shape the optimum pupil is to be. By use of the 2-D representation of the Schwarz inequality, as in steps (11)-(16), we establish that

$$|\tau(\omega_1,\omega_2)|^2 \le [\tau_{re}(\omega_1,\omega_2)]^2. \tag{36}$$

In (36), τ_{re} results from the real pupil function $U_{re} = |U|$, while τ results from U. Since (36) holds true at all (ω_1,ω_2), h in Eq. (29) is maximized by a real pupil function.

The statistically best pupil has radial symmetry--The statistical average over all possible objects is an omnidirectional distribution: there is no preferred direction of resolution or contrast. Therefore, the average bandpass region is circular. Let this have radius Ω, so that in Eq. (34),

$$\Omega_1 = \Omega_2 = \Omega \tag{37}$$

Substitution of (37) into (34) results in $h(\Omega_1,\Omega_2) \to h(\Omega)$. Since there is therefore no constraint on h_{max} or on τ_{max} which favors a particular direction in (ω_1,ω_2)-space, necessarily

$$\tau_{max}(\omega_1,\omega_2) = \tau_{max}(\omega) \tag{38a}$$

where

$$\omega = (\omega_1^2 + \omega_2^2)^{\frac{1}{2}} \tag{38b}$$

This is also the best choice from a practical standpoint. That is, for a random assortment of objects the best average resemblance between image and object occurs when resolution is the same in all directions.

Substitution of constraints (37) and (38) into Eq. (34) leads to the simplification

$$h(\Omega) = \pi\Omega^{-2} \int_0^{\Omega} \omega \log \tau(\omega)d\omega \tag{39}$$

Because of Eqs. (38), the pupil $U_{max}(\beta_1,\beta_2)$ must be radially symmetric, or

$$U_{max}(\beta_1,\beta_2) = U_{max}(\rho) \tag{40a}$$

where

$$\rho = (\beta_1^2+\beta_2^2)^{\frac{1}{2}} \tag{40b}$$

In all practical cases a pupil is of finite extent. Hence the maximum extent of U_{max} is described by a reduced radius, ρ_o.

Method of solution

We represent $U(\rho)$ as a Fourier-Bessel series

$$U(\rho) \equiv U^{(M+1)}(\rho) = \sum_{m=0}^{M} b_m J_o(\mu_m \rho) \tag{41a}$$

where the μ_m are defined by

$$J_1(\mu_m \rho_o) = 0, \quad m = 1,2,\ldots, M \tag{41b}$$

and

$$\mu_o = 0 \quad . \tag{41c}$$

Series (41a) is sometimes known as Dini's expansion.[16]

We seek the b_m that satisfy

$$\partial h/\partial b_m = 0, \quad m = 0,1,2,\ldots, M \quad , \tag{42}$$

thereby making h an extremum. The resulting $U^{(M+1)}(\rho)$ is called $U_{max}^{(M+1)}(\rho)$, where

$$\lim_{M\to\infty} U_{max}^{(M+1)} = U_{max} \quad . \tag{43}$$

Representation (41) for $U^{(M+1)}(\rho)$ is the circular analogy to representation (19) for $U^{(M+1)}(\beta)$. Thus, $U_{max}^{(M+1)}(\rho)$ has the property

$$\partial U^{(M+1)}(0)/\partial\rho = \partial U^{(M+1)}(\rho_o)/\partial\rho = 0 \quad . \tag{44}$$

The contrast transfer function corresponding to $U_{max}^{(M+1)}(\rho)$ is denoted as $\tau_{max}^{(M+1)}(\omega)$, the information transfer as $h_{max}^{(M+1)}$.

As in the 1-D development, we must relate $\tau^{(M+1)}(\omega)$ to the b_m and then substitute into Eq. (39), so as to enable extremal conditions (42) to be enforced. Unfortunately, the link between $\tau^{(M+1)}(\omega)$ and $U^{(M+1)}(\beta)$, Eq. (35), is now a 2-D integral, and this is a time-consuming evaluation, especially when performed repeatedly in a recursive evaluation. We chose the following alternative route.

The point amplitude u at radial coordinate r in image space is given by[2]

$$u(r) = \int_0^{\rho_o} d\rho \; \rho J_o(\rho r) U(\rho) \tag{45}$$

in the case of rotational symmetry. By substituting (41a) into (45) we have the analytic expression

$$\underset{r \neq \mu_m}{u(r)} = \sum_{m=0}^{M} b_m f(r, \mu_m) \tag{46a}$$

where

$$f(r,\mu_m) = \rho_o(\mu_m^2 - r^2)^{-1} \left[\mu_m J_o(\rho_o r) J_1(\rho_o \mu_m) - r J_o(\rho_o \mu_m) J_1(\rho_o r)\right]. \tag{46b}$$

It is known that $\tau(\omega)$ is related to u(r) through the sampling expression[17]

$$\tau(\omega) = \sum_{p=1}^{\infty} \frac{u(\lambda_p)^2}{J_1(2\rho_o\lambda_p)^2} J_o(\lambda_p \omega) \Bigg/ \sum_{p=1}^{\infty} \frac{u(\lambda_p)^2}{J_1(2\rho_o\lambda_p)^2} \tag{47a}$$

where the λ_p are defined by

$$J_o(2\rho_o\lambda_p) = 0, \quad p = 1,2,... \tag{47b}$$

Hence, we evaluate $\tau^{(M+1)}(\omega)$ by calculation of the $u(\lambda_p)$ in Eq. (46a), with substitution into (47a). The only approximation derives from truncation of the two series in (47a), which we enforced beyond 20 terms. This error is insignificant.

Solution for the b_m satisfying (42) is by the "merit function" approach described in the 1-D calculation. Formulae equivalent to Eqs. (24)-(27) were programmed for calculation on an IBM 7072 computer. In the interest of brevity, these will not be further described.

Computer output

Once a solution b_m, $m = 0,1,...,M$ is found for a given ratio Ω/Ω_o, the following quantities are printed out: $h_{max}^{(M+1)}$, the final b_m, normalized $b_m(\equiv b_m')$ such that the largest value of $U_{max}^{(M+1)}$ is unity, the normalized pupil function $U_{max}^{(M+1)}(\rho)$, $\tau_{max}^{(M+1)}(\omega)$, Strehl intensity ratio S, and the total energy relative to the energy through the uncoated aperture, E. The last two quantities are computed by the formulae

$$S = [b_o' + 2 \sum_{m=1}^{M} b_m' J_1(\mu_m)/\mu_m]^2 \tag{48a}$$

and

$$E = \sum_{m=0}^{M} [b_m' J_o(\mu_m)]^2. \tag{48b}$$

These are easily derived from pupil representation (41). In Eqs. (48) The particular value $\rho_o = 1$ is implicit. (This value was used throughout the calculation.)

Numerical results

We show results for the case M = 4, which represents a 5-term pupil coating. The acquired values of $-h^{(5)}_{max}(\Omega/\Omega_o)$ are believed to exceed $-h_{max}(\Omega/\Omega_o)$ by no more than 0.5% for $0.1 \le \Omega/\Omega_o \le 0.9$.

In Fig. 4 we show the results $U^{(5)}_{max}(\rho;\Omega/\Omega_o)$ and $\tau^{(5)}_{max}(\omega;\Omega/\Omega_o)$ for values of Ω/Ω_o = 0.2, 0.4, 0.6, 0.8 and 0.99. For comparison, the corresponding curves $U_o(\rho)$ and $\tau_o(\omega)$ due to uncoated, diffraction-limited optics are also shown, where[2]

$$\tau_o(\omega) = \begin{cases} 2\pi^{-1}[\cos^{-1}(\omega') - \omega'(1-\omega'^2)^{\frac{1}{2}}], & |\omega'| \le 1 \\ 0, \quad |\omega'| > 1 & (\omega' \equiv \omega/\Omega_o) \end{cases} \tag{49}$$

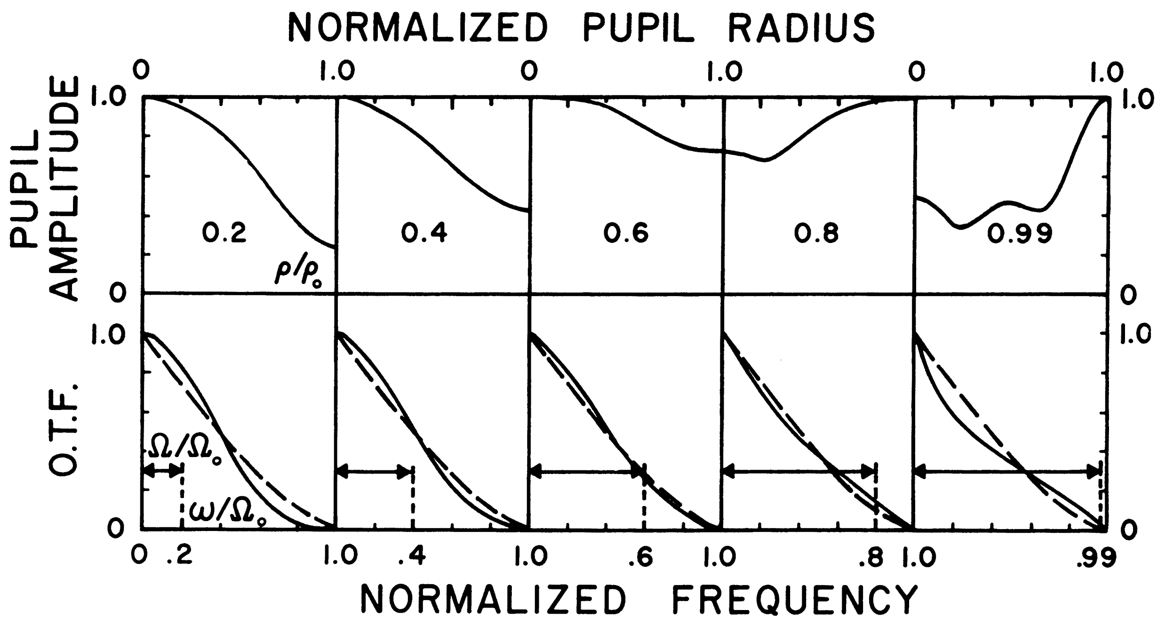

Fig. 4. Two-dimensional case. Optimum 5-term pupil functions $U^{(5)}_{max}(\rho)$, (top row), and resulting optical transfer functions $\tau^{(5)}_{max}(\omega)$ (bottom row, solid) for a subdivision of bandpass values Ω/Ω_o. Comparison of $\tau^{(5)}_{max}(\omega)$ with transfer function $\tau_o(\omega)$ (dashed curves) due to uncoated, diffraction-limited optics, shows that low values of Ω/Ω_o emphasize the low-frequency band of $\tau^{(5)}_{max}(\omega)$; high values emphasize the high-frequency band. These are the known properties of apodizers and superresolvers, respectively.

The behavior of the curves of Fig. 4 is analogous to that of those in Fig. 1, the one-dimensional result.

Table 2 lists the coefficients b_m' for the subdivision $\Omega/\Omega_o = 0.1p$, $p = 1,2,\ldots,9$. These determine the normalized optimal 5-term pupil coatings through Eqs. (41). It appears that, by interpolation, the b_m' for any Ω/Ω_o in the continuum between 0.1 and 0.5 may be found with fair accuracy.

Table 2. Fourier-Bessel coefficients for determination of the optimum 5-term pupil coating at given Ω/Ω_o

Ω/Ω_o	b_o'	b_1'	b_2'	b_3'	b_4'
0.1	0.502329	0.642166	-0.184746	0.086046	-0.045794
0.2	0.534371	0.571775	-0.148746	0.057869	-0.015268
0.3	0.583486	0.505800	-0.109870	0.020558	0.000026
0.4	0.619810	0.421183	-0.057414	0.018104	-0.001682
0.5	0.643823	0.310939	-0.009604	0.027549	0.027293
0.6	0.827147	0.227032	-0.044027	-0.030385	0.020232
0.7	0.922954	0.019256	-0.113058	-0.051821	0.074643
0.8	0.905181	-0.235214	-0.024464	0.023404	0.060776
0.9	0.738436	-0.382093	0.203178	-0.114802	0.082563

Fig. 5 shows the gain in information transfer due to $U_{max}^{(5)}$ for a continuum of Ω/Ω_o values. Comparison is made between $h_{max}^{(5)}(\Omega/\Omega_o)$ and $h_o(\Omega/\Omega_o)$, information transfer due to uncoated, diffraction-limited optics. The latter is computed by substitution of Eq. (49) into Eq. (39), with subsequent 10-point Gauss quadrature. We see that the main advantage in information relay due to $U_{max}^{(5)}$ occurs for $0 < \Omega/\Omega_o \lesssim 0.5$.

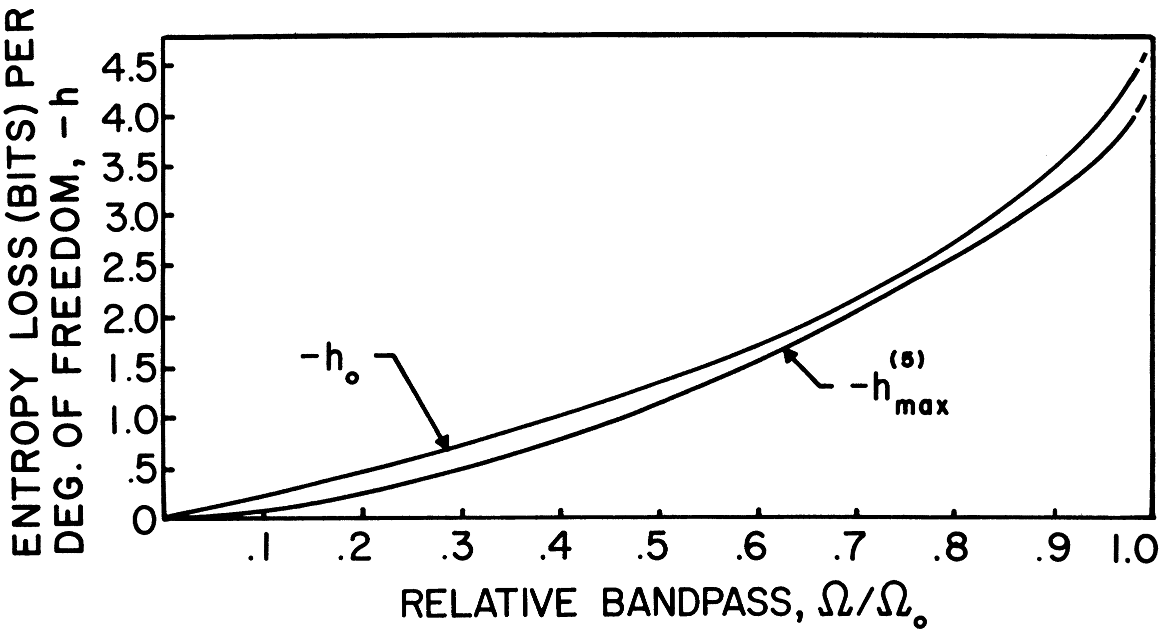

Fig. 5. In answer to the question posed by the title, we show the curve of minimum information loss per degree of freedom, $-h_{max}^{(5)}(\Omega/\Omega_o)$. *This is compared with* $-h_o(\Omega/\Omega_o)$, *information loss per degree of freedom due to uncoated, diffraction-limited optics. For* Ω/Ω_o *in the range 0.1 through 0.9,* $h_{max}^{(5)} = h_{max}$, *the absolute optimum value, to better than 0.5% accuracy. The advantage in information relay of* h_{max} *over* h_o *is seen to be mainly for* $\Omega/\Omega_o \leq 0.5$.

Fig. 6 shows the variation of relative energy and Strehl intensity with Ω/Ω_o. These show a more serious loss of light than their 1-D counterparts, except for the higher peaks in the region $0.7 \leq \Omega/\Omega_o \leq 0.8$.

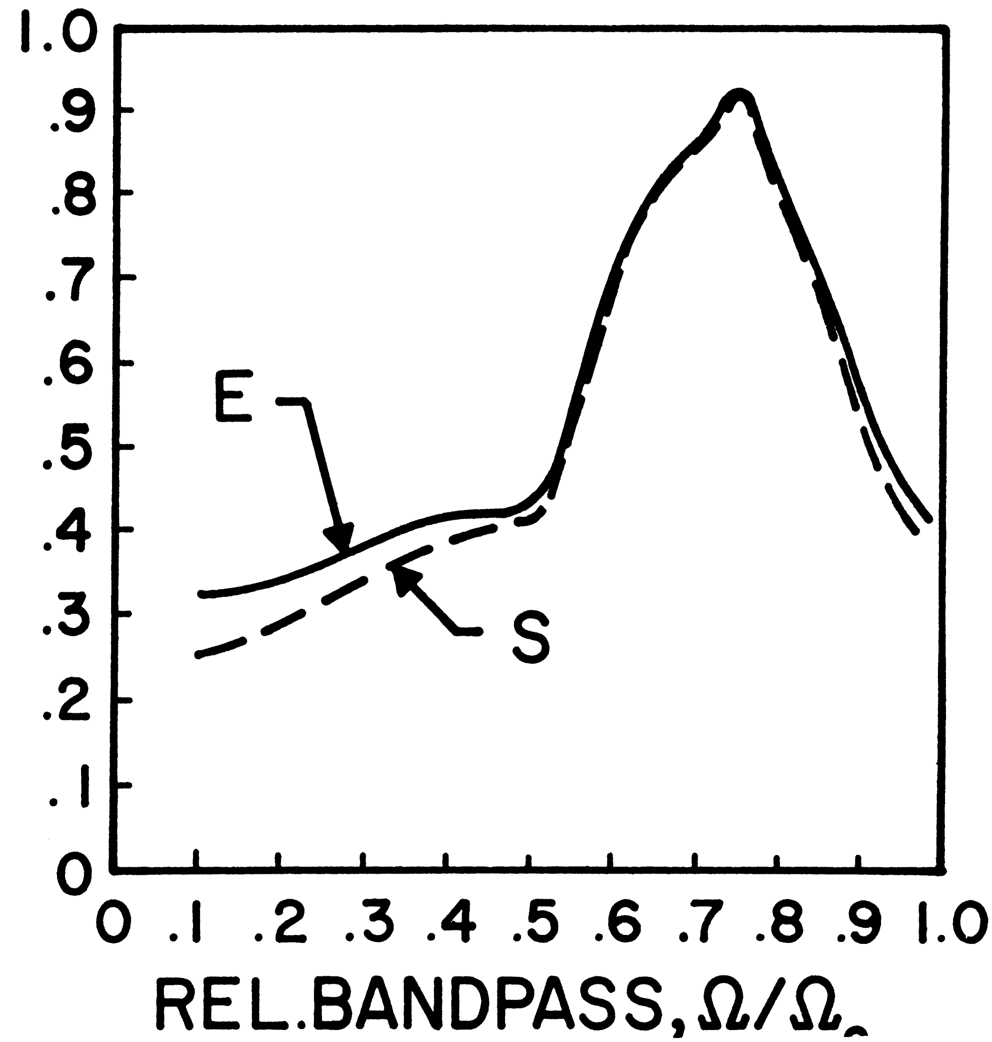

Fig. 6. Total relative energy E and Strehl intensity ratio S resulting from use of coatings $U^{(5)}_{max}(\rho)$ *at* $0.1 \le \Omega/\Omega_0 \le 0.9$. *Parameters E and S are observed to be smallest for* $\Omega/\Omega_0 \lesssim 0.5$, *where the advantage of* $h^{(5)}_{max}$ *over* h_0 *is the greatest (see Fig. 5). The curves of E and S indicate a moderate light loss owing to the optimum coatings.*

THE CONCEPT OF "ZERO INFORMATION TRANSFER" AND ITS RELATION TO SYSTEM RESOLUTION

We may use the curves of Fig. 2 or Fig. 5 to quantitatively relate the relatively new concept of "information transfer" to that of "resolving power," the latter being a traditional measure of system performance. More precisely, we show that the condition of *zero* information transfer to the image space occurs when the independent degrees of freedom (i.e., the sources of information) in the object are spaced closer than a fundamental length Δx which is proportional to the classical resolution limit of the optics. The proportionality constant varies in an inverse way with the amount of information available at each degree of freedom.

Let the optics have bandpass Ω_o. Suppose the object $o(x)$ is bandwidth-limited to region $|\omega| \leq \Omega$ (e.g., $o(x)$ might be an image). Then $o(x)$ obeys a sampling theorem,

$$o(x) = \sum_{n=-\infty}^{\infty} o(n\pi/\Omega) \text{ sinc } (\Omega x - n\pi) \tag{50a}$$

where points $n\pi/\Omega$ locate the independent degrees of freedom. The spacing is

$$\Delta x = \pi/\Omega \quad . \tag{50b}$$

Assume $o(x)$ is practically zero for $|x| > X$. Then

$$\mathcal{O}(\omega) = \sum_{n=-\infty}^{\infty} \mathcal{O}(n\pi/X) \text{ sinc } (X\omega - n\pi) \quad . \tag{51}$$

Hence, object frequency space is also determined by a subdivision of independent degrees of freedom, $\mathcal{O}(n\pi/X)$.

Now, suppose each value of $o(n\pi/\Omega)$ has only two possible values, and that these are equally likely--the case of "binary transmission." Then, by definition (1) there is

$$1 \text{ bit of information/degree of freedom} \tag{52}$$

in x-space. But, the information/degree of freedom is the same in ω-space as in x-space.[18] Hence, (52) describes the information density in frequency space as well.

We now pose the question, "What is the maximum possible object bandpass, specified by ratio Ω/Ω_o, for which the image receives some information?" The curve of $h_{max}^{(5)}(\Omega/\Omega_o)$ in Fig. 2 may be used to find the answer. Observing the abscissa for which $h_{max}^{(5)} = -1$ bit/degree of freedom, we have

$$\Omega/\Omega_o = 0.573 \quad . \tag{53}$$

Noting the definition of h_{max}, (53) is the largest value of Ω/Ω_o allowed by all possible 1-D optical systems. We now use (50b) and the well-known relation

$$\Omega_o = 2\pi/\lambda f\# \tag{54}$$

(λ is the wavelength of light, f# is the optical f/number) in (53) to find the minimum allowed spacing Δx_{min} of the object information:

$$\Delta x_{min} = 0.87\lambda f\# \quad . \tag{55}$$

This is almost precisely $\lambda f\#$, the Rayleigh resolution limit. To summarize, if binary object information is spaced any closer than 0.87 times the Rayleigh resolution limit, no information will be received in the image. This figure assumes the use of the optimal lens system (for information relay); therefore for any other lens system Δx_{min} must exceed $0.87\lambda f\#$.

The analysis may be carried through for any amount h of information/ degree of freedom in the object, a practical limitation being the finite extent of the curve $h_{max}^{(5)}$ in Fig. 2. For example, the minimal Δx for an h-value of 2 bits per degree of freedom is $0.57\lambda f\#$. This illustrates the general trend--as h increases, Δx_{min} decreases--which is intuitively expected.

CONCLUSIONS

Of all uncoated optical systems, the diffraction-limited system relays the most information. The information relay can be further increased by coating the diffraction-limited system in an optimal way. The resulting gain in transmitted information is especially large if the information is contained within a small bandpass region (relative to the optical cutoff frequency). For example, the gain exceeds 100% of the uncoated value when the bandpass region is less than 20% of the system cutoff frequency.

The optimum pupil coatings may be qualitatively described as follows. They are real, and therefore purely of absorption type. In the 1-D case they are even functions of the pupil coordinate, and in the 2-D case of a circular bandpass region they are radially symmetric. When the information bandpass region is small, they are apodizers; when large, they are super-resolvers. They are generally smooth functions of the pupil coordinate, and therefore seem to pose no unusual problems of fabrication.

The amount of coating required for the 1-D optimum pupil is relatively minimal. For example, the total energy passing through the 1-D pupil is never less than 0.5 × the incident energy for the cases considered. Also, the Strehl intensity ratio is, for most cases, more than 0.50. On the other hand, for the circular pupil the corresponding figures are both 0.30, which represents a more serious light-loss problem.

Comparing Figs. 2 and 5, we see that for a given bandpass ratio Ω/Ω_o there is generally less loss of information for the 1-D pupil than for the circular pupil. This fact might be exploited when there is choice of the dimensionality of the object.

As an illustrative use of the results, we have shown that a binary object (as defined) cannot have its independent degrees of freedom spaced more closely than 0.87 times the Rayleigh resolution length. At this spacing (and less) no information is transferred to the image.

ACKNOWLEDGMENT

The author wishes to thank R. V. Shack for his valuable comments on the subject.

APPENDIX

Formal solution to the optimization problem by the calculus of variations

It is possible to derive an integral equation involving $U_{max}(\beta)$ and $\tau_{max}(\omega)$. We do this, and then use it to check the independently calculated solution $U^{(15)}_{max}(\beta;\ 0.40)$, $\tau^{(15)}_{max}(\omega;\ 0.40)$. The formal solution for the 2-D problem is also given.

Combining Eqs. (10) and (18),

$$h(\Omega) = \int_0^{\Omega} d\omega \ln \int_{\omega-\beta_o}^{\beta_o} d\beta U(\beta)U(\beta - \omega) - \int_0^{\Omega} d\omega \ln \int_{-\beta_o}^{\beta_o} d\beta [U(\beta)]^2. \qquad \text{(a)}$$

In (a) we neglect factor $2\Omega^{-1}$, as it is immaterial to the procedure, and we use the natural logarithm because of its convenient properties. Finally, we restrict consideration to real functions $U(\beta)$ because we have shown that $U_{max}(\beta)$ must be real.

Let

$$U(\beta) = U_{max}(\beta) + \varepsilon\zeta(\beta) \qquad \text{(b)}$$

where $\zeta(\beta)$ is an arbitrary function and ε is any real number. We substitute (b) into (a), and then supress the subscript "max" for brevity of notation. Then

$$h = h(\varepsilon) = \int_0^{\Omega} d\omega \ln \int_{\omega-\beta_o}^{\beta_o} d\beta [U(\beta) + \varepsilon\zeta(\beta)] [U(\beta - \omega) + \varepsilon\zeta(\beta - \omega)] - \int_0^{\Omega} d\omega \ln \int_{-\beta_o}^{\beta_o} d\beta [U(\beta) + \varepsilon\zeta(\beta)]^2 . \qquad \text{(c)}$$

Because $U(\beta)$ maximizes h by hypothesis, the functional dependence $h(\varepsilon)$ must attain its maximum at $\varepsilon = 0$. Then

$$\partial h(\varepsilon)/\partial\varepsilon \Big|_{\varepsilon=0} = 0. \tag{d}$$

Enforcing condition (d) upon (c),

$$\int_0^{\Omega} d\omega[\tau(\omega)]^{-1} \int_{\omega-\beta_o}^{\beta_o} d\beta[\zeta(\beta)U(\beta - \omega) + \zeta(\beta - \omega)U(\beta)] - 2\int_0^{\Omega} d\omega \int_{-\beta_o}^{\beta_o} d\beta\zeta(\beta)U(\beta) = 0 \quad , \tag{e}$$

where equality (10) was used once more.

We now interchange integration limits in (e). The following identities may be established in a straightforward manner:

$$\int_0^{\Omega} d\omega \int_{\omega-\beta_o}^{\beta_o} d\beta = \int_{-\beta_o}^{\beta_o} d\beta \int_0^{\{\beta+\beta_o,\Omega\}} d\omega \tag{f}$$

where $\{a,b\} \equiv$ the smaller of a,b in general;

$$\int_{\omega-\beta_o}^{\beta_o} d\beta\zeta(\beta - \omega)U(\beta) = \int_{-\beta_o}^{\beta_o-\omega} d\beta\zeta(\beta)U(\beta + \omega) \tag{g}$$

and

$$\int_0^{\Omega} d\omega \int_{-\beta_o}^{\beta_o-\omega} = \int_{-\beta_o}^{\beta_o} d\beta \int_0^{\{\beta_o-\beta,\Omega\}} \tag{h}$$

The use of (f), (g) and (h) in (e) leads to the required form

$$\int_{-\beta_o}^{\beta_o} d\beta\zeta(\beta) \left\{ \int_0^{\{\beta+\beta_o,\Omega\}} d\omega[\tau(\omega)]^{-1}U(\beta - \omega) + \int_0^{\{\beta_o-\beta,\Omega\}} d\omega[\tau(\omega)]^{-1}U(\beta+\omega) - 2\Omega U(\beta) \right\} = 0 \quad . \tag{i}$$

By the fundamental theorem of the calculus of variations,[19] since $\zeta(\beta)$ is arbitrary, the entire factor of $\zeta(\beta)$ must be zero at all β. This yields the final result,

$$\int_0^{\{\beta+\beta_o,\Omega\}} d\omega \frac{U(\beta - \omega)}{\tau(\omega)} + \int_0^{\{\beta_o-\beta,\Omega\}} d\omega \frac{U(\beta + \omega)}{\tau(\omega)} = 2\Omega U(\beta) \quad . \tag{j}$$

We note that although Eq. (10) may be substituted into (j) in order to arrive at an equation in but one unknown function--$U(\beta)$--the resulting expression is sufficiently complicated to encourage recourse to other methods of solution.

Eq. (j) was used to check many of the solutions $U_{max}^{(M+1)}(\beta)$, $\tau_{max}^{(M+1)}(\omega)$ discussed previously. As would be expected, the agreement between the left- and right-hand sides of (j) improves as M increases. Table 3 (next page) shows a typical result of these checks, the case $\Omega/\Omega_o = 0.4$, $M + 1 = 15$. Because of the good agreement at large M, the assumption of evenness implicit in functions $U_{max}^{(M+1)}(\beta)$ is probably correct.

In the case of statistically random, two-dimensional information, where the optimum pupil is radially symmetric, a derivation similar to the previous one results in the condition

$$\int_0^{\{\rho+\rho_o,\Omega\}} d\omega \frac{\omega U(|\rho - \omega|)}{\tau(\omega)} = \frac{\Omega^2}{2} U(\rho) \tag{k}$$

on the optimum functions $U(\rho)$ and $\tau(\omega)$.

Table 3. A check on the solution $U_{max}^{(15)}(\beta)$, $\tau_{max}^{(15)}(\omega)$, $\Omega/\Omega_o = 0.4$ by substitution into Eq. (j)

β/β_o	Left-hand side	Right-hand side
0.0	1.288	1.292
0.1	1.283	1.281
0.2	1.265	1.258
0.3	1.182	1.184
0.4	1.102	1.102
0.5	1.020	1.019
0.6	0.935	0.935
0.7	0.848	0.851
0.8	0.761	0.756
0.9	0.673	0.676

REFERENCES

1. P. M. Duffieux, *L'integrale de Fourier et ses applications à l'optique*, (Besançon, privately printed, 1947).

2. E. L. O'Neill, *Introduction to Statistical Optics* (Reading, Mass., Addison-Wesley Publ. Co., 1963).

3. C. E. Shannon, Bell System Tech. J. 27:379 and 623 (1948).

4. P. B. Fellgett and E. H. Linfoot, ~~Proc.~~ PHIL. TRANS. Roy. Soc. 247A:369-407 (1955).

5. This is normally called the "optical transfer function." We substitute "contrast" to emphasize that $\tau(\omega)$ represents the transfer of contrast as distinguished from $h(\Omega)$, which represents transfer of information.

6. Stanford Goldman, *Information Theory* (New York, Prentice-Hall, 1955).

7. Ref. 6, p. 145.

8. P. Jacquinot and B. Roizen-Dossier, in *Progress in Optics*, ed. by E. Wolf (Amsterdam, North-Holland Publ. Co., 1964), Vol. III, p. 32.

9. Ref. 8, p. 122.

10. Brigitte Dossier, Revue d'Optique 33:67-70 (1954).

11. As, for example, in solution of the famous "Luneberg apodization problems," republished in R. K. Luneberg, *Mathematical Theory of Optics* (Berkeley, Univ. of California Press, 1964).

12. This general approach was suggested in ref. 8.

13. See, for example, F. B. Hildebrand, *Introduction to Numerical Analysis* (New York, McGraw-Hill Book Co., 1956), p. 447, where it is called the Newton-Raphson method.

14. See, for example, R. Courant, *Differential and Integral Calculus* (London, Blackie and Son, Ltd., 1942), Vol. II, pp. 188-199.

15. Ref. 13, p. 320, Eq. (8.4.6).

16. R. V. Churchill, *Fourier Series and Boundary Value Problems* (New York, McGraw-Hill Book Co., 1941), p. 164.

17. R. Barakat, J. Opt. Soc. Am. 54(7):920-930 (July 1964).

18. As may be verified from Eqs. (50a) and (51), the total number of degrees of freedom in x-space equals the total number in ω-space. On the other hand, in ref. 6, p. 141, it is noted that the total information is the same in x-space as in ω-space. Therefore the information per degree of freedom is the same in both spaces.

19. H. Lass, *Elements of Pure and Applied Mathematics* (New York, McGraw-Hill Book Co., 1957), p. 291.

jal