AFCRL-67-0070

NUMERICAL RESTORATION OF OPTICAL OBJECTS OBSCURED BY DIFFRACTION AND NOISE

Richard W. Harris
Craig K. Rushforth

Electro-Dynamics Laboratories
Departments of Electrical Engineering, Physics, and Chemistry
UTAH STATE UNIVERSITY
Logan, Utah

Contract No. AF19(628)-3825
Project No. 8663

SCIENTIFIC REPORT NO. 8

31 December 1966

Dr. G. A. Vanasse

This research was sponsored by the Advanced Research Projects Agency under ARPA Order No. 450

Prepared for
Air Force Cambridge Research Laboratories
Office of Aerospace Research
UNITED STATES AIR FORCE
Bedford, Massachusetts

AFCRL-67-0070

NUMERICAL RESTORATION OF OPTICAL OBJECTS OBSCURED BY DIFFRACTION AND NOISE

Richard W. Harris
Craig K. Rushforth

Electro-Dynamics Laboratories
Departments of Electrical Engineering, Physics, and Chemistry
UTAH STATE UNIVERSITY
Logan, Utah

Contract No. AF19(628)-3825

Project No. 8663

SCIENTIFIC REPORT NO. 8

31 December 1966

Dr. G.A. Vanasse

This research was sponsored by the Advanced Research Projects Agency
under ARPA Order No. 450

Prepared for

Air Force Cambridge Research Laboratories
Office of Aerospace Research
UNITED STATES AIR FORCE
Bedford, Massachusetts

ABSTRACT

The restoration of incoherent optical objects which have been diffracted by an optical system and corrupted by detector and additive background noise is considered. The approach to the problem is basically numerical and considers operating directly on the noisy image and point spread function rather than the Fourier transform of these quantities. The effects of noise and the use of a priori information in the restoration process are given particular attention.

Several "optimum" estimates of the object intensity distribution are considered. Based on statistics which have been verified in practice, the Bayes, maximum a posteriori, maximum likelihood and mean square error estimates of the object intensity distribution are obtained. These statistical estimates are compared mathematically and in many cases numerically to other non-statistical estimates formulated from control theory and dynamic programming. Extensive numerical results have been obtained for the restoration of various one-dimensional objects in the presence of noise. Two monochromatic "point sources" in the presence of noise are shown to be resolved when separated by 1/5 of the Rayleigh criterion distance. Numerical results are also shown for the mean square error as a function of a priori information, the measuring scheme chosen and diffraction.

TABLE OF CONTENTS

LIST OF FIGURES

Figure | Page

Figure Page

Figure Page

INTRODUCTION

Statement of the Problem

All objects which are inputs to a physically realizable optical imaging system emerge from the system as imperfect representations of the original object. This object corruption may be due to several effects inherent in the optical system, such as diffraction and aberration, as well as background and detector noise. In fact, the effects of diffraction alone can severely limit system resolution and effectively destroy, insofar as the observer is concerned, the original object detail. These effects are so pronounced that in the past, as Toraldo di Francia (1952) points out, the classical Rayleigh resolution limit has been accepted as a theoretical limit. However, in 1955 Toraldo suggested that a priori information about the object could be used to alleviate the apparent theoretical limits imposed by diffraction.

In 1964 J. L. Harris showed, by using some well-known results, that in the absence of noise a priori knowledge that the object is of finite extent is sufficient to enable the exact restoration of a diffracted optical object. This work theoretically establishes that object restoration is limited only by system noise.

With this background we state that the problem treated in this paper is that of restoring an optical object which has been diffracted and corrupted by background and detector noise. Particular emphasis is placed on the trade-off between restoration detail and the limiting noises.

Approaches to the Problem

Although a unique theoretical solution has been shown to exist, the application of a working restoration procedure which performs well for a variety of objects and optical configurations is quite another matter. Basically, three approaches have been considered in solving this problem.

One approach has been to utilize the Fourier transform relationships that exist between the object and image. Harris (1966), Osterberg (1966) and Wolter (1961) have considered this approach. However, this technique appears to be difficult to analyze when applied to the noisy observed quantities, and as yet no results appear to be available which use Fourier techniques in such an environment.

A second approach, used by Barnes (1966a), is one of analytically solving for the object by using the eigenfunctions of an integral operator involving an analytical point spread function. The drawbacks of this approach, from a practical point of view, are: (1) the extreme difficulty, at present, in obtaining eigenfunctions for a general point spread function

(one that is measured), and (2) the fact that, to date, system noise has not been included in the analytical restoration process.

The third approach which is developed in this paper is that of using numeric . methods to perform the restoration. We reiterate the practical advantages of this procedure over those above. First, in an actual real-world situation only discrete quantities are available to operate on and effect a solution; i.e., the point spread function and image are usually not analytically known and so must be discretely measured and then operated on to perform the restoration. Secon ', the observed quantities are not deterministic but must be trea ed as random variables. This randomness may be due to additive background fluctuations, film granularity, photomultiplier multiplicative noise or some other detector noise. These effects can be taken into account by using numerical techniques and appealing to the statistically optimum minimum mean square error (MSE) estimation procedure.

Plan of the Paper

The first section presents a review of the development of the image-object relationships and presents the integral equation that is to be solved to effect the restoration. Emphasis is placed on reviewing the assumptions involved in developing the imaging equation and in discussing the concept of diffraction and how it is to be varied in this

paper. Emphasis is also placed on the transfer of the continuous version of the imaging equation to its discrete representation so that numerical methods can be used in its solution.

As a preview to the numerical work that will follow later, the second section after the introduction reviews previous numerical work in the solution of the basic integral equation developed in the first section. Also presented in this section is a dynamic programming solution, developed by the author, which utilizes an "area" constraint in the solution.

The third section treats two analytical approaches which assess the difficulty in regaining object information which has been obscured due to diffraction and noise. Particular emphasis is placed on the trade-off between SNR and diffraction in obtaining this information.

The fourth section discusses the noise models to be used and presents the development of the minimum MSE object estimate. This estimate and the computational features discussed in the last two sections constitute the basic restoration procedure as presented in this paper.

Results of performing actual simulated restorations and of the study of errors involved in using the MSE estimate are presented in the last two sections. Here again, as in the section on analytical results, particular emphasis has been placed on the effects of noise and diffraction in restoring object detail. In addition, emphasis has been placed on

presenting actual simulations for various noise levels so that one may visually judge the restoration improvement. These simulations augment the important but somewhat abstract results obtained in the analytical section and serve to give one a more complete picture of the restoration procedure.

THE IMAGING EQUATION

In this section we derive and discuss the relationship between the radiation from an incoherent source intensity distribution and the corresponding image intensity distribution. Near the end of the section the imaging equation in both continuous and discrete form is stated. In order to clarify what is meant by diffraction as it applies to the restoration problem, the concept of diffraction and its relationship to the imaging equation will be discussed. With this background the reader should be able to more clearly assess what is meant by object restoration in the presence of diffraction.

The Optical Configuration

The optical configuration to be used in the derivation is shown in Figure 1. However, it should be mentioned that the imaging equation developed here with the subsidiary definitions necessary to utilize this equation can be derived from several optical configurations. Two such derivations are mentioned below.

Helstrom (1964) derives this equation with suitable geometric assumptions which make it possible for a lens or absorber to replace the aperture. His approach is quite natural if one wants to develop the

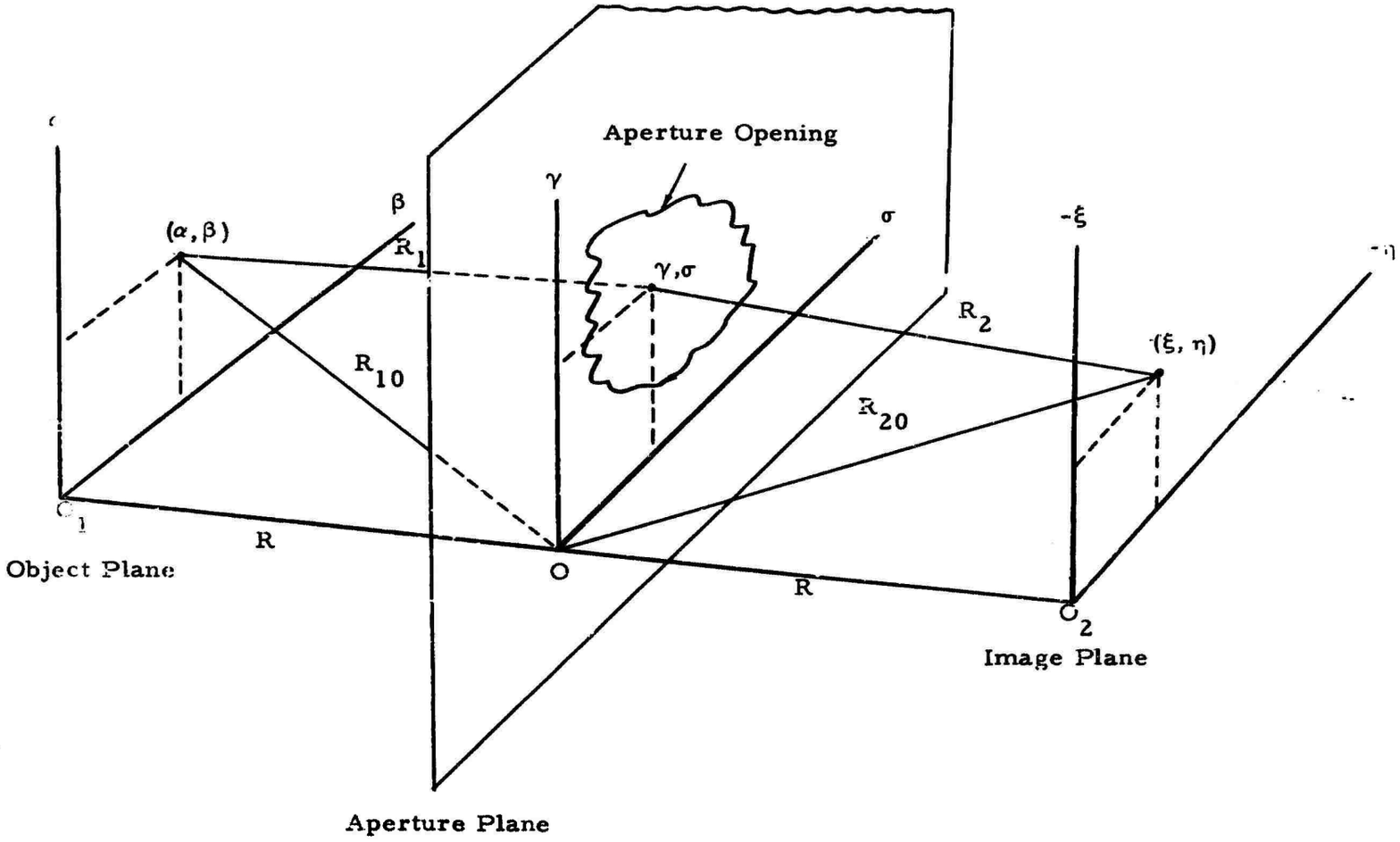

Figure 1. Optical configuration.

practical aspect of an instrument which observes objects at great distances from the aperture.

Cutrona and others (1960) use a configuration composed of appropriately spaced lenses which is often used in the laboratory. Using this system, the only assumptions necessary to derive the imaging equation are that the lenses together serve to image the object plane onto the image plane and that the lenses are aberration free.

The Scalar Theory

To begin with, the disturbance in the object plane is taken as an electromagnetic field at optical frequencies. The complete derivation of the imaging equations by finding appropriate solutions of Maxwell's equations has been accomplished only in a few idealized situations (Stone, 1963). However, an approximate theory has been developed which allows one to solve the imaging relations for a variety of aperture shapes. The approximate theory treats the electromagnetic disturbance as a scalar field and neglects effects due to polarization. Fortunately, the imaging equations developed from this theory predict the associated distributions to high accuracy when certain restrictive although not impractical assumptions are made.

The wave equation, which the amplitude of the wave in the scalar theory must obey, is

$$\nabla^2 u + \frac{1}{c^2} \frac{\partial^2 u}{\partial t^2} = 0 \qquad (1)$$

where u is the scalar amplitude of the wave and c is the free space velocity of light. The monochromatic (wavelength λ) solution of this equation in spherical coordinates may be written

$$u(R, t) = \frac{a}{R} \cos [k(R-ct) + \theta] \qquad (2)$$

where a is the constant fixing the amplitude, R is the spherical radial variable, $k = 2\pi/\lambda$, and θ is a phase angle. This solution represents the wave propagated from a point in the object plane.

It is desired to consider many such points located at (α, β) in the object plane, and so we write for the object amplitude distribution

$$u(R, \alpha, \beta, t) = c\, R_e [A(\alpha, \beta, t) e^{ikR}] \qquad (3)$$

where $A(\alpha, \beta, t)$ is the complex amplitude, R_e means "real part of," and c is a constant which takes the inverse distance relationship into account.

Here we make the assumption as discussed by Helstrom (1964) that the observation time is to be small compared with the reciprocal bandwidth of the radiation and that no information regarding the instantaneous time dependence will be used. Dropping the time dependence and for our purposes the unimportant constant c, we have

$$u(R,\alpha,\beta) = A(\alpha,\beta)e^{ikR}. \tag{4}$$

Use of Huygen's Principle and the Geometric Assumptions

The Fourier transform relations which will be developed are essentially those established by Huygen, Fresnel, and Kirchoff. For a more detailed treatment, the reader is referred to Stone (1963), O'Neill (1963), Born and Wolf (1959), and Beran and Parrent (1964).

Consider a wave due to a point source in the object plane which propagates toward the aperture. In essence, Huygen's principle states that once this wave has propagated to points which are just through the aperture opening, then these points themselves may be regarded as sources of secondary wavelets that again propagate with the usual solution (2) for a point source. The disturbances due to all of these "secondary" sources which reach the image plane then may be appropriately summed to obtain the image plane distribution. Finally, to obtain the image distribution due to several point sources or a continuum of point sources

in the object plane we merely sum or integrate over the object plane coordinates.

Following the above procedure we seek to find the amplitude at an arbitrary point (γ,σ) just to the right of the aperture opening. At a point to the left of the aperture

$$u(R_1,\alpha,\beta) = A(\alpha,\beta)e^{ikR_1}. \tag{5}$$

Dealing only with the exponential,

$$e^{ikR_1} = e^{ikR_{10}}\, e^{ik(R_1-R_{10})}. \tag{6}$$

Referring to Figure 1 the length R from O_1 to O can be written in two ways:

$$R = R_1^2 - (\gamma-\alpha)^2 - (\sigma-\beta)^2 = R_{10}^2 - \alpha^2-\beta^2, \tag{7}$$

also

$$R_1^2 - R_{10}^2 = (R_1-R_{10})(R_1 + R_{10}) = \gamma^2 + \sigma^2-2(\alpha\gamma + \sigma\beta). \tag{8}$$

Thus

$$R_1 - R_{10} \approx \frac{\gamma^2 + \sigma^2}{R_1 + R_{10}} - \frac{2(\gamma\alpha + \sigma\beta)}{R_1 + R_{10}}. \qquad (9)$$

We now disregard the squared term in the exponential on the grounds that the Fraunhofer far-field conditions hold and that the open portion of the aperture plane is small. However, we still make the assumption that the dimensions of the aperture are many wavelengths, an assumption which is made in order that the scalar wave theory will agree with experimental results (Stone, 1963). We also assume that the source is sufficiently far removed from the aperture so that $R_1 + R_{10} \approx 2R$ in the denominator. Thus, the electric field just to the left of the aperture is

$$A(\alpha,\beta)e^{ikR_{10}}e^{-ik\frac{(\gamma\alpha + \sigma\beta)}{R}}. \qquad (10)$$

In order to simplify the above expression (10) the following spatial frequency components are defined:

$$f_\gamma = \frac{\gamma}{\lambda R}, \qquad (11)$$

and

$$f_\sigma = \frac{\sigma}{\lambda R}.$$

Expression (10) becomes

$$A(\alpha,\beta)e^{ikR_{10}}e^{-i2\pi(f_\gamma\alpha + f_\sigma\beta)}. \qquad (12)$$

We now define a transfer function $G(f_\gamma, f_\sigma)$ and obtain

$$v_{\alpha,\beta}(f_\gamma, f_\sigma) = G(f_\gamma, f_\sigma)\, A(\alpha,\beta)\, e^{ikR_{10}}e^{-i2\pi(f_\gamma\alpha + f_\sigma\beta)} \qquad (13)$$

for the electric field at a point (γ,σ) or (f_γ, f_σ) just to the right of the aperture. The magnitude $G(f_\gamma, f_\sigma)$ cannot be greater than unity, for the aperture cannot amplify the incoming field (Helstrom, 1964). We also define the transfer function so that it is zero for points outside the open portion of the aperture plane.

At this point we apply Huygen's principle and treat each point on the right side of the aperture as a "secondary" source. Following the same reasoning used in arriving at the disturbance at the left of the aperture plane due to a point source in the object plane, the electric field in the image is

$$v_{\alpha,\beta}(f_\gamma, f_\sigma)\, e^{ikR_{20}}e^{i2\pi(f_\gamma\xi + f_\sigma\eta)} \qquad (14)$$

where again we have assumed the Fraunhofer conditions.

To obtain the total electric field due to a point source in the object plane we must integrate over all of the "secondary" point sources in the aperture plane. Thus, the electric field in the image plane due to a point source in the object plane is

$$y_{\alpha,\beta}(\xi,\eta) = e^{ikR_{20}} \iint_{-\infty}^{\infty} v_{\alpha,\beta}(f_\gamma, f_\sigma)\, e^{i2\pi(f_\gamma \xi + f_\sigma \eta)}\, df_\gamma df_\sigma. \qquad (15)$$

Substituting for $v_{\alpha,\beta}(f_\gamma, f_\beta)$

$$y_{\alpha,\beta}(\xi,\eta) = e^{ik(R_{10} + R_{20})} \qquad (16)$$

$$\cdot \iint_{-\infty}^{\infty} A(\alpha,\beta)\, G(f_\gamma, f_\sigma)\, e^{i2\pi[f_\gamma(\xi-\alpha) + f_\sigma(\eta-\beta)]}\, df_\gamma df_\sigma.$$

Thus far we have dealt entirely with the amplitude of the electromagnetic field. However, throughout the derivation that follows it will be assumed that we are dealing with incoherent light and that the measured quantities are intensities, which add linearly. Coherent light could also be assumed throughout and the form of the imaging equation would be the same as that derived for the incoherent case. In this respect the numerical techniques used could be applied to either coherent or incoherent illumination, but the simulated results in the last section apply only to the incoherent case.

The intensity (watts) at (ξ, η) due to a point source at (α, β) is

$$b(\alpha, \beta, \xi, \eta) = y_{\alpha,\beta}(\xi, \eta)\, y^{*}_{\alpha,\beta}(\xi, \eta) = \left| y_{\alpha,\beta}(\xi, \eta) \right|^2 \qquad (17)$$

where the star as used here denotes complex conjugate. Performing this multiplication

$$b(\alpha, \beta, \xi, \eta) = \int_{-\infty}^{\infty} \left| A(\alpha, \beta) \right|^2 G(f_\gamma, f_\sigma)\, G(f'_\gamma, f'_\sigma) \cdot e^{-i2\pi [\alpha (f_\gamma - f'_\gamma) + \beta (f_\sigma - f'_\sigma)]} \cdot e^{+i2\pi [\xi (f_\gamma - f'_\gamma) + \eta (f_\sigma - f'_\sigma)]} \, df_\gamma\, df'_\gamma\, df_\sigma\, df'_\sigma . \qquad (18)$$

To obtain the total intensity due to a continuum of point sources in the object plane we integrate over the object plane. Thus

$$b(\xi, \eta) = \int_{-\infty}^{\infty} b(\alpha, \beta, \xi, \eta)\, d\alpha\, d\beta , \qquad (19)$$

and using (18)

$$b(\xi, \eta) = \int_{-\infty}^{\infty} G(f_\gamma, f_\sigma)\, G^{*}(f'_\gamma, f'_\sigma)\, e^{i2\pi [\xi (f_\gamma - f'_\gamma) + \eta (f_\sigma - f'_\sigma)]} \cdot \left[\int_{-\infty}^{\infty} \left| A(\alpha, \beta) \right|^2 e^{-i2\pi [\alpha (f_\gamma - f'_\gamma) + \beta (f_\sigma - f'_\sigma)]} \, d\alpha\, d\beta \right] df_\gamma\, df'_\gamma\, df_\sigma\, df'_\sigma . \qquad (20)$$

The integral in brackets is defined as the spatial frequency spectrum of the object intensity distribution. The quantity $|A(\alpha,\beta)|^2$ is defined as the object intensity distribution and will be denoted $x(\alpha,\beta)$. The integral in brackets is also the two-dimensional Fourier transform of $x(\alpha,\beta)$. We denote it as $X(f_\gamma - f'_\gamma,\ f_\sigma - f'_\sigma)$. In general, capital letters will denote the spatial frequency transform of the analogous function in lower case letters.

To facilitate further simplification let $t_\gamma = f_\gamma - f'_\gamma$ and $t_\sigma = f_\sigma - f'_\sigma$. Thus $f'_\gamma = f_\gamma - t_\gamma$, $df'_\gamma = -dt_\gamma$, and $f'_\sigma = f_\sigma - t_\sigma$, $df'_\sigma = -dt_\sigma$, and

$$b(\xi,\eta) = \iint_{-\infty}^{\infty} X(t_\gamma, t_\sigma)\, e^{i2\pi[\xi t_\gamma + \eta t_\sigma]} \cdot \left[\iint_{-\infty}^{\infty} G(f_\gamma, f_\sigma)\, G^*(f_\gamma - t_\gamma,\ f_\sigma - t_\sigma)\, df_\gamma df_\sigma \right] dt_\gamma dt_\sigma. \tag{21}$$

The integral in brackets above is defined as the Fourier transform of the point spread function. It is to be denoted as $H(t_\gamma, t_\sigma)$. It will be assumed throughout for the configuration considered that this same point spread function applies for all point sources in the object plane. This is the spatial invariance assumption one uses in an "idealized" optical configuration (O'Neill, 1963). With this definition

$$b(\xi,\eta) = \iint_{-\infty}^{\infty} X(t_\gamma, t_\sigma)\, H(t_\gamma, t_\sigma) e^{i2\pi[\xi t_\gamma + \eta t_\sigma]}\, dt_\gamma dt. \tag{22}$$

Taking the Fourier transform of both side of the above equation,

$$B(f_\gamma, f_\sigma) = H(f_\gamma, f_\sigma)\, X(f_\gamma, f_\sigma). \tag{23}$$

By the convolution theorem

$$b(\xi, \eta) = \int_{-\infty}^{\infty}\int h(\xi-\alpha,\ \eta-\beta)\, x(\alpha,\beta)\, d\alpha d\beta. \tag{24}$$

The one-dimensional version of (24) is

$$b(\xi) = \int_{-\infty}^{\infty} h(\xi-\alpha)\, x(\alpha) d\alpha. \tag{25}$$

Since it is assumed that the object is of finite spatial extent, Equations (24) and (25) may be rewritten as

$$b(\xi, \eta) = \int_{R_2}\int h(\xi-\alpha,\ \eta-\beta)\, x(\alpha,\beta) d\alpha d\beta, \tag{26}$$

$$b(\xi) = \int_{-a}^{a} h(\xi-\alpha)\, x(\alpha) d\alpha. \tag{27}$$

where R_2 denotes the region in two-dimensional object space where $x(\alpha,\beta)$ is nonzero.

It is our objective to solve the integral Equations (26) and (27) for the object when the image and point spread function are specified. These equations represent the diffraction of an object by an optical system and are known as the one and two dimensional versions of the imaging equation. Of course, it is obvious to this point that no noise is represented in Equations (26) and (27). Noise effects are treated in a later section.

The Concept of Diffraction

When the point spread function is much larger in spatial extent than the object, then the image is nearly equal to the point spread function. This is a case of large diffraction, and the image bears little resemblance to the original object. Alternatively, when the point spread function is small spatially in comparison with the object the image is nearly equal to the object and the diffraction effects are almost negligible.

In order to provide a measure of diffraction we define

$$R \triangleq \frac{\text{Point spread extent}}{\text{Object extent}} \tag{28}$$

where by point spread extent we refer to the "Airy disk" of the point spread function and by object extent we refer to a measure of the outside size of the object (where the object is nonzero); i.e., the size of a two-point source object is taken as the separation between the points.

The ratio R may be varied in two ways. We can either fix the aperture and vary the object or fix the object and vary the aperture. Both methods of varying R produce the same qualitative results. In this paper we have considered, so far as the simulated results are concerned, a fixed optical system which views various objects. This is generally the case treated in practice since the cause of a "diffraction limited system" is usually due to the fact that either the aperture was fixed when the measurements were made or, on realizing that diffraction was severe, the aperture was opened as far as possible but diffraction was still apparent.

Since we are presenting results for a fixed optical system that views various objects, it was convenient to fix the aperture in reference with the classical Rayleigh criterion for resolving two point sources. The Rayleigh criterion resolution distance was set at unity, and thus

$$R = \frac{2.0}{\text{object size}} \tag{29}$$

and R = 2.0 corresponds to the spacing of two point sources which Rayleigh proposed were resolved. However, as we shall demonstrate, two sources can be resolved even in the presence of noise when R = 10.0.

Now that the concept of diffraction has been reviewed and put in context with regard to the restoration problem, we consider next the

presentation of the numerical technique to be used in solving the imaging equation.

The Numerical Technique

We desire to solve the imaging equation by a numerical technique. Thus, we are led to consider approximate discrete representations of the various functions involved. The one-dimensional version (27) for a single point ξ_i may be written as

$$b_i = \sum_{i=1}^{m} (h_{ij} w_j) x_j + \epsilon_i \qquad (30)$$

where $b_i = b(\xi_i)$, $h_{ij} = h(\xi_i, \alpha_j)$, $x_j = x(\alpha_j)$, w_j is the quadrature weight, and ϵ_i is the quadrature error involved in transferring from continuous to discrete quantities.

The quadrature error ϵ_i is very important in the restoration process, particularly if it is assumed that the object, point spread function and image are all continuous functions. The quadrature error can be regarded as depending primarily upon the number of points (M) chosen and the type of quadrature weight (w_i) used. Evidence indicating the proper choice of these quantities for the one-dimensional case is presented in a later section. Thus, at present, we neglect the quadrature error and defer this discussion to a later section.

Generalizing (30) to N image points $\xi_1, \xi_2, \ldots \xi_N$, we have N equations of the form (30) which may be written in matrix form as

$$b = Ax \tag{31}$$

where the elements a_{ij} in the A matrix are defined as

$$a_{ij} = h_{ij} w_j \qquad \begin{matrix} i = 1, \ldots N \\ j = 1, \ldots M \end{matrix}$$

and b and x are, respectively, Nx1 and Mx1 dimensional column vectors. The two-dimensional version (26) may also be represented in the form of (31) by a suitable definition of the points in two dimensional space. This definition is illustrated in the Appendix.

Equation (31) is the discrete numerical representation of Equations (26) and (27) used throughout the paper. Observation of (31) in light of the objective (that of solving for x given A and b), the obvious solution for the case when M = N and A is nonsingular is

$$x = A^{-1} b \tag{32}$$

where A^{-1} is the inverse matrix for A.

However, such a simple solution is essentially of no value when applied to the practical optical problem. The two outstanding difficulties experienced by the author and others (Phillips, 1962; Bellman, 1965) are stated below. First, in arriving at (32) it is assumed that no noise is present in the system. The use of this assumption is very important since excessive accuracy may have to be used in order to obtain the correct solution. Second, as the diffraction becomes appreciable, the A matrix approaches a singular matrix, and from a computational point of view the successful inversion of the A matrix is nearly impossible. Both of these difficulties can be alleviated by taking the noise and the a priori information into account. This fact is demonstrated in the simulated restorations presented in a later section.

Phillips (1962) has also developed a technique based on control theory which overcomes the problems mentioned above. Because of the similarity of his solution, as developed further by Twomey (1963), and the minimum mean square error (MSE) solution, in the next section a review of their approach to the problem is presented. Bellman (1965) has further extended the computational procedure developed by both Phillips and Twomey and has applied dynamic programming to the solution of the problem. Hence, we shall also consider the dynamic programming approach and attempt to evaluate its merits and limitations as it applies to the restoration problem.

NUMERICAL SOLUTIONS AND METHODS IN THE ABSENCE OF NOISE

Review of the Control Theory Approach

Because the final form of the solutions developed by Phillips (1962) and Twomey (1963) is very similar mathematically and conceptually to the MSE solution derived later, a rather extensive review of these papers is presented in this section.

Recall the one-dimensional version of the imaging equation (27). Phillips (1962) mentions that this integral equation of the first kind can be unstable in that infinitesimal changes in $b(\xi)$ can cause large changes in $x(\alpha)$ and that the success in solving this equation by any method depends largely on the accuracy of $b(\xi)$ and the shape of the kernel $h(\xi-\alpha)$.

Since the solution will depend on the accuracy of $b(\xi)$, Phillips suggests that (27) be altered to read

$$b(\xi) + e(\xi) = \int_{-a}^{a} h(\xi-\alpha)x(\alpha)\,d\alpha \qquad (33)$$

where $e(\xi)$ is an unknown arbitrary bounded function. The solution to this equation is not unique, but now we seek to find the best solution from a family of solutions. The "best solution" is referred to in the

sense that further constraints are placed on the solution which enable one to solve the problem.

Phillips introduced a smoothness constraint, which is that the quantity

$$\int_{-\infty}^{\infty} [x''(\xi)]^2 d\xi \tag{34}$$

be small, where $x''(\xi)$ denotes the second derivative of $x(\xi)$. Numerically (34) may be approximated by

$$\sum_{i=1}^{M} (x_{i-1} - 2x_i + x_{i+1})^2. \tag{35}$$

Phillips then considers the function $e(\xi)$ to be bounded in the following way:

$$\int_{-a}^{a} e^2(\xi)\, d\xi \leq E \tag{36}$$

or equivalently

$$\sum_{i=1}^{M} e_i^2 \leq E. \tag{37}$$

Now we can prop. e the following minimization:

$$\min_{\{x_i\}} \sum_{i=1}^{M} (x_{i-1} - 2x_i + x_{i+1})^2, \tag{38}$$

subject to

$$\sum_{i=1}^{N} \left(\sum_{j=1}^{M} a_{ij} x_j - b_i \right)^2 - E = 0. \tag{39}$$

Following the general minimization procedure using Lagrange multipliers the problem posed in (38) and (39) is equivalent to seeking the $\{x_i\}$ which minimize

$$R(x) = \sum_{i=1}^{M} (x_{i-1} - 2x_i + x_{i+1})^2 + \lambda^{-1} \left[\sum_{i=1}^{N} \left(\sum_{j=1}^{M} a_{ij} x_j - b_i \right)^2 - E \right] \tag{40}$$

where λ^{-1} is the Lagrangian multiplier (Courant and Hilbert, 1953; Hildebrand, 1952).

If we continued on with the Lagrangian method we would find partial derivatives of R(x) with respect to the x_i and λ^{-1}. From these equations one then solves for the x_i and λ^{-1}. In this case, however, we cannot solve explicitly for λ^{-1} unless the bound E is specified. Thus, to make the problem more mathematically tractable we consider the quantity λ^{-1} to be known and E to be unknown. This is often done in practice,

and one can reason that the parameter λ^{-1} is to be varied until some specified bound is attained (Bellman, 1962).

When λ^{-1} is known we can equivalently choose x_i to minimize

$$R(x) = \sum_{i=1}^{N} \left(\sum_{j=1}^{M} a_{ij} x_j - b_i\right)^2 + \lambda \sum_{i=1}^{M} \left(x_{i-1} - 2x_i + x_{i+1}\right)^2. \tag{41}$$

Now the second term can be regarded as the constraint, while we seek to minimize the squared error.

Note that if λ is allowed to be negative then $R(x)$ can be made zero regardless of $\{x_i\}$. Thus only non-negative values of λ are to be considered.

To find $\{x_i\}$ we differentiate $R(x)$ with respect to x_k and equate the result to zero. Thus,

$$\frac{\partial R(x)}{\partial x_k} = \sum_{i=1}^{N} a_{ik} \left(\sum_{j=1}^{M} a_{ij} x_j - b_i\right) \tag{42}$$

$$+ \left(x_{k-2} - 4x_{k-1} + 6x_k - 4x_{k+1} + x_{k+2}\right) = 0$$

where $x_o = x_{M+1} = 0$, $k = 3, 4, \ldots M-2$. In matrix form this equation can be written

$$A'Ax - A'b + \lambda Hx = 0 \tag{43}$$

where $H = \begin{vmatrix} 5 & -4 & 1 & 0 & 0 & 0 & . & . \\ -4 & 6 & -4 & 1 & 0 & 0 & . & . \\ 1 & -4 & 6 & -4 & 1 & 0 & . & . \\ 0 & 1 & -4 & 6 & -4 & 1 & . & . \\ . & . & . & . & . & . & . & . \\ . & . & . & . & . & . & . & . \end{vmatrix}$

The prime is used to indicate a matrix or vector transpose. Thus, the minimizing vector x is

$$x = (A'A + \lambda H)^{-1} A'b. \tag{44}$$

This is the result obtained by Twomey (1963) and is a slightly modified form of the original expression developed by Phillips (1962).

From (44) we see that if λ is zero the solution reverts back to the overdetermined pseudoinverse solution (M<N) or the inverse solution (M=N). Values of $\lambda>0$ weight the constraint more heavily and are chosen in accordance with the amount of "smoothing" necessary to control the unstability of the system.

Once the solution vector x is found the individual "errors" (e_i) may be determined from

$$e = Ax - b. \tag{45}$$

These values may be used as a criterion to choose λ (Phillips, 1962). For $\lambda = 0$, the e_i are zero.

The choice of a suitable value for λ may lead one to be rather skeptical about the use of such a parameter. Experience has shown, however, that in applying this technique to the optics problem the ill-conditioning of the original system can lead to solutions which conflict with the a priori knowledge that the incoherent object intensity distribution cannot be negative or exhibit large positive or large positive and negative oscillations. Based on this rather limited prior information, a practical choice of λ can be made.

In practice we can, by obtaining a solution for several values of λ, determine a minimum value λ_{min} so that the solutions do not conflict radically with this a priori knowledge. For a range of values above this minimum, the solutions do not change appreciably, and thus are all "acceptable solutions." With this knowledge and the use of further computational schemes we can from a practical point of view circumvent the arbitrariness in the choice of λ.

Other Constraints

Twomey (1963) and Bellman, et al. (1964, 1965) utilize other constraints which are applicable to their particular problems. Twomey (1963) suggested a constraint which minimizes the sum of squares of the differences between the actual solution vector x and an a priori vector p. In this case we seek to minimize

$$R(x) = \sum_{i=1}^{N} \left(\sum_{j=1}^{M} a_{ij} x_j - b_i \right)^2 + \lambda \sum_{i=1}^{M} (x_i - p_i)^2. \quad (46)$$

The vector solution (obtained in the same straightforward way as before) is

$$x = (A^t A + \lambda I)^{-1} (A^t b + \lambda p) \quad (47)$$

where I is an MxM identity matrix. Twomey states that a general form of the solution when many different constraints are considered is

$$x = (A^t A + \lambda H)^{-1} (A^t b + \lambda p). \quad (48)$$

where H is a matrix and p is a vector, and they are to be specified by the constraint used.

Here we call attention to the fact that (47) and (48) are very similar mathematically to the MSE solutions (149) and (155) which are derived later. Conceptually, the methods used in deriving these equations are also very similar in that the basic motivation in each case was to utilize a priori information in producing a better solution. Further analogies between the MSE and control theory approach are deferred to the Object Estimation Section.

Sequential Approximations

The method of sequential approximations or iteration shown here was first mentioned by Bellman, et al., (1965). It was found to be very useful in the restoration problem and was used extensively in obtaining the simulated restorations presented later. This method was applied to both the MSE solution (15[illegible]) and the control theory solution (47). The method consists simply of continuously replacing the a priori mean or vector by an updated version as shown below:

$$x_n = (A^tA + \lambda I)^{-1} A^tb + (A^tA + \lambda I)^{-1} x_{n-1}. \tag{49}$$

This is a sequence in which x_n approaches the vector

$$x = A^{-1}b \tag{50}$$

for A nonsingular. The convergence of this sequence is seen as follows.

For convenience we define the matrix

$$B = (A^tA + \lambda I)^{-1} \tag{51}$$

and the vector

$$d = A'b. \tag{52}$$

Now for $\lambda>0$ and A nonsingular the eigenvalues of λB are all less than unity (Bellman, et al., 1965). To show that the sequence (49) converges, we note that it is a Cauchy sequence

$$\begin{aligned} x_n - x_{n-1} &= \lambda B(x_{n-1} - x_{n-2}) \\ &= (\lambda B)^2(x_{n-2} - x_{n-3}) \\ &\vdots \\ x_n - x_{n-1} &= (\lambda B)^n[Bd + x_o(I + \lambda B)], \end{aligned} \tag{53}$$

whence, as n approaches infinity, $x_n - x_{n-1}$ approaches zero and x_n and x_{n-1} approach x. Thus

$$x = (A'A + \lambda I)^{-1} A'b + \lambda(A'A + \lambda I)^{-1}x \tag{54}$$

and

$$x = A^{-1}b. \tag{55}$$

Again we stress that unless we can successfully invert A (from a computational point of view) and possess essentially deterministic knowledge of the vector h, then the vector x will not correspond to the object vector x. Procedures have been designed into the solutions which alleviate the lack of this knowledge. However, this does not minimize the importance of the iteration technique, since its use allows one to effectively vary the a priori information in a much easier way than by inverting a matrix for each new value of λ.

Dynamic Programming Solutions

Review of dynamic programming

To present some continuity with regard to dynamic programming procedures, the following brief review is presented. The reader is referred to Bellman (1957, 1960, 1962) for further details. A classically simple yet exhaustive example involving the principles of dynamic programming would serve this purpose, but such an example is hard to find for the procedure described.

Dynamic programming is a term used to describe the mathematical theory of performing a sequence of decisions, or more formally, the theory of multistage decision processes. In this paper it is used as a mathematical tool to sequentially compute the vector components of the object vector x.

The major advantage of dynamic programming over the solutions just considered is that no matrix inverses are needed. Thus, this procedure can alleviate matrix inversion errors for large or nearly singular matrices. However, the procedure also has limitations for the problem being considered. One such limitation is due to the use of time as an artificial index of available information, a facet to be discussed later in this section.

Dynamic programming solutions inherently depend upon the optimality principle, which is: An optimal policy has the property that whatever the initial state and initial decision are the remaining decisions must constitute an optimal policy with regard to the state resulting from the first decision. The general proof of this principle is proved by a combination of induction and contradiction. A specific proof is found in Bellman (1962).

For a more detailed review we consider the discrete deterministic process; deterministic in the sense that the result of a decision is uniquely determined by the decision and discrete in the sense that the process consists of a finite number of stages.

Bellman (1957) defines a state vector $p = (p_1, \ldots, p_m)$ which is a member of a set D and a sequence of transformations $T = \{T_q\}$. The transformations have the property that $p \in D$ implies $Tq(p) \in D$ also. That is, the transformations have the property of transforming the state vector at any stage of the process into its original set.

A policy is any choice of the set of variables which yields an allowable sequence of decisions. An optimal policy is the choice of q $q_1, \ldots, q_N$, which in our case minimizes a preassigned function of the final state p_N. Corresponding to each q_i we have T_{q_i}, and thus we may equivalently regard a policy as a selection of transformations. The preassigned function $R(p_N)$ of the final state is denoted as the criterion function.

Now we define

$$f_N(p) = \min R\,(p_N) \tag{56}$$

where $f_N(p)$ is the N stage return obtained starting from an initial state p and using an optimal policy.

We can use the optimality principle to obtain a recurrence relationship for the functions of the set $f_N(p)$. Suppose a transformation T_q is chosen as a result of the first decision. The new state vector is $T_q(p)$. The minimum value of the criterion function as a result of the next N-1 stages is $f_{N-1}(T_q(p))$ from (56) above. Using the optimality principle, if it is desired to minimize the total N-stage criterion function, q must be chosen so that

$$f_N(p) = \min f_{N-1}\,(T_q(p)) \tag{57}$$

for $N \geq 2$, with

$$f_1(p) = \min T_q(p). \tag{58}$$

Equations (57) and (58) are the crucial relationships used in discrete deterministic dynamic programming. Once these relationships are established for the particular problem at hand, the solution is well on its way to completion.

Stochastic dynamic programming

A similar formulation is available for a discrete stochastic process in which the basic difference is that a decision results in a distribution of transformations. Several additive noise forms were considered under this formulation, but for the forms considered only first moments of the additive noise entered directly into the solution, while second moments and first moments entered into the final error function. Since we can assume without loss of generality that the noise mean is zero, these solutions, except for the error term, were the same as the deterministic solutions.

Deterministic solution with area constraint

Bellman, et al. (1964, 1965) has obtained dynamic programming solutions for two constraints, one which involves the use of an a priori

vector p as in (47) and another which utilizes a smoothness constraint. Here we obtain a solution which uses an area constraint and is suited to the a priori information available in the problem.

We first show that the area under the object is attainable from the image and point spread function. The area under the object distribution is the product of the areas under the image and point spread functions. This is easily demonstrable. Consider the one-dimensional version of the imaging Equation (25) which is

$$b(\xi) = h(\xi) * x(\xi) \tag{59}$$

where the star as used here denotes convolution. The analogous expression in the spatial frequency domain is

$$B(f) = H(f)X(f) \tag{60}$$

where

$$f = \frac{\gamma}{\lambda R}.$$

Expression (60) holds for all f, thus when f = 0, we have

$$\int_{-\infty}^{\infty} b(\xi)d\xi = \int_{-\infty}^{\infty} x(\xi)d\xi \int_{-\infty}^{\infty} h(\xi)d\xi \tag{61}$$

which is the desired product of areas. The numerical approximation for these integrals yields

$$c = \sum_{i=1}^{M} x_i w_i = \sum_{i=1}^{N} b_i s_i / \sum_{i=1}^{N} h_i v_i \tag{62}$$

where w_i, s_i and v_i are quadrature weights.

Since c is attainable from (62) we can use (62) as the constraint and minimize

$$R_M(x) = \lambda\left(\sum_{i=1}^{M} w_i x_i - c\right)^2 + \sum_{i=1}^{N}\left(\sum_{j=1}^{M} a_{ij} x_j - b_i\right)^2. \tag{63}$$

Before proceeding with the dynamic programming solution, we shall consider the vector solution. Differentiating $R_M(x)$ with respect to x_k and equating the result to zero,

$$\frac{\partial R_M(x)}{\partial x_k} = \lambda\left(\sum_{i=1}^{M} w_i x_i - c\right) w_k + \sum_{i=1}^{N} a_{ik}\left(\sum_{j=1}^{M} a_{ij} x_j - b_i\right) = 0. \tag{64}$$

In matrix form

$$\lambda\, ww'x - \lambda cw + A'Ax - A'b = 0 \tag{65}$$

where $w^t = (w_1, \ldots, w_M)$ is the vector of quadrature weights. Solving for the vector x,

$$x = (A^tA + \lambda ww^t)^{-1} (A^tb + \lambda cw). \tag{66}$$

If we define the unit vector $e^t = (1, \ldots 1)$, then the vector form of the area constraint solution is

$$x = (A^tA + \lambda ee^t)^{-1} (A^tb + \lambda ce) \tag{67}$$

where $w_i = 1$, $i = 1, \ldots, M$.

Evaluating $R_M(x)$ by substitution of (67) into (63) we obtain the following quadratic form for the error:

$$\begin{aligned} e = {} & b^t \{ I + A[A^tA + \lambda ee^t]^{-1} A^t \} b \\ & + 2b\{ \lambda A[A^tA + \lambda ee^t]^{-1} e\} c \\ & + \{ \lambda + \lambda^2 e^t[A^tA + ee^t]^{-1} e\} c^2. \end{aligned} \tag{68}$$

This may be written in the form,

$$e = b'Q_M b + 2b'p_M c + r_M c^2 \tag{69}$$

which is referred to in the dynamic programming solution.

To begin the derivation of the dynamic programming solution, we specify that $1 \leq M \leq N$. For M = 1 there results the simple problem of minimizing

$$R_1(x) = \lambda(x_1 w_1 - c)^2 + \sum_{i=1}^{N} (a_{i1} x_1 - b_i)^2. \tag{70}$$

The minimizing value of x_1 is

$$x_1 = \frac{\lambda c w_1 + b' a_{(1)}}{\lambda w_1^2 + a'_{(1)} a_{(1)}} \tag{71}$$

where $a'_{(1)} = (a_{11}, a_{21}, \ldots, a_{N1})$ and in general $a'_{(M)} = (a_{1M}, \ldots, a_{NM})$ and represents the M^{th} column of the A matrix. $R_1(x)$ evaluated at the minimizing value of x_1 may be written in the form

$$\min_{x_1} R_1(x) = b'Q_1 b + 2b' p_1 c + r_1 c^2 \tag{72}$$

where

$$Q_1 = \frac{I - a_{(1)} a'_{(1)}}{\lambda w_1^2 + a'_{(1)} a_{(1)}},$$

$$p_1 = \frac{-\lambda w_1 a_{(1)}}{\lambda w_1^2 + a'_{(1)} a_{(1)}},$$

and

$$r_1 = \frac{\lambda - \lambda^2 w_1^2}{\lambda w_1^2 + a'_{(1)} a_{(1)}}.$$

The recurrence relationship

$$f_M(b, c) = \min_{x_M} [f_{M-1}(b - a_{(M)} x_M, c - w_M x_M)] \tag{73}$$

is stated for $M \geq 2$, with

$$f_1(b, c) = \min_{x_1} R_1(x). \tag{74}$$

This recurrence relation is derived in the Appendix using the optimality principle.

Referring to the definitions previously discussed, the parameters b and c represent the state variables, the functions $\{ f_i(b, c) \}$ correspond to the criterion functions, and the optimal policy corresponds to the choice of x_i or transformations $\{ b - a_{(i)} x_i, \; c - w_i x_i \}$.

The general form of the error for the vector solution (69) is similar to $f_1(b, c)$ shown in (74) and (72). It is a simple matter to prove

inductively that the general form of $f_M(b, c)$ is

$$f_M(b, c) = b'Q_M b + 2b'p_M c + r_M c^2. \tag{75}$$

To solve for the general <u>scalar value</u> x_M, the recurrence relation (73) is used with (75)

$$\begin{aligned} b'Q_M b + 2b'p_M c + r_M c^2 = \min_{x_M}\{ & (b-a_{(M)}x_M)' Q_{M-1}(b-a_{(M)}x_M) \\ & + 2(b - a_{(M)}x_M)' p_{M-1}(c-w_M x_M) \\ & + r_{M-1}(c-w_M r_M)^2\}. \end{aligned} \tag{76}$$

Expanding the expression on the right,

$$\begin{aligned} b'Q_M b + 2b'p_M c + r_M c^2 = \min_{x_M}\{ & b'Q_{M-1}b + 2b'p_{M-1}c + r_{M-1}c^2 \\ & + x^2_M[r_{M-1}w^2_M + 2w_M a'_{(M)}p_{M-1} + a'_{(M)}Q_{M-1}a_{(M)}] \\ & - 2x_M[cw_M r_{M-1} + ca'_{(M)}p_{M-1} + w_M b'p_{M-1}] \}. \end{aligned} \tag{77}$$

For convenience the first expression on the right in brackets is defined as G and the second expression in brackets is defined as D. Differentiating the expression on the right with respect to x_M and equating the result

to zero,

$$x_M = \frac{G}{D} = \frac{cw_M r_{M-1} + ca'(M)p_{M-1} + w_M b' p_{M-1} + a'(M)Q_{M-1} b}{r_{M-1} w_M^2 + 2w_M a'(M)p_{M-1} + a'(M)Q_{M-1} a(M)}. \quad (78)$$

All of the quantities in this expression are known except Q_{M-1}, p_{M-1}, and r_{M-1}, but recurrence relations for these quantities can be established by evaluating $f_M(b, c)$. Substituting (78) into (77),

$$b'Q_M b + 2b'p_M c + r_M c^2 = b'Q_{M-1} b + 2b'p_{M-1} c + r_{M-1} c^2 - \frac{G^2}{D}. \quad (79)$$

Expanding G^2 and equating quadratic coefficients the following recurrence relations are evident:

$$Q_M = Q_{M-1} - \frac{(w_M p_{M-1} + \alpha(M))(w_M p_{M-1} + \alpha(M))'}{r_{M-1} w_M^2 + 2w_M \rho_M + k_M}$$

$$p_M = p_{M-1} - \frac{(w_M p_{M-1} + \alpha(M))(r_{M-1} w_M + \rho_M)}{r_{M-1} w_M^2 + 2w_M \rho_M + k_M} \quad (80)$$

$$r_M = r_{M-1} - \frac{(r_{M-1} w_M + \rho_M)^2}{r_{M-1} w_M^2 + 2w_M \rho_M + k_M}$$

where for computational purposes we have defined auxiliary quantities

$$\alpha_{(M)} = Q_{M-1}\, a_{(M)},$$

$$\rho_M = p'_{M-1}\, a_{(M)},$$

and

$$k_M = a'_{(M)}\, Q_{M-1}\, a_{(M)}.$$

Recalling the expressions for Q_1, p_1, and r_1 in (72) we see that if $Q_o = I$, $p_o = 0$, and $r_o = \lambda$, then the relations just developed (80) hold for $1 \leq M \leq N$.

If the stability control parameter $\lambda = 0$, then (78) reduces to

$$x_M = \frac{a'_{(M)} Q_{M-1}}{a'_{(M)} Q_{M-1} a_{(M)}}. \tag{81}$$

Since the scalars $r_i = 0$, $i = 1, \ldots, M$, and the vectors $p_i = 0$, $i = 1, \ldots, M$, then for $\lambda = 0$ the recurrence relations (80) reduce to

$$Q_M = Q_{M-1} - \frac{\alpha_{(M)} \alpha'_{(M)}}{k_M} \tag{82}$$

for $1 \leq M \leq N$, with $Q_o = I$.

The general relations (78) and (80) constitute the dynamic programming solution for the restoration problem using the area constraint as

previously described. These relations along with the relations developed by Bellman, et al. (1965) for the prior vector constraint were used in actually simulating the restoration process. The results of these simulations appear in a later section.

The major advantage of dynamic programming over the matrix inverse solutions is that no matrix inverses are needed. As a further evaluation, an apparent limitation is discussed below.

Pseudo-time and dynamic programming

Time as used in dynamic programming is usually used as an index to indicate the availability of information. In applying dynamic programming to the restoration problem, time in this sense is an artificial index since in most cases all of the information needed to effect a solution is available at the same time. An elaboration of these statements follows.

In business problems, as well as in many other problems because of the sequence of time, only limited information may be available to operate upon at a certain stage of the problem, and in these cases one can do no better than act on the available information. For these cases the dynamic programming solution is optimum in the sense of using all of the available information. However, the stepwise minimization as noted in (70) is accomplished as if only the first column of the A matrix is available, when actually the entire A matrix was available. There is some cost for

this neglect. If this cost does not compensate for the computational error involved in matrix inversion, then, neglecting other factors, one would not use the dynamic programming approach. This tradeoff has not been investigated in its entirety, but the evidence from the results available indicates that the dynamic programming solutions are not as well suited to the optics problem as the matrix inverse solutions.

Combination of dynamic programming and matrix inversion

As a continuation of the above discussion, one logically may ask if there is a procedure one can use to combine the properties of dynamic programming (which enable one to treat large dimensionality) and yet use more A matrix information per object estimate. The author has developed a generalization of the dynamic programming procedure which allows one to restore the object vector as groups of subvectors. The solution for each subvector involves the inverse of an increasingly greater submatrix of the original A matrix and thus utilizes more available information than the scalar by scalar solution previously considered. It appears that this generalization may be applied to all of the dynamic programming constraints used thus far.

Because of the computer programming complexity involved and the fact that the matrix inverse procedures were computationally sound for the A matrix dimensionality considered, this generalization was not investigated further. However, future efforts should not overlook this

generalization, particularly in regard to the greater A matrix dimensionality inherent in two-dimensional object restoration.

ANALYTIC RESULTS WITH NOISE FOR SPECIFIC CASES

In addition to the theoretical implications that this section conveys, this section is presented in order to emphasize the trade-off between diffraction and noise in regaining object information which has been corrupted by these effects. Two rather separate looks at this trade-off have been investigated. In each case different amounts of a priori information about the object are assumed.

The first case involves the problem of discriminating between two different objects. It is known a priori that there is a choice between one of two objects and further that the objects are specified a priori as a one-point source and as a two-point source object. In the second case it is assumed that the object consists of two point sources, but because of diffraction effects the separation between the two point sources is not known. For the second case the relationship between the error variance of an estimate of the separation and diffraction is shown. It is important that the a priori information used in each case is emphasized so that we can better evaluate, qualitatively at least, the cost involved in regaining various amounts of information from the observed image.

A second point to be mentioned is that the results presented in this section are largely theoretical, while those that follow treat the more

practical aspects involved in the general restoration problem when there is virtually no a priori object information.

Detection Error vs. Diffraction

The problem of optical discrimination or detection has been considered by several authors (Helstrom, 1964; Harris, 1964b; Rushforth and Harris, 1966; [illegible]n, 1966). In this treatment several assumptions are made which enable the analytical solution to the detection problem to be obtained. This is done to present a rather complete example of diffraction and noise effects in the detection problem.

The noise model and detection scheme

To begin the problem, we assume that the image intensity distribution is due to one of two known objects plus background noise. We consider only the one-dimensional version of Figure 1 and that the objects are the point source $\delta(\alpha)$ located at the origin and the sum of two point sources $\frac{1}{2}\delta(\alpha-\eta_1) + \frac{1}{2}\delta(\alpha-\eta_2)$ located at η_1 and η_2 respectively. The images due to these objects are denoted respectively $p(\xi)$ and $q(\xi)$. To enable analytic results to be obtained, the point spread function is taken as

$$h(\xi) = \frac{1}{\sqrt{2\pi\sigma^2}}\, e^{-\xi^2/2\sigma^2}. \qquad (83)$$

Thus using (27) the images are

$$p(\xi) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\xi^2/2\sigma^2} \tag{84}$$

and

$$q(\xi) = \frac{1}{\sqrt{2\pi\sigma^2}} [e^{-(\xi-\eta_1)^2/2\sigma^2} + e^{-(\xi-\eta_2)^2/2\sigma^2}]. \tag{85}$$

With this formulation the two states of nature are

$$\omega_1: \; b(\xi) = p(\xi) + n(\xi) \tag{86}$$

$$\omega_2: \; b(\xi) = q(\xi) + n(\xi) \tag{87}$$

where $n(\xi)$ is the additive background noise. The block diagram of state of nature ω_2 is shown in Figure 2.

It is further assumed that the noise vector n (the discrete version of $n(\xi)$) is white and Gaussian with mean $\bar{n} = 0$ and covariance matrix $K_n = \sigma_n^2 I$. The more practical Poisson noise model used with a detector is discussed in the Object Estimation section. Since n is Gaussian, the conditional density functions of the image are also Gaussian and are

$$f(b/\omega_1) \sim N(p, \sigma_n^2 I) \tag{88}$$

$$f(b/\omega_2) \sim N(q, \sigma_n^2 I) \tag{89}$$

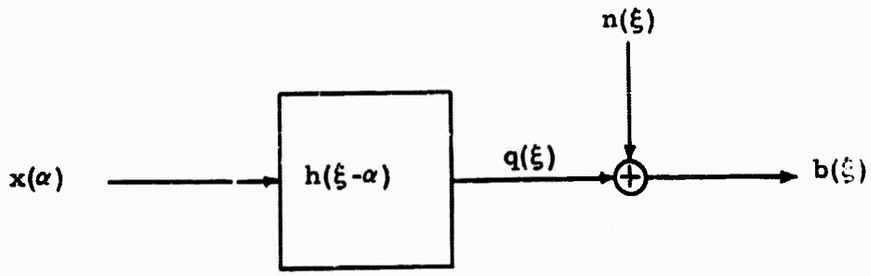

Figure 2. Block diagram of the optical configuration.

where $N(\mu, K_n)$ indicates a joint Gaussian density with mean vector μ and covariance matrix K_n, and b, p and q are, respectively, the discrete versions of $b(\xi)$, $p(\xi)$ and $q(\xi)$.

Using the techniques of statistical decision theory (Middleton, 1960), we can at this point specify the optimum discrimination procedure. If the two states of nature are equally likely a priori, and if the costs associated with the two types of error (i.e., choosing ω_1 when ω_2 is true and vice versa) are equal, then it can be shown that the optimum decision scheme (in the sense of minimizing the average cost or, in this case, the error probability) is the following:

$$\begin{aligned} &\text{choose } \omega_1 \text{ if } \ell(b) \geq 1 \\ &\text{choose } \omega_2 \text{ if } \ell(b) < 1 \end{aligned} \tag{90}$$

where

$$\ell(b) = f(b/\omega_1)/f(b/\omega_2) \tag{91}$$

is the likelihood ratio of the observed vector b. Since the natural logarithm is a monotonically increasing function of its argument, an equivalent statement is:

$$\text{choose } \omega_1 \text{ if } L(b) \geq 0$$

$$\text{choose } \omega_2 \text{ if } L(b) < 0 \tag{92}$$

where

$$L(b) = \ell n\, \ell(b). \tag{93}$$

Noting that $f(b/\omega_1)$ and $f(b/\omega_2)$ are Gaussian, substitution of (88) and (89) into (93) results in the decision procedure:

$$\text{choose } \omega_1 \text{ if } (p-q)^t b \geq \tfrac{1}{2}(q^t q - p^t p)$$

$$\text{choose } \omega_2 \text{ otherwise.} \tag{94}$$

Equation (94) is just the discrete form of the correlation detector or matched filter, matched in this case to the difference between the diffracted objects.

Equation (94) specifies the data processing necessary to make an optimum decision. However, we wish to investigate more thoroughly the consequences of using this procedure as diffraction and noise vary. To accomplish this we evaluate the probability of error associated with the discrimination procedure just described.

Evaluation of the discrimination procedure

There are two types of error which a discrimination procedure can lead to: Choosing ω_1 when ω_2 is the true state of nature (i.e., calling the two point source a one point source in our case), and choosing ω_2 when ω_1 is true. If decision procedure (94) is used, the error probability for the equally likely objects is

$$P_e = \tfrac{1}{2}P(G \geq \rho/\omega_2) + \tfrac{1}{2}P(G < \rho/\omega_1)$$

$$= \tfrac{1}{2}\int_{\rho}^{\infty} f(G/\omega_2)dG + \tfrac{1}{2}\int_{-\infty}^{\rho} f(G/\omega_1)dG \qquad (95)$$

where $G = (p-q)'b$ is the test statistic, $\rho = \frac{1}{2}(q'q-p'p)$ is the threshold, and $f(G/\omega_i)$ is the conditional density of G given that ω_i is the true state of nature. To evaluate the error probability we must find the conditional densities $f(G/\omega_2)$ and $f(G/\omega_1)$.

The Gaussian noise model allows these conditional densities to be found by noting that the scalar statistic G is a linear combination of independent Gaussian random variables, and thus G itself is Gaussianly distributed. To specify completely the densities $f(G/\omega_i)$, we need only find the conditional means $E(G/\omega_i)$ and the conditional variances $var(G/\omega_i)$. These conditional means and variances are, in this case, easily found. They are:

$$\mu_1 = E(G/\omega_1) = (p-q)^t p$$

$$\mu_2 = E(G/\omega_2) = (p-q)^t q$$

$$\sigma_1^2 = \text{var}\,(G/\omega_1) = \sigma_n^2\,(p-q)^t(p-q)$$

$$\sigma_2^2 = \text{var}\,(G/\omega_2) = \sigma_n^2\,(p-q)^t(p-q). \qquad (96)$$

Using (96), together with the Gaussian probability density function, the probability of error is

$$P_e = \tfrac{1}{2}\int_{\rho}^{\infty} \frac{e^{-(u-\mu_2)^2/2\sigma_2^2}}{\sqrt{2\pi\sigma_2^2}}\,du + \tfrac{1}{2}\int_{-\infty}^{\rho} \frac{e^{-(u-\mu_1)^2/2\sigma_1^2}}{\sqrt{2\pi\sigma_1^2}}\,du. \qquad (97)$$

Further algebraic manipulation yields the simpler form

$$P_e = \int_{d/2}^{\infty} \frac{e^{-u^2/2}}{\sqrt{2\pi}}\,du \qquad (98)$$

where

$$\frac{d}{2} = \frac{1}{2}\left\{\frac{1}{\sigma_n^2}(p-q)^t(p-q)\right\}^{\frac{1}{2}}. \qquad (99)$$

Equations (98) and (99) are the important relations. In the work that follows several quantities are defined which will enable us to extract a good deal of information from these relations about the probability of error as the diffraction and noise vary.

In order to simplify (99) recall that (84) and (85) are the continuous versions of p and q. Using a suitable definition as to the transformation of continuous quantities to discrete and vice versa, it can be shown that, when p and q are transformed back to their continuous versions (99) becomes

$$\frac{d}{2} = \frac{1}{2}\left\{\frac{1}{\sigma_n^2}\int_{-\infty}^{\infty}[p(\xi) - q(\xi)]^2\, d\xi\right\}^{\frac{1}{2}}. \tag{100}$$

At this point note that the quantity in braces may be defined as the "energy-to-noise" ratio

$$\text{ENR} = d^2 = \frac{1}{\sigma_n^2}\int_{-\infty}^{\infty}[p(\xi) - q(\xi)]^2\, d\xi \tag{101}$$

where the term "energy" refers to the integral square of the difference image. (The usual definition of this quantity as it applies to electronic signals is the signal-to-noise ratio (SNR). However, in this paper we reserve SNR for the more literal case where signal refers to the true observable image intensity or counts and noise refers to the sample mean absolute variation of the noise intensity or counts.)

The analytical formulation of the detection problem is found by substituting the images (84) and (85) into (100). Thus d/2 becomes

$$\frac{d}{2} = \left\{ \frac{1}{\sigma_n^2 \sqrt{\pi\sigma^2}} \left[3 + e^{\frac{-(\eta_2-\eta_1)^2}{4\sigma^2}} - 2\left(e^{-\eta_1^2/4\sigma^2} + e^{-\eta_2^2/4\sigma^2} \right) \right] \right\}^{\frac{1}{2}}. \qquad (102)$$

Note that d/2 depends upon both η_1 and η_2 and not merely on their separation (which we denote $\gamma = \eta_2 - \eta_1$), as might have been predicted. Thus d/2 and P_e both depend not only on the noise variance and diffraction but also on the a priori choice of η_1 and η_2.

It is not immediately clear from (102) just how diffraction plays a role in d/2 and P_e. In order to see this role more clearly consider the case when $\eta_1 = -\eta_2$.

$$\frac{d}{2} = \left\{ \frac{1}{\sigma_n^2 \sqrt{\pi\sigma^2}} \left[3 + e^{-\gamma^2/4\sigma^2} - 4e^{-\gamma^2/4\sigma^2} \right] \right\}^{\frac{1}{2}}. \qquad (103)$$

Now the importance of the diffraction ratio R as defined in (28) becomes more apparent since R is in this case

$$R = \frac{k\sigma}{\gamma} \qquad (104)$$

where k is an arbitrary constant which together with σ specifies the

"width" of the point spread function (i.e., $k\sigma$ controls the aperture opening) and γ being the separation is defined as the source width.

The basic trade-off between diffraction, noise, and P_e can be seen rather easily now. When R becomes large ($\gamma \to 0$ for fixed σ), then, as one would suspect d/2 approaches zero which drives the P_e to 1/2--the worst case. Alternatively, when R becomes small ($\gamma \to \infty$ for fixed σ) then

$$\frac{d}{2} \to \left[\frac{3}{\sigma_n^2 \sqrt{\pi\sigma^2}} \right]^{\frac{1}{2}} \tag{105}$$

and the probability of error reduces to a quantity which depends only upon the noise variance σ_n^2. Not until $\sigma_n^2 \to 0$ does P_e approach zero in this last case. (The ratio R may also be varied by fixing γ and varying σ. However, as discussed previously, the qualitative results are similar, and we have considered throughout the paper a fixed optical system which views various objects.)

In addition to the above discussion on limiting cases, we wish to consider a more detailed presentation of the variation between the pertinent quantities. Figures 3-6 comprise the resulting presentation. In order to vary the a priori information, i.e., η_1 and η_2, three cases were considered: Case 1, $\eta_1 = -\eta_2$, case 2, $\eta_1 = -1/4\ \eta_2$ and case 3, $\eta_1 = 0$ and $\eta_2 = \gamma$. The variation of R was appropriately accomplished

by fixing the aperture ($\sigma = 1/2$) and setting $k = 4$; thus the point spread function (83) contains 95% of its area within the "aperture width" of $4\sigma = 2.0$.

The variation of the error probability with diffraction is shown in Figure 3. The limiting cases discussed previously are evident: (1) As R becomes smaller then P_e for all of the curves approaches a quantity which depends upon the noise; note that this quantity changes noticeably as σ_n^2 for case 1 decreases from 1.0 to .001; (2) for R large then P_e approaches 0.5. An interesting feature shown in this figure is the trade-off between a priori information, diffraction and P_e for constant noise ($\sigma_n^2 = 1.0$). For R small case 1 (the two symmetrical point sources compared with a source at the origin) has the lowest P_e, but for R large case 1 exhibits the highest P_e.

This result may contradict one's intuition at first glance, but as the images are studied for the 3 cases it may be visually seen that for R small it is easier to discriminate between the images for case 1 than those for case 3 or case 2. On the other hand, for R large enough that the two symmetrical sources of case 1 are located "under" the point spread function ($\gamma < 2.0$) then it is easier to discriminate between the images of case 2 or case 3 than those of case 1 because one of the asymmetrical sources of case 2 and case 3 remains outside of the point spread function width longer. That is, R has to become larger (for

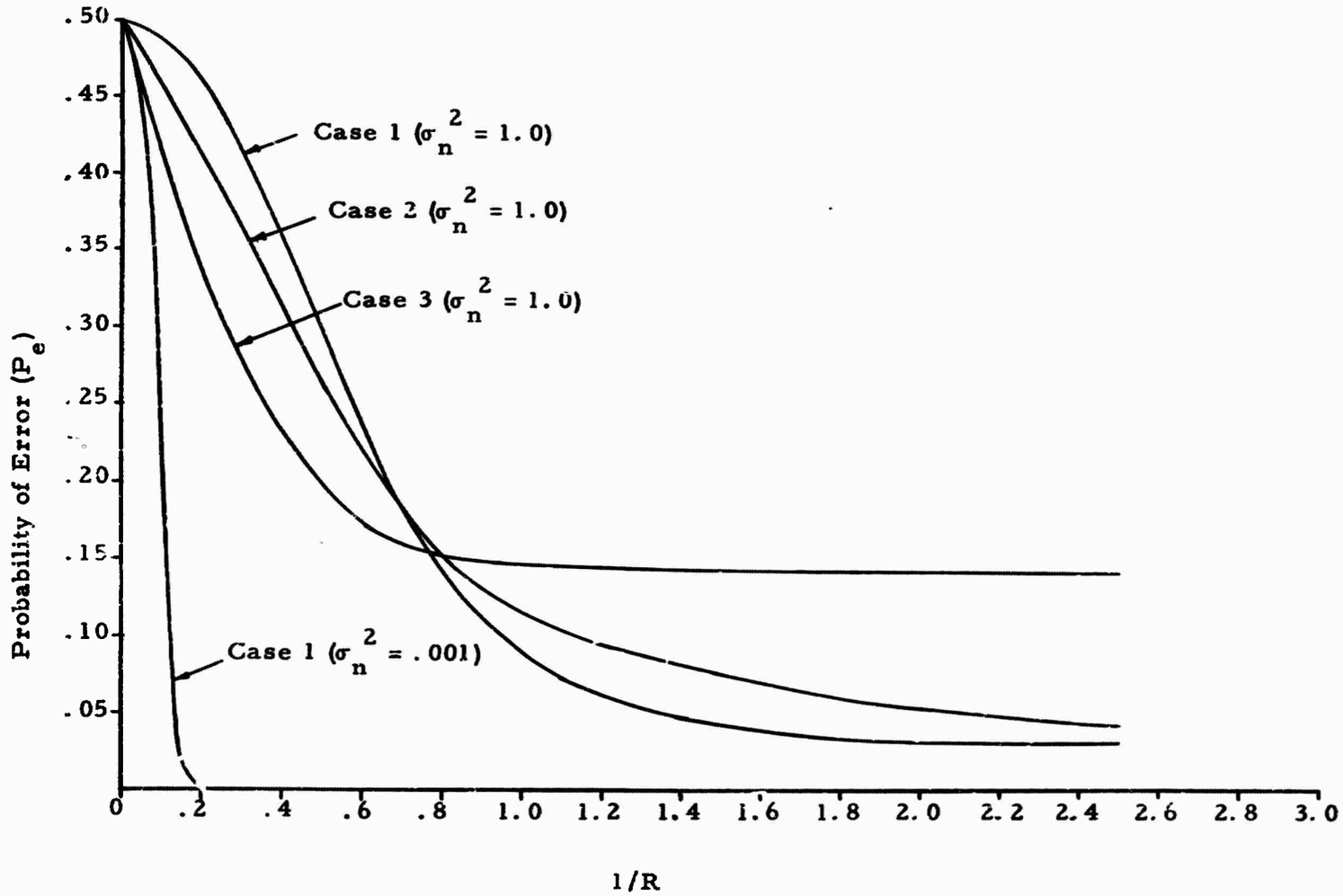

Figure 3. Probability of error vs. the reciprocal of the diffraction ratio R.

case 1 than case 2 or case 3) for both of the sources in case 2 or case 3 to be located within the point spread function width.

To further illustrate the effects of the noise level σ_n^2 on P_e Figure 4 shows P_e vs. log $(1/\sigma_n^2)$ for fixed large value of diffraction (R = 10.0). This figure clearly illustrates that even for R large, $P_e \to 0$ for sufficiently small σ_n^2, thus confirming the notion that noise not diffraction is the true limitation. It is also interesting to note the advantage of a priori knowledge in that the figure indicates a savings of about 10^2 in σ_n^2 (ENR) when case 3 is compared to case 1 for $P_e = 0.0$.

Figure 5 indicates the db loss incurred by increasing diffraction while maintaining a P_e referenced at R = 0.0. This figure shows that the diffraction effect acts essentially like a noise amplifier and indicates that, regardless of the initial noise level, the effects of diffraction alone can cause considerable ENR loss. For example, if the discrimination procedure produced a satisfactory P_e for R = .4, then to maintain this same P_e for R = 4.0 would require a 21.8 db increase in ENR.

In reiteration of the concept of diffraction, a small R does not imply that $P_e \to 0$ since as $R \to 0$ the P_e still depends upon the uncertainty caused by noise. In this respect the P_e for R small could still be rather high (perhaps even .5). Thus the db loss shown in Figure 5 is that loss incurred or similarly the increase necessary to maintain a P_e referenced only with respect to diffraction.

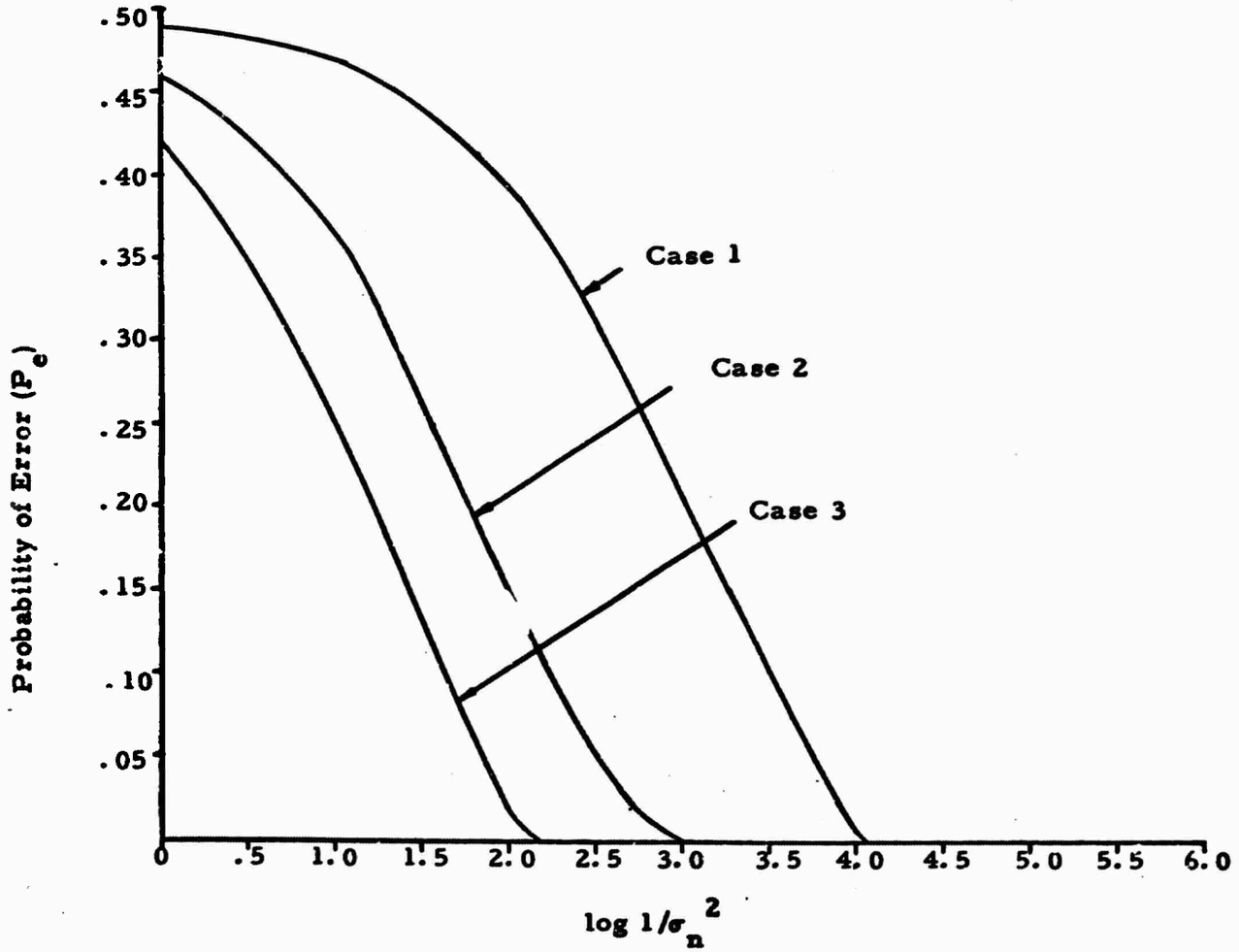

Figure 4. Probability of error vs log ($1/\sigma_n^2$) for R = 10.0.

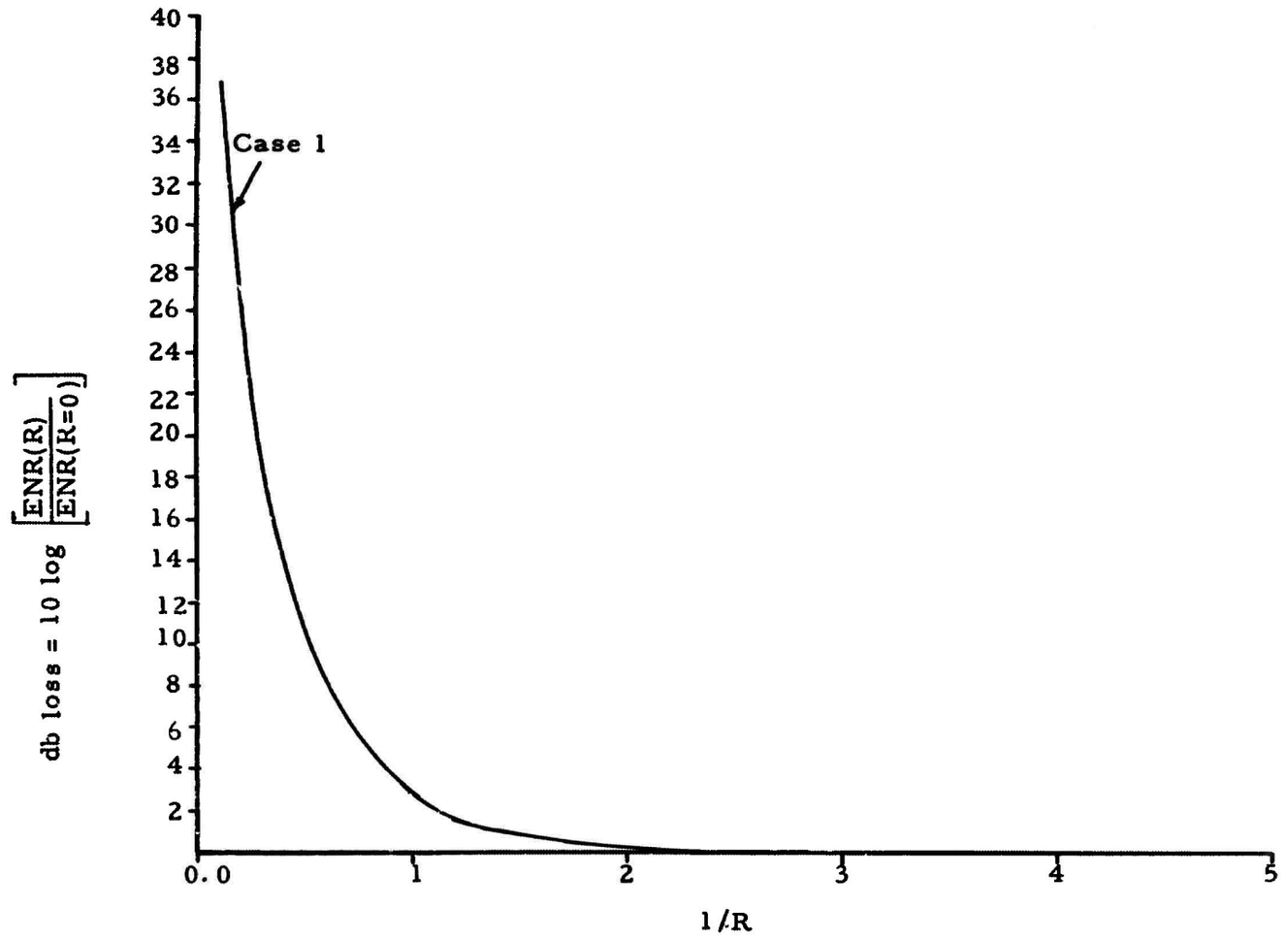

Figure 5. The db loss, referenced at R =0, vs the reciprocal of the diffraction ratio R.

We also desire to indicate the variation of the db loss incurred when the noise (σ_n^2) as well as diffraction varies. Figure 6 illustrates this variation. In this figure the ENR is referenced by fixing both R and σ_n^2 such that $P_e = .05$. Two points are illustrated by this figure: (1) We can, as was also illustrated in Figure 3, trade σ_n^2 for diffraction and still maintain the same P_e (this is seen as we observe for the 0 db reference points that by decreasing σ_n^2 from 1.0 to .001 we can increase R by a factor of 10 while still maintaining a $P_e = .05$). (2) The large diffraction effects are rather severe and could impose a severe limitation to the detection process unless the noise can be greatly reduced.

The results of the detection problem just considered afford a rather complete look at the basic problems imposed by diffraction and of the basic relationship that noise (σ_n^2) can be traded for diffraction (R). The next section asks a related question with regard to the information to be regained from the image and provides further insight as to the rather broad applicability of the relationship just mentioned.

Separation Error Bound vs. Diffraction

In this section we investigate the error variance associated with estimating the separation of two point sources when no a priori information is given as to the location of these two sources. The problem is formulated as follows.

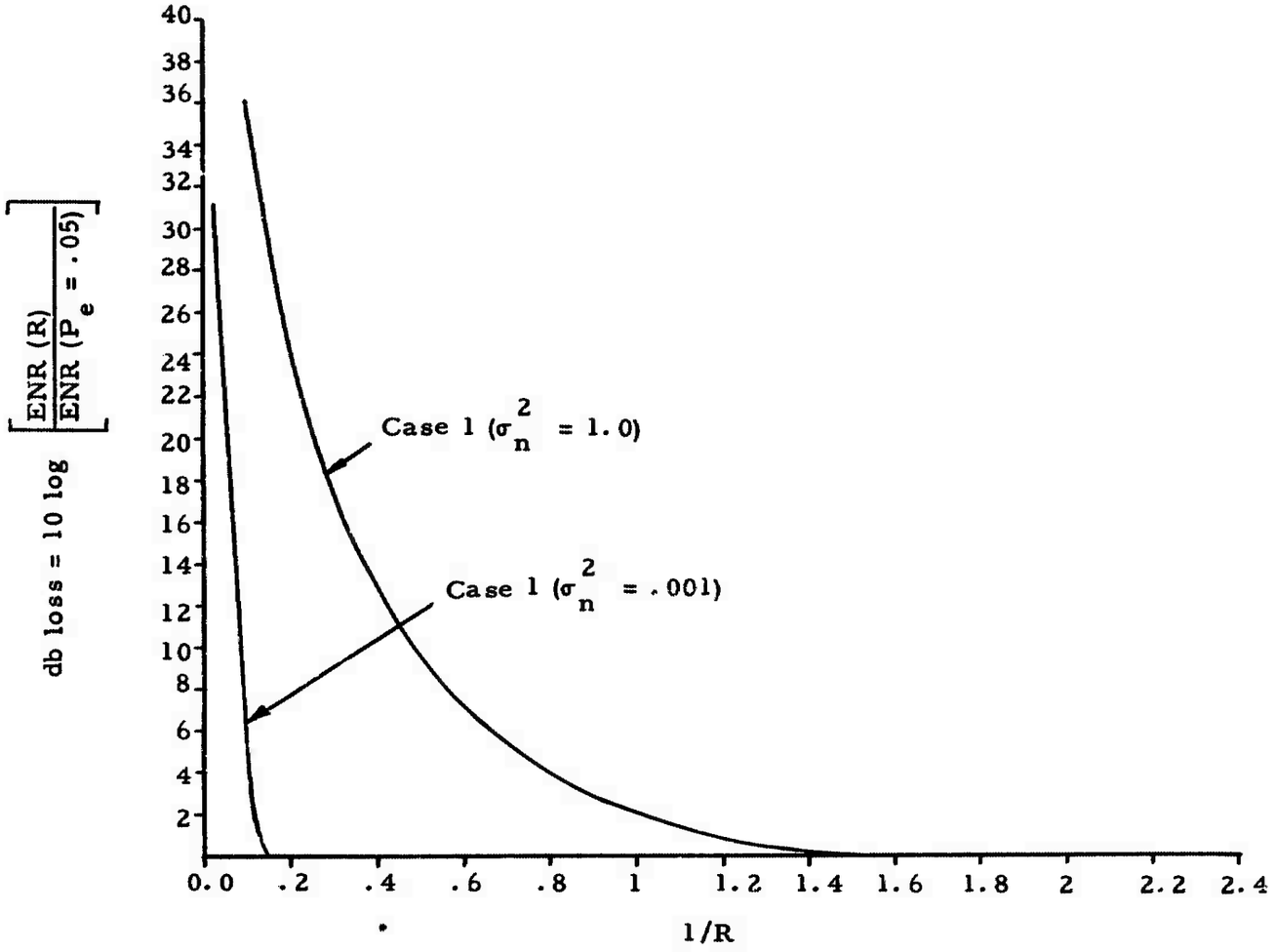

Figure 6. The db loss, referenced at P_e = .05, vs the reciprocal of the diffraction ratio R.

We consider the image due to two point sources located at η_1 and η_2 and determine the covariance matrix of the errors in estimating η_1 and η_2. We then define as a new random variable the separation $\gamma = |\eta_2 - \eta_1|$ and find the error statistics of this new random variable from those of η_1 and η_2.

The model considered is

$$b(\xi) = q(\xi, \eta_1, \eta_2) + n(\xi) \tag{106}$$

where the observed signal is $b(\xi)$ and $n(\xi)$ is Gaussian white noise. For these conditions Swerling (1964) has shown that the limiting elements for the inverse of the error covariance matrix of η_1 and η_2 are

$$B_{ij} = \frac{2}{N_o} \int_{-\infty}^{\infty} \frac{\partial q(\xi, \eta_1, \eta_2)}{\partial \eta_i} \frac{\partial q(\xi, \eta_1, \eta_2)}{\partial \eta_j} d\xi. \tag{107}$$

Again we consider the case when the point spread function is specified by (83). Thus when $x(\alpha) = \delta(\alpha - \eta_1) + \delta(\alpha - \eta_2)$

$$q(\xi, \eta_1, \eta_2) = \frac{1}{\sqrt{2\pi\sigma^2}} \left\{ e^{-(\xi - \eta_1)^2/2\sigma^2} + e^{-(\xi - \eta_2)^2/2\sigma^2} \right\}. \tag{108}$$

For convenience let the covariance matrix of η_1 and η_2 be K. Then from (107)

$$(K^{-1})_{11} = (K^{-1})_{22} = \frac{1}{2\sqrt{\pi}\, N_o \sigma^3} \tag{109}$$

and

$$(K^{-1})_{12} = (K^{-1})_{21} = \frac{e^{-\gamma^2/4\sigma^2}}{4\, N_o \sqrt{\pi}\, \sigma^5} [2\sigma^2 + \gamma^2]. \tag{110}$$

From (109) and (110) var (η_i) and cov (η_i, η_j) can be determined. Then, as referred to previously, we define $\gamma = \eta_2 - \eta_1$ and determine the error variance of this parameter (the separation) from

$$\text{var}\,(\gamma) = \text{var}\,(\eta_1) + \text{var}\,(\eta_2) - 2\,\text{cov}\,(\eta_1, \eta_2) \tag{111}$$

The resulting limiting error variance of the separation is

$$V = 4\sqrt{\pi}\, N_o \sigma^3 \cdot \left\{1 - e^{-\gamma^2/4\sigma^2}(1 - \gamma^2/2\sigma^2)\right\}^{-1}. \tag{112}$$

This expression may be regarded as the lower bound variance for all of the estimates of the separation which are minimum variance unbiased estimates.

We seek to investigate the variation of this error variance with diffraction and noise. As in the detection example we fix $\sigma = 1/2$ and consider that $R = 2.0/\gamma$. The limiting cases are comparable to those resulting in the detection example. For R large ($\gamma \to 0$) then the error variance approaches infinity and the separation between the point sources cannot be regained. When R is small ($\gamma \to \infty$) then

$$V \to \sqrt{\pi}\,\frac{N_o}{2} \tag{113}$$

and here as in the detection problem the error depends upon the noise variance (in this case $N_o/2 = \sigma_n^2$). Not until $N_o \to 0$ does the error variance approach zero.

To further illustrate the trade-off between the error variance and diffraction for a fixed noise level we consider Figure 7 which shows the db loss vs. 1/R (the separation in this case). The db loss is referenced for R = 1.64. This reference was chosen because the general curve of error variance vs. R has a weak minimum at R = 1.64. The existence of this minimum appears to violate the general relationship that the error is proportional to R, and as yet no physical or mathematical reason is apparent which explains the existence of the minimum. However, once R > 1.64 the same general relationship shown previously for the discrimination problem exists between V and R. For example, suppose

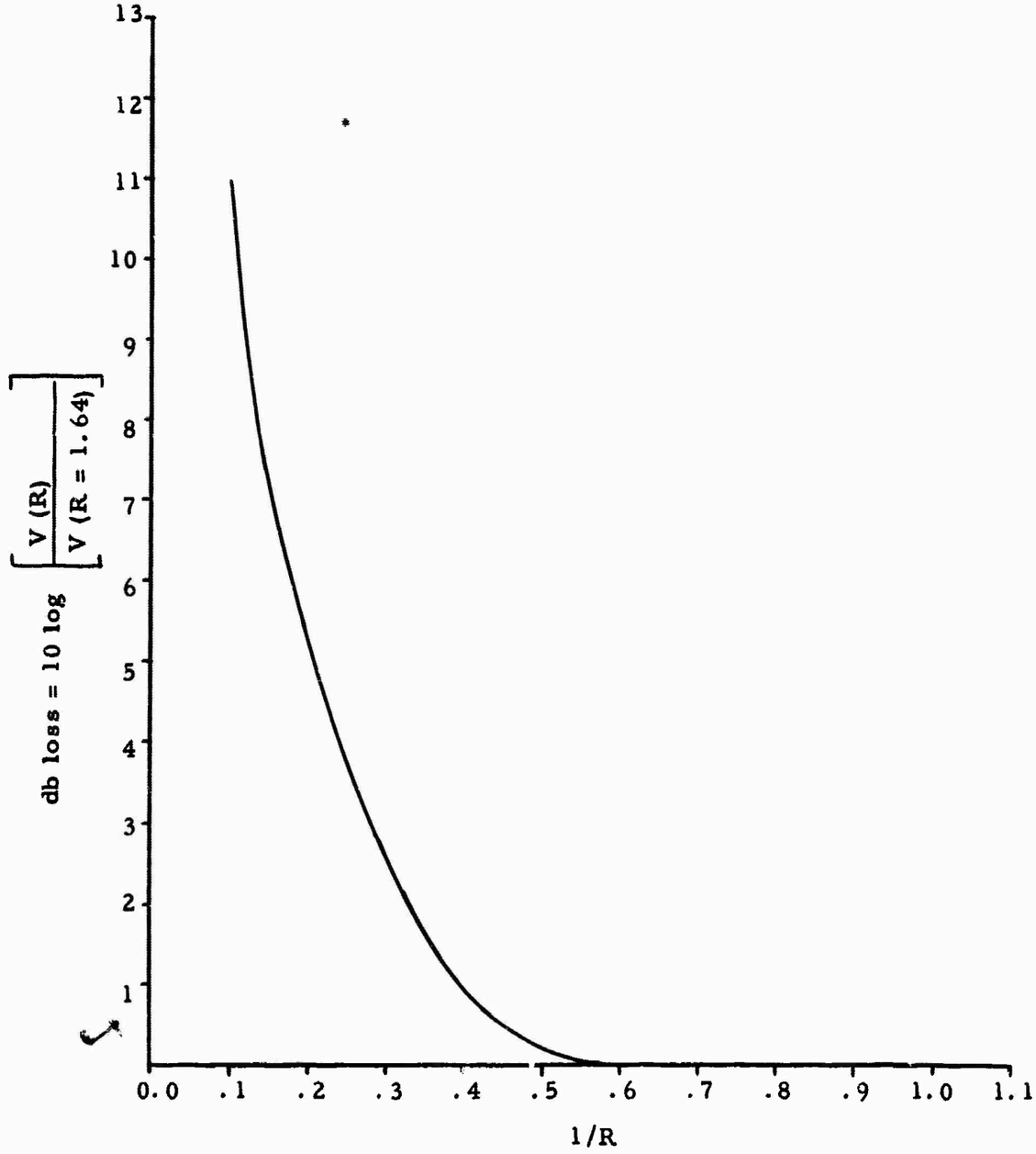

Figure 7. The db loss vs 1/R for a fixed aperture and noise variance.

that we were satisfied with the error variance achieved for R = 2.0; then in order to maintain this error variance for R = 10 we must provide for an increase of 10.8 db. Thus, as Figure 5 illustrates, the effect of increasing R (decreasing the object size) is essentially one of amplifying the existing noise level, and in order to maintain a constant error variance we must increase the SNR. Also noted is the rather severe limitation imposed by excessive diffraction. This same comment, as we shall demonstrate, applies to the numerical solution when restoring incoherent objects.

The next section presents and discusses several noise models which are applicable to the restoration problem and develops the minimum MSE estimate, which is used as the basic solution in this paper to restore objects.

OBJECT ESTIMATION

The sequence of this section is as follows. First we discuss and present the noise models associated with the restoration problem. The basic sources of noise considered are the additive background noise and the multiplicitative detector noise.

Next we consider using the functional form of the multiplicitative density function in estimating the object vector from the Baysean approach. This approach offers some insight as to how a priori object information might be encoded, but as will be discussed two basic drawbacks to this approach are that the additive noise is considered as a non-random variable and that the use of a priori object information in the inverse operator is not possible for the mathematical model considered.

From the Baysean approach we move on to the minimum MSE approach. Here we consider the additive noise as a random variable and are able to utilize a priori information in the inverse operator in a strikingly similar manner to that shown by Phillips and Twomey. With a reasonable assumption for practical problems, it is easily demonstrated that the MSE estimate, when multiplicative effects are considered, differs only by a multiplicative constant from the case when these effects are neglected. Since both of the noise models can be effectively represented in the MSE estimate and since a priori information can be profitably

encoded, the basic tool for object restoration used in this paper is considered to be the MSE estimate.

Although the MSE estimate effectively accounts for noise and uses a priori information, it is not the ultimate estimate. As an indication of one of the areas where the MSE estimate could be improved upon, we conclude this section by discussing the minimum distance estimate for the object $x(\alpha)$.

Noise Models

Figure 8 shows the block diagram of the imaging process and indicates the sources of noise.

The additive model

The general term additive noise, as previously used, has referred to the observed quantity, either intensity or mean counts, which is present when the object radiation is absent. The radiation which causes this intensity or mean count fluctuation is often referred to as 'background' radiation and is that radiation which accounts for stray light fluctuations as well as the background intensity in which the object intensity distribution is immersed. We assume here as Helstrom (1964) does that this noise is basically additive. In discrete form the additive model is

$$b = q + n = Ax + n \tag{114}$$

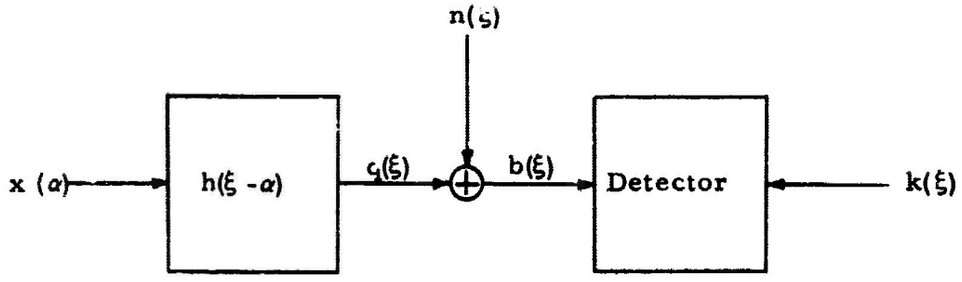

Figure 8. Block diagram of the imaging process with noise sources.

$$b = q + n = Ax + n \tag{114}$$

where n is the additive noise vector.

The multiplicative model

When the radiation strikes the detector, which is often a photomultiplier tube, a single photoemissive surface or photographic film, the general nature of the reaction must be described in statistical terms. To more accurately describe this model the surface of the detector is considered as being divided into N cells each with incremental area dA small enough such that the intensity impinging upon any one cell is constant throughout that cell. In practice, it is convenient to characterize the detector surface by a scalar parameter known as the quantum efficiency, which we shall denote as η. Under suitable practical conditions the mean number of counts observed in a time interval of length τ in each cell may be taken as

$$z_i = \frac{\eta \tau}{h\nu} b_i \tag{115}$$

where b_i is defined as the i^{th} incoming intensity component due to signal q_i plus noise n_i, h is Plank's constant and ν is the mean frequency of the quasichromatic radiation. Furthermore, the counts (number of resulting photoelectrons) k_i can be assumed for suitable conditions which have been

verified in practice (Mandel, 1959; Goodman, 1965) to obey the Poisson distribution

$$P(k_i/b_i) = \frac{(Db_i)^{k_i} e^{-Db_i}}{k_i!} \tag{116}$$

where

$$D = \frac{\eta}{h\nu}. \tag{117}$$

It also appears to be reasonable to assume that the counts in each cell are statistically independent from those in another cell (Helstrom, 1964). Thus

$$P(k/b) = \prod_{i=1}^{N} \frac{(Db_i)^{k_i} e^{-Db_i}}{k_i!} \tag{118}$$

where k is the count vector with components k_i.

Equations (115), (116), and (118) constitute the various stages of the detector noise model.

Temporal variation

As a further discussion of the noise models we elaborate on the assumption that is made with regard to temporal variation. In the Imaging Equation section it was stated that the scalar quantity as

described by (2) specifies the electromagnetic field radiating from a point in the object plane as a function of distance and time. In carrying on with the derivation of the imaging equation we assumed that no information regarding time variation would be considered. To make this assumption more explicit we state that the reason for this assumption is that instruments for measuring intensity generally cannot follow the instantaneous fluctuations at optical frequencies (Helstrom, 1964). Thus, in the strict sense the observed intensities (b and n) and the estimated intensity x are time averages. This does not imply that the observed quantities b and n are fixed and known. That is, we consider that a priori information is available about these observable quantities, and in the general case we consider that a priori information may be available about the object x. In fact, in practice the time average will only be taken over a finite observational interval, so that from a practical point of view b, n and x are only sample means and not true means. This consideration allows us to consider a more flexible model of additive noise variation in that the mean noise intensity n may vary from observation interval to observation interval as well as spatially. In the practical sense then, n is still to be considered as a random vector with covariance matrix K_n in which any one or all of the terms in K_n are nonzero and reflect our uncertainty about the true noise vector. Further discussion concerning time variation is presented after the derivation of the MSE estimate near the end of this section.

Discussion of the Baysean Approach

The motivation for discussing this approach is due to the general insight which this method of estimation theory has in the past provided for problems of this general type, particularly with regard to the profitable use of a priori information.

The term Baysean as used here refers to the general estimation schemes that have been developed using the density functions in Bayes' theorem. For our problem Bayes' theorem is,

$$f(x/k) = \frac{P(k/x)\, f(x)}{\int_{-\infty}^{\infty} P(k/x) f(x) dx} \tag{119}$$

where f(x/k) is the a posteriori density function, P(k/x) the likelihood function, and f(x) the a priori density function.

Three estimates, the maximum likelihood (ML), maximum a posteriori, and Bayes' may in principle be obtained from (119) once P(k/x) and f(x) are specified. We first present the derivation of the ML estimate. Next we discuss a problem which arises as we attempt to determine the Bayes' and maximum a posteriori object estimates. We will indicate how this problem, which is one of mathematical complexity, can be alleviated at the cost of losing the use of a priori information in the inverse operator. In conclusion we will present the form of the

resulting Bayes' and maximum a posteriori estimates and discuss the insight it provides as to the weighting of a priori information.

The ML estimate is defined as that set of parameters x (which for this estimation procedure are regarded as fixed but unknown), such that the likelihood function P(k/x) is maximized (Mood and Greybill, 1963). In obtaining this estimate as well as those which follow in this discussion the additive noise is assumed to be fixed and known.

Finding the object estimate which maximizes P(k/x) in (118) is accomplished as follows. Since the natural logarithm is a monotonic function of its argument, we can equivalently maximize

$$L(x) = \ell n\, P(k/x). \tag{120}$$

Upon differentiating L(x) with respect ao an arbitrary component of x, say x_m, and setting the result equal to zero we obtain

$$\frac{\partial L(x)}{\partial x_m} = \sum_{i=1}^{N} \left[\frac{k_i}{q_i + n_i} - D \right] a_{im} = 0 \tag{121}$$

where $q_i = \sum_{j=1}^{M} a_{ij} x_j$. Noting that this equation holds for all $m = 1, \ldots, M$, then

$$vA = 0 \tag{122}$$

where v is an N x 1 vector whose i^{th} component is

$$v_i = \frac{k_i}{q_i + n_i} - D. \tag{123}$$

Clearly, if A is square (M=N) and nonsingular, then the estimate x can be found from

$$v = 0. \tag{124}$$

Since we have M=N equations of the form

$$\sum_{j=1}^{M} a_{ij}\hat{x}_j = \frac{k_i}{D} - n_i \tag{125}$$

then we can solve for the ML estimate, which is

$$\hat{x} = A^{-1}(k/D-n). \tag{126}$$

The ML estimate does not provide any new information which would overcome the basic problem of an unstable A matrix whose inverse operates on a noisy observed vector. It does, however, indicate that the ML procedure results in the straightforward inverse operator which operates

on a modified count vector (modified by dividing out the multiplicative detector effect and subtracting the fixed and known background intensity n).

As was demonstrated by Phillips (1962) the addition of a priori information does alleviate the unstable A matrix problems; thus we seek to investigate the Bayes' and maximum a posteriori estimates which use the a priori density f(x). The Bayes' estimate considered here is obtained by minimising the quadratic loss function, while the maximum a posteriori estimate is designated as the vector x which maximizes the a posteriori density f(x/k). The success in obtaining these estimates, as far as mathematical tractability is concerned, depends largely upon the form of f(x). Herein lies the problem, referred to previously. The form of f(x) must be such that one can mathematically solve for either the a posteriori mean in the Bayes' case or determine a unique vector x which will maximize f(x/k). Thus far we have been unable to specify a nontrivial f(x) to enable these estimates to be determined. The apparent difficulty is due to the functional relationship between the object x and the image Ax.

We have, however, been able to specify an a priori density for the image vector b which enables the determination of the optimum estimate for the image b. This density is the joint gamma density, which is

$$f(b) = \prod_{i=1}^{N} \frac{c_i^{u_i} (b_i)^{u_i - 1} e^{-b_i c_i}}{\Gamma(u_i)} \qquad (127)$$

where the parameters c_i and u_i specify the mean $\bar{b}_i = \frac{u_i}{c_i}$ and the variance $\text{var}(b_i) = u_i/c_i^2$. Here we have assumed that the intensity b_i is independent from b_j for all i and j.

Utilization of f(b) as specified above with the Poisson distribution allows the determination of the a posteriori density f(b/k), which is again a joint gamma distribution with parameters $D + c_i$ and $k_i + u_i$ which replace, respectively, the parameters c_i and u_i in (127). With the specification of f(b/k) in the form of the joint gamma distribution, we can solve for either the Bayes' or the maximum a posteriori estimate for the image $\hat{b}$. Using this estimate we consider that a reasonable, although suboptimum, procedure is to operate on $\hat{b}$ to obtain an estimate for x.

This procedure is easily accomplished for the Bayes' case. Since the Bayes' estimate is the a posteriori mean we have

$$\hat{b}_i = E[b_i/k_i] = \frac{k_i + u_i}{D + c_i}. \tag{128}$$

Using the proecdure just discussed, we have N equations of the form

$$\hat{b}_i = \sum_{j=1}^{M} a_{ij}x_j + n_i = \frac{k_i + u_i}{D + c_i}. \tag{129}$$

For A nonsingular and (M=N) then the suboptimum estimate for x is

$$\hat{x} = A^{-1}[k_c + \bar{b}_c - n] \tag{130}$$

where k_c and $\bar{b}_c$ are, respectively, N x 1 dimensional vectors with components

$$k_{c_i} = \frac{k_i}{D + c_i}$$

and

$$\bar{b}_{c_i} = \frac{c_i \bar{b}_i}{D + c_i}$$

with $\bar{b}_i = u_i/c_i$. The suboptimum maximum a posteriori estimate for x is found in a similar manner as that demonstrated in obtaining the ML estimate. Its form is essentially that of the Bayes' estimate shown in (130).

In order to discuss the insight Bayes' estimate provides in using a priori information, we consider

$$\hat{x}_i = \sum_{j=1}^{N} a_{ij}^{-1} \left[\frac{D}{D + c_j} \left(\frac{k_j}{D} \right) + \frac{c_j}{D + c_j} \bar{b}_j - n_j \right] \tag{131}$$

where the a_{ij}^{-1}'s are the elements of A^{-1}.

Recall that the constant D is defined in (117) and contains the observation period τ. If τ becomes large, D becomes large and the observed counts k_j approach the mean counts z_j and thus k_j/D approaches the mean intensity b_j.

The mean and variance of the j^{th} prior gamma distribution are u_j/c_j and u_j/c_j^2. Thus, generally speaking small c_j values correspond to large prior uncertainty, since the variance is large and the a priori distribution is rather broad and flat.

Now refer to (131) and consider the case when c_j is small and D is large. Here we see that the estimation procedure weights the observation much more than the a priori mean component $\bar{b}_j$. This case corresponds to very large a priori uncertainty, and naturally we have little confidence in the a priori mean.

Alternatively, consider the case when c_j is large and D is small. The situation is the reverse of that just considered. We weight the a priori mean $\bar{b}_j$ heavily now and almost disregard the observation k_j. Here the estimation procedure reduces to one of essentially "cleaning up" or enhancing the a priori mean $\bar{b}_j$.

Minimum MSE Approach

In this section the estimation problem is approached from a different point of view, which does not require the specific functional

form of the a priori density or the likelihood function to be specified. We now assume that the additive noise vector is random and treat two basic types of MSE estimation. The first is the "classical" least squares approach, which assumes that no a priori information about the object is available and that the additive noise has zero mean and covariance matrix $\sigma_n^2 I$. Secondly, we consider the more general case when the noise covariance matrix is K_n, and we take advantage of encoding a priori object information in the form of the prior mean $\bar{x}$ and the covariance matrix K_x.

Before beginning the two approaches the noise model is more explicitly stated. Recall (115) and note that

$$z_i/D = \sum_{j=1}^{M} a_{ij}x_j + n_i \tag{132}$$

or

$$b_i = \Sigma\, a_{ij}x_j + n_i. \tag{133}$$

Since n_i is a random variable, the quantity z_i/D or b_i is now a random variable. If n_i has mean zero, which we can assume without loss of generality, then the model is

$$b = \bar{b} + n \tag{134}$$

where $\bar{b} = Ax$ and $\bar{b}$ is the mean of b.

Classical least squares

The classical least squares approach presented here has been treated by several authors. The reader is referred to Mood and Greybill (1963) and Deutsch (1965) for further discussion. It is instructive to mention that in the general case treated by Deutsch (1965), the assumption of a linear relationship may be necessary between the unknown parameter x and the observed random vector b. However, in the restoration problem, this assumption is not necessary since the imaging process is inherently linear.

We consider here the case for $K_n = \sigma_n^2 I$. The parameter x is chosen to minimize

$$R(x) = (Ax-b)'(Ax-b) \tag{135}$$

or alternatively

$$R(x) = \sum_{i=1}^{N} \left(\sum_{j=1}^{M} a_{ij} x_j - b_i \right)^2. \tag{136}$$

The parameter x is determined in the usual manner by differentiating R(x) with respect to an arbitrary element of x, say x_m, and setting the result equal to zero. In summation form

$$\sum_{i=1}^{N} a_{im} \left(\sum_{j=1}^{M} a_{ij} \hat{x}_j - b_i \right) = 0, \tag{137}$$

since this expression must hold for all m = 1, . . . , M,

$$A^t A \hat{x} = A^t b. \tag{138}$$

If $(A^t A)$ is nonsingular, then the estimate is

$$\hat{x} = (A^t A)^{-1} A^t b. \tag{139}$$

When A is square and nonsingular

$$\hat{x} = A^{-1} b. \tag{140}$$

Thus, the classical least squares approach results in the same basic inversion form as shown in the ML case and as stated in the Imaging Equation section. Even though statistical methods have been used in attaining this basic form, we still have not been able to gain much

additional insight which would enable us to overcome the problems of instability and noise referred to earlier. The inclusion of a priori information as shown next does, however, provide considerable insight into these problems.

The general minimum MSE estimate

Now we assume that a priori information is available in the form of the prior object mean $\bar{x}$ and the covariance matrix K_x. Information about the noise vector n is assumed to be available as encoded in the covariance matrix K_n, and we assume that the noise mean is zero. We further assume that x and n are independent. The procedure which follows in obtaining the general result (149) was first applied to the restoration problem by Rushforth (1965).

We postulate a linear model which enables the use of K_x, $\bar{x}$ and K_n. We assume that x and the random intensity vector b are related by

$$x = Hb + g \tag{141}$$

where H is an MxN matrix and g is an Mxl column vector. We seek H and g such that the quantity

$$MSE = E[(\hat{x}-x)^t (\hat{x}-x)] \tag{142}$$

is minimized where x represents the true object. Once H and g are determined, then the optimum estimate $\hat{x}$ is found by substituting these quantities into the expression for $\hat{x}$ above.

Recall (114) that b = Ax + n. It can be shown (Rushforth, 1965) that the vector g can be written

$$g = (I-HA)\bar{x}. \tag{143}$$

Substitution of (141) and (143) into (142) yields

$$\begin{aligned} MSE = tr\{ & [H(Ax + n) + (I-HA)\bar{x}-x] \\ & [H(Ax + n) + (I-HA)\bar{x}-x]' \} \end{aligned} \tag{144}$$

where tr denotes the trace of a matrix. Upon expanding (144) and performing the expectation over x and n we can write (144) in the following form:

$$MSE = tr\{H[A\,K_x A' + K_n]H' - K_x A'H' - HA\,K_x + K_x\}. \tag{145}$$

To solve for H we complete the square in (145) and obtain

$$e = \mathrm{tr}\Big\{[H(A K_x A' + K_n)^{\frac{1}{2}} - K_x A'(A K_x A' + K_n)^{-\frac{1}{2}}]$$

$$[H(A K_x A' + K_n)^{\frac{1}{2}} - K_x A'(A K_x A' + K_n)^{-\frac{1}{2}}]'$$

$$+ K_x - K_x A'(A K_x A' + K_n)^{-1} A K_x\Big\}. \tag{146}$$

Now the minimizing value of H is easily determined by equating either one of the two expressions in brackets above equal to zero. Thus,

$$H = K_x A'(A K_x A' + K_n)^{-1}. \tag{147}$$

Further manipulation reveals that H may be represented in the alternate form

$$H = (A' K_n^{-1} A + K_x^{-1})^{-1} A' K_n^{-1}. \tag{148}$$

Using equations (148) and (143) for H and g the general expression for $\hat{x}$ is

$$\hat{x} = (A' K_n^{-1} A + K_x^{-1})^{-1}(A' K_n^{-1} b + K_x^{-1} \bar{x}). \tag{149}$$

An important quantity to consider in the evaluation of how well $\hat{x}$ performs is the MSE which results when x is used. The resulting MSE

is found as follows. By substituting H into (146) and noting the equality of (147) and (148), the MSE expression becomes

$$MSE = tr\left\{K_x - (A^t K_n^{-1} A + K_x^{-1})^{-1} A^t K_n^{-1} A K_x\right\}. \quad (150)$$

Thus,

$$MSE = tr\left\{(A^t K_n^{-1} A + K_x^{-1})^{-1}[(A^t K_n^{-1} A + K_x^{-1}) K_x - A^t K_n^{-1} A K_x]\right\}. \quad (151)$$

The quantity in brackets in (151) reduces to I; thus, the general expression for the MSE reduces to

$$MSE = tr\left\{(A^t K_n^{-1} A + K_x^{-1})^{-1}\right\}. \quad (152)$$

Multiplicative effects for the high SNR case

We assume in this paper that the observed quantity is either intensity or mean counts. That is, the observation interval is assumed to be long enough so that when the detector is present the actual number of counts k_i is a very good approximation for the true mean counts z_i. Thus when

$$k_i \approx z_i, \quad (153)$$

the optimum estimate for x is

$$\hat{x} = [A^t K_n^{-1} A + K_x^{-1}]^{-1} [A^t K_n^{-1} (z/D) + K_x^{-1} \bar{x}]. \qquad (154)$$

In practice when k_i deviates significantly from z_i, we must consider a more sophisticated model such as that treated by Austin (1966) in which he assumed that we observe counts as they emerge from the detector but estimate the mean intensity x of the object. This treatment adds insight (in the limiting case for $\tau \to \infty$ the estimates are the same), but indicates that much more complexity is involved in the restoration process when counts are observed.

An increase in τ essentially implies an increase in SNR. The evidence presented in the Analytic Results with Noise for Specific Cases section and the Simulated Object Restoration section indicates that for reasonable restoration (for $R > 1.0$) we must consider the high SNR case. These results indicate that the SNR must be high enough so that we can neglect the more complex model and just consider the general case (154). By high SNR we are implying, generally speaking, that for the diffraction range considered ($R > 1.0$) the SNR must be high enough to neglect the detector except for a multiplicative constant.

Another way of stating the high SNR assumption is to state that the SNR is high enough so that the central limit theorem applies (Parzen, 1960)

and we have Gaussian rather than Poisson statistics. In this case it is easily demonstrated that if $k_i \approx z_i$ then the use of Gaussian distribution functions for f(z/x) and f(x) result in a Bayes' estimate or a maximum a posteriori estimate that is equivalent to (154). Thus, the assumption of high SNR ties the Baysian and MSE approaches together.

To follow up the earlier discussion on temporal variation, the high SNR assumption implies that the observation interval is long enough so that n, b and x are sample means with respect to time that are near, but not equal, to the true temporal mean. Thus, K_x and K_n reflect our ignorance of the true quantities. As is the usual case, large elements in K_n and K_x imply a greater uncertainty than small elements.

Next we consider special cases of the general estimate (149) and the MSE (152) and discuss the trade-off between a priori information and noise.

Discussion of the MSE estimates--special cases

The trade-off between a priori information and noise is easily seen when $K_x = \sigma_x^2 I$ and $K_n = \sigma_n^2 I$. The general estimate (149) becomes

$$\hat{x} = (A'A + \sigma_n^2/\sigma_x^2 I)^{-1}(A'b + \sigma_n^2/\sigma_x^2 \bar{x}). \qquad (150)$$

Consider the case when $\sigma_x^2 >> \sigma_n^2$. This case corresponds to large a priori uncertainty. Here, as before in the discussion of the Bayes and maximum a posteriori estimates, we tend to weight the observation vector b much more heavily than the prior mean $\bar{x}$. On the other hand, when $\sigma_n^2 >> \sigma_x^2$, we have a noisy observation vector b and we weight $\bar{x}$ heavily.

Expressions (149) and (155) gives us more insight than (130) in that the basic inversion operator is also modified in accordance with a priori information. Note the similarity in form of (155) and solution (47). The analogy between the various quantities in these two equations is also evident. Particularly is this so when the covariance matrix K_x is compared with H and the a priori mean $\bar{x}$ is compared with the a priori vector p. The fact that both of these approaches result in the same basic form further enhances the use of a priori information in the inverse operator, and it will be demonstrated that an essential part of the restoration process is the stability control offered by use of this information.

In order to show how the general case (149) reduces to the classical MSE estimate (139) and to introduce the problem of measurement selection, consider the case when there is essentially no contribution due to a priori information. In this case the estimate (149) may be written

$$\hat{x} = (A^t K_n^{-1} A)^{-1} (A^t K_n^{-1} b). \qquad (156)$$

This is known as the Markov estimate and corresponds to the minimization of the weighted least squares

$$R(x) = (Ax-b)\,K_n^{-1}\,(Ax-b) \tag{157}$$

where as before K_n is the noise covariance matrix (Deutch, 1965). When $K_n = \sigma_n^2 I$, the white noise case, the estimate in expression (156) reduces as it should to the classical least squares estimate of (139).

When the elements in K_x^{-1} are small compared with those in $A^t K_n^{-1} A$ the error (152) reduces to

$$MSE = tr\left[(A^t K_n^{-1} A)^{-1}\right]. \tag{158}$$

When $K_n = \sigma_n^2 I$ we have

$$MSE = \sigma_n^2\, tr\left[(A^t A)^{-1}\right], \tag{159}$$

or the equivalent form

$$MSE = \sigma_n^2 \sum_{i=1}^{N} \left|1/\lambda_i^2\right| \tag{160}$$

where the λ_i are the eigenvalues of the A matrix. This form (160) was also suggested by Barnes (1966b) as an important form to consider when studying the effects of diffraction when using the classical inversion proce s.

When K_n is specified in (158), only the A matrix can be varied to reduce the error. This is both a nontrivial and important problem. It is nontrivial because it is very difficult to solve for optimal conditions that will specify the A matrix for each restoration. It is important because, as will be shown, the MSE can be reduced considerably (by several orders of magnitude) by just knowing where to measure the image and where to predict the object.

Austin (1966) has been able to solve for the optimal condition in a special case. The special case is one in which the image is assumed to be the summation of equally spaced diffraction patterns from known point sources which are separated by the Rayleigh distance and $K_n = \sigma_n^2 I$. For this special case he showed that the optimum image measurement locations are above the known point source locations which are transferred to the image plane. In this case A = I, and the MSE becomes

$$MSE = \sigma_n^2 M. \tag{161}$$

In this paper we consider the more general and more important case since we consider here that the object is to be both unknown and continuous and that the diffraction effects are severe enough so that we must predict at points separated by a distance which, for most cases, is only a small fraction of the Rayleigh distance. To further specify the case we consider in this paper, we reiterate that the value of R is to be greater than unity. (The case Austin (1966) considered holds for $R \leq 1.0$. For $R = 1.0$ the number of point sources $M = 3.0$.) For $R > 1.0$, the optimum scheme for choosing the A matrix parameters has not been solved. However, in order to provide a rather complete computational procedure, the variation of the parameters which determine the A matrix, and thus the MSE has been extensively studied by numerical methods. Results of this study and the computational insight which they provide for the working restoration procedure are presented in the section entitled "MSE Variation--Choosing the A Matrix Parameters."

Perhaps the outstanding problem which remains to be solved is that of initially estimating the size of the diffracted object. This important problem is discussed further below.

Minimum Distance Estimation

Minimum MSE estimation considers the minimization of functional differences between the predicted object and the true object. The integral

equation, however, requires that we also choose other parameters before the vector x can be estimated. For example, M, N, the α_i's, the ξ_i's and the w_i must all be chosen before estimating x. Evidence for making intelligent choices for all of these parameters is discussed later. Here we present a minimum distance criterion which accounts for both the size and functional shape of the object.

Visually, diffracted objects appear smeared or spread out. The size of the true object is not known. To effect a better restoration than the MSE estimate we consider choosing α and x such that the following mean square distance is a minimum:

$$d^2 = E\left\{[\hat{x}(\hat{\alpha})-x(\alpha)]^t[\hat{x}(\hat{\alpha})-x(\alpha)] + [\hat{\alpha}-\alpha]^t[\hat{\alpha}-\alpha]\right\}. \tag{162}$$

The minimization of this expression implies choosing α and x together. No analytic results have been obtained for accomplishing this task. The problem thusfar has been that the size and object functional values are so intimately related that we cannot estimate one without knowledge of the other.

In the absence of an optimum size estimate, the object size and likewise the diffraction have been initially estimated by subtracting the point spread size from the image size. Using this initial size estimate,

we have then sequentially estimated the object size and shape together. However, it appears that we can do better, and future efforts undoubtedly should consider this problem.

SIMULATED OBJECT RESTORATIONS

In this section we present the results obtained from actual numerical computations performed on a digital computer which simulate the various restoration solutions developed previously. The major computer program used in performing these simulations is presented in the Appendix. All of the solutions which utilize a priori information in the inversion process have been investigated. The MSE solution with the additional computational feature of iteration has been used most extensively.

For all of the simulations we have considered the one-dimensional incoherent point spread function

$$h(\xi-\alpha) = \text{sinc}^2(\xi-\alpha) \tag{163}$$

where we have fixed the aperture, as previously discussed, so that the Rayleigh separation is unity. (Equation (163) describes an optical configuration with unity magnification. This involves no loss in generality since it is always possible to normalize the object-image coordinates so that the magnification is unity.) Using (163) the imaging equation is

$$b(\xi) = \int_{-a}^{a} \text{sinc}^2(\xi-\alpha)x(\alpha)d\alpha . \tag{164}$$

Results are presented on the interrelationship between the diffraction level (R), the noise level in the image and point spread function, the choice of parameters which govern the A matrix, and the amount and type of a priori information used. Also considered is a method of smoothing the image when we encounter excessive noise.

In this section the results of the interrelationship just mentioned are discussed as the section proceeds. For a summary of these interrelationships and the computational scheme which has been developed the reader is referred to the Summary and Conclusions section. The computational scheme is also discussed in the next section entitled "MSE Variation--Choosing the A Matrix Parameters."

Before presenting the simulations, it is necessary to present a clearer picture of the simulated restoration process by discussing three general areas. These areas consider the objects to be restored, the image and A matrix accuracy, and the choice of a restoration improvement criterion.

Objects to be Restored

Basically four objects have been considered. These are the uniform pulse,

$$x(\alpha) = \begin{cases} 1.0 \text{ for } -a \leq \alpha \leq a \\ 0 \text{ elsewhere} \end{cases} \qquad (165)$$

the smooth pulse,

$$x(\alpha) = \begin{Bmatrix} \cos k\pi\alpha \text{ for } -a \leq \alpha \leq a \\ 0 \text{ elsewhere} \end{Bmatrix} \tag{166}$$

two uniform pulses,

$$x(\alpha) = \begin{Bmatrix} 1.0 \text{ for } -a_1 \leq \alpha \leq a_1 \\ \qquad\quad -a_2 \leq \alpha \leq a_2 \\ 0 \text{ elsewhere} \end{Bmatrix} \tag{167}$$

and two smooth pulses,

$$x(\alpha) = \begin{Bmatrix} \cos k\pi(\alpha-\varphi) \text{ for } -a_1 \leq \alpha \leq a_1 \\ \qquad\qquad\quad -a_2 \leq \alpha \leq a_2 \\ 0 \text{ elsewhere.} \end{Bmatrix} \tag{168}$$

The motivation for considering these basic objects was as follows. First, recall that the solutions investigated utilize a smoothing parameter which in the MSE case depends upon the noise variance σ_n^2 and the a priori variance σ_x^2. Phillips (1962) has mentioned that his solution should work for $x(\alpha)$'s that are smooth. In order to gain an estimate of the loss involved when the objects are not smooth, various smooth ($\cos k\pi\alpha$) and non-smooth (the uniform pulse) objects have been chosen. Secondly, it is interesting to determine how well the restoration procedure can restore split sources. Such objects are certainly difficult to restore (perhaps the "worst case" would be a large number of split sources

which are very closely spaced), and they provide a more absolute criterion as to how well the restoration procedure performs. Furthermore, the resolution of split sources which have been excessively diffracted provides a good comparison with the classical Rayleigh criterion.

Image and A Matrix Accuracy

As has been previously emphasized, noise is the limiting factor in the restoration of optical objects of finite extent. Thus, in the simulations which are to be presented it is important that we specify the procedures used to obtain the image and A matrix and to provide measures as to the accuracy of these quantities. This section presents this information.

Obtaining the images

The images used for the restoration process were obtained by numerical integration using Simpson's quadrature. The motivation for using numerical methods was essentially twofold. The integrations for a variety of sources one may wish to consider cannot always be performed analytically. Also, since the restoration process is one of "inverse-integration," then it is natural that one would learn from the forward process, particularly in regard to the choice of the number of points for a desired accuracy. A third consideration is that a numerical procedure allows one to determine the image for arbitrary arguments, while the analytic image (if available) may be difficult to obtain for some arguments.

All of the images were noisy in the sense that only a finite number of significant figures were known. Specific bounds or limits were not derived which would enable us to explcitly state the exact number of significant figures used. However, the number of significant figures available for use in the restoration was estimated in each case.

In some cases the image accuracy available was limited only by the number of points chosen to approximate the integration. These images were accurate to at least 4 or more significant figures and are referred to as the "no-noise" images. To obtain images with less accuracy than the no-noise cases, the accurate images were perturbed by an additive random variable.

The determination of an estimate for the available significant figures and other measures of the amount of noise used in the various restorations are discussed below.

Image accuracy--"no-noise" cases

Here we define for convenience the quantity MSF as the minimum number of significant figures available in the image. This number (MSF) is recorded on all of the restorations. The estimation of the MSF for the various sources follows.

The images due to the uniform sour can be found in terms of the integral

$$b(\xi) = \int_{-a}^{a} \operatorname{sinc}^2(\xi-\alpha)\,d\alpha \qquad (169)$$

which is

$$b(\xi) = \frac{(\xi-a)\cos 2\pi(\xi+a) - (\xi+a)\cos 2\pi(\xi-a)}{2\pi^2(\xi+b)(\xi-a)}$$

$$+ \frac{1}{\pi}\left\{\mathrm{Si}[2\pi(\xi+a)] - \mathrm{Si}[2\pi(\xi-a)]\right\} \qquad (170)$$

where

$$\mathrm{Si}(x) = \int_{o}^{x} \frac{\sin u}{u}\,du.$$

Using (170), the numerical image accuracy from the computer was compared with the true value (170) for at least one image argument. The minimum number of significant figures tc which these two values agreed is defined as the MSF.

To obtain the images resulting from the smooth sources, one must perform the integration

$$b(\xi) = \int_{-a}^{a} \operatorname{sinc}^2(\xi-\alpha)\cos k\pi(\alpha-\varphi)\,d\alpha. \qquad (171)$$

This integration is not easily performed analytically, nor is an estimate of the error for Simpson's summation easily obtained.

For these images a rough one significant figure approximation was determined analytically which insured that the image values were of the right order of magnitude. As a further check, the number of increments (points) in the integration interval was increased while comparing the image values obtained. The MSF was recorded as the minimum number of significant figures to which the two numerical values agreed.

Image accuracy--additive noise cases

In the simulated restorations, the noisy image was obtained using the additive model

$$\mathbf{b} = \bar{\mathbf{b}} + \mathbf{n} \tag{172}$$

where n is a Gaussian random vector of mean zero and covariance matrix $\sigma_n^2 \mathbf{I}$. Thus, b is a Gaussian random vector with mean $\bar{\mathbf{b}}$ and covariance matrix $\sigma_n^2 \mathbf{I}$. In each case where noise was added, the true "no-noise" image value ($\bar{b}_i$) was known accurately and then perturbed by the random noise sample. Thus, when noise is not added, it is assumed that the MSF described previously are available for the restoration process.

It is desirable to determine a criterion which relates σ_n^2 to the number of significant figures available for various image strengths b_i.

Several criteria could be established to find such a relationship. The criterion chosen is explained below and is one which agrees closely with the actual simulated results.

Let the true value $\bar{b}_i = 1.0$, and consider choosing σ_n^2 such that

$$P[-\tfrac{1}{2} < n_i \leq \tfrac{1}{2}] = \tfrac{1}{2} \tag{173}$$

and

$$P[n_i \leq -\tfrac{1}{2}] + P[n_i > \tfrac{1}{2}] = \tfrac{1}{2}.$$

Thus

$$P[\tfrac{1}{2} < b_i \leq 1.5] = \tfrac{1}{2} \tag{174}$$

and

$$P[b_i \leq \tfrac{1}{2}] + P[b_i > 1.5] = \tfrac{1}{2}.$$

Now consider the following rounding procedure. (The image values were not actually rounded but the procedure is considered here purely for establishing a reasonable σ_n^2 vs. significant figure relationship.) Suppose we set $b_i = 1.0$ if $.5 < b_i \leq 1.5$. If $b_i \leq .5$ we set $b_i = 0.0$, and if $b_i > 1.5$ we set $b_i = 2.0$. Thus there is a probability of .5 that b_i will

have two significant figures and a probability of .5 that b_i will have no significant figures. In this case the mean number of significant figures is 1.

To find σ_n^2 we have

$$P[\tfrac{1}{2} \geq n > \tfrac{1}{2}] = \int_{-.5/\sigma_n}^{.5/\sigma_n} \frac{e^{-u^2/2}}{\sqrt{2\pi}}\, du = \tfrac{1}{2} \tag{175}$$

and $\sigma_n^2 = .55$. For 2 significant figures and $b_i = 1.0$, we find σ_n^2 such that

$$P[.05 \geq n_i > .05] = \tfrac{1}{2}$$

and

$$P[n_i \leq -.05] + P[n_i > .05] = \tfrac{1}{2}. \tag{176}$$

The noise variance is $\sigma_n^2 = .055$.

Figure 9 presents log $[1/\sigma_n^2]$ vs. the mean number of significant figures for the various image strengths used. The cross hatched region is the region used most frequently in this paper.

Since each image varies in strength as a function of its argument, then for σ_n^2 fixed the number of significant figures available varies over

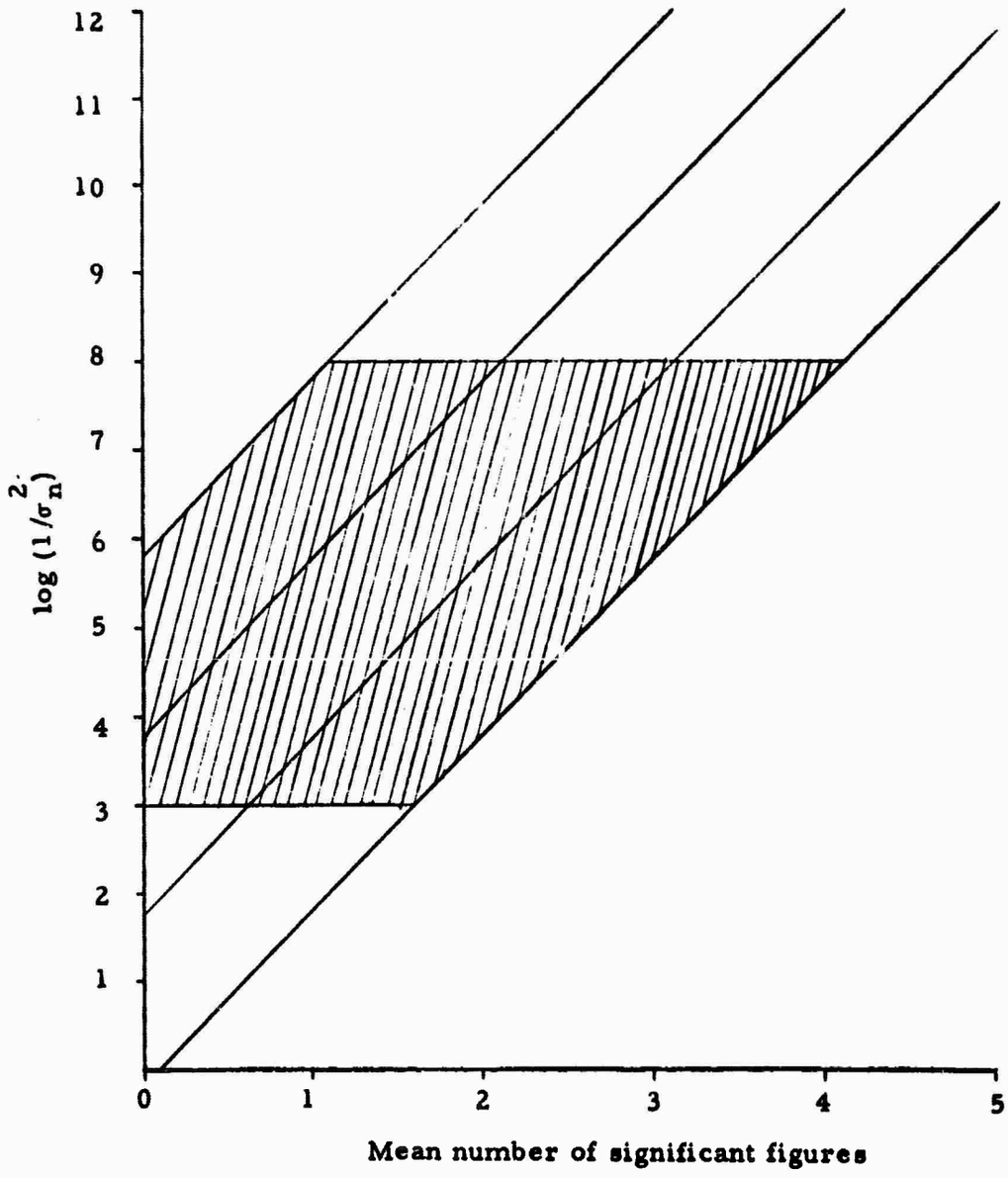

Figure 9. The log $(1/\sigma_n^2)$ vs the mean number of significant figures.

a range of values. This range corresponding to a fixed value of σ_n^2 is defined as SFR, and when $\sigma_n^2 > 0$ it is recorded on each restoration.

As a further measure of image noise level we use the following sample mean absolute error:

$$\epsilon_s = \sum_{i=1}^{N} \frac{|b_i - \bar{b}_i|}{N}. \tag{177}$$

To indicate the relative magnitude of the noise the following percentages are defined,

$$P_{max} = \frac{\epsilon_s \times 100}{\bar{b}_i\,(max)},$$

and (178)

$$P_{min} = \frac{\epsilon_s \times 100}{\bar{b}_i\,(min)}.$$

In (178) the sample mean error is computed as a percentage of the true "no-noise" maximum and minimum image measurements used in the restoration process. The reciprocals of these quantities are regarded as indications of the maximum and minimum SNR's used.

Obtaining the A matrix

In practice the point spread function could be either measured or analytically approximated. In either case some error is involved. However, it is well to keep in mind that the A matrix would be less noisy, for many applications, than the image because the point spread function could be accurately determined from many samples of the transfer function in a controlled low noise laboratory situation.

We have considered two general cases of A matrix accuracy, one which uses the full analytical accuracy of the computer to determine the A matrix, and the more important practical case when the A matrix is noisy. In the latter case several levels of noise were considered which exceeded the image noise.

A matrix accuracy

The additive model used in obtaining the noisy A matrix may be considered from two equivalent viewpoints. First suppose we have a noisy A matrix consisting of the true A matrix plus a B matrix composed of Gaussian random variables. Thus the image is

$$b = (A + B)x, \tag{179}$$

or

$$b = Ax + Bx. \tag{180}$$

Since Bx is a vector whose elements are linear combinations of Gaussian random variables, the elements of Bx are also Gaussianly distributed. Thus, Equation (180) may be alternatively written as

$$b = Ax + n \tag{181}$$

where n = Bx. This model is equivalent to the original additive model in Equation (114). Alternatively, if the A matrix or the point spread function is determined by measuring the image response to a point source, then we may write the imaging equation as

$$a_{ij} = \bar{a}_{ij} + n_i \tag{182}$$

where the $\bar{a}_{ij}$ are the true values and n_i is the additive background noise referred to previously.

To obtain the noisy A matrices Equation (182) was used with $n_i \sim N(0, \sigma_A^2)$ and $a_{ij} \sim N(\bar{a}_{ij}, \sigma_A^2)$. The $\bar{a}_{ij}$ values were the analytic values obtained by the computer. When $\sigma_A^2 = 0$ the full analytic accuracy of the computer (usually 8 significant figures) was used.

For the noisy A matrix cases we have chosen to indicate the level of noise by defining

$$\epsilon_A = \sum_{i=1}^{N} \sum_{j=1}^{M} \frac{|a_{ij} - \bar{a}_{ij}|}{NM} \tag{183}$$

as the sample mean absolute error between the random variable a_{ij} and the true analytic value $\bar{a}_{ij}$. Then we compute

$$P_{max_A} = \frac{\epsilon_A \times 100}{a_{ij}(max)}. \tag{184}$$

The minimum percentage is not computed because the $a_{ij}(min)$ values were often zero or very close to zero. Thus the effective SNR for the noisy A matrix elements ranges from a minimum of zero to a maximum of $1/P_{max_A}$.

Improvement Criterion

The choice of a measure which indicates restoration improvement over the initial object estimate in the absence of data processing (which we have assumed throughout is the image) will likely depend upon the application. In the absence of a specific application we consider two improvement measures, the ad-hoc or intuitative visual measure and a mathematical measure. Generally speaking, these two measures should agree, at least on a qualitative basis.

In many problems one can assess mathematical improvement by using the "standard" squared error (SE) criterion which is

$$SE = \sum_{i=1}^{M} (\hat{x}_i - x_i)^2 \qquad (185)$$

where $\hat{x}_i$ is the estimate resulting from the restoration process and x_i is the corresponding value of the true object. We have defined the general term squared error (SE) so that the term for mean square error (MSE) can be used exclusively as defined previously. The major difference between the MSE and the SF as shown above is that when σ_x^2 is finite we are inferring the possession of a priori information which we do not actually have. If the prior guess for the object is the image, then a finite σ_x^2 will weight the image so that the estimate will, roughly speaking, lie somewhere between the image and the true object, but because σ_x^2 is finite we do not know the exact error involved.

To continue the discussion, it was evident that visual improvement could not be assessed only by SE improvement. In fact, as will be illustrated, there were several cases when visual improvement was apparent, but there was actually a decrease in SE improvement. The difference is due to the fact that the SE does not account for object size. The squared distance (162), on the other hand, does include size error, and it agreed more closely with the intuitative visual error. Because of this agreement we have chosen to indicate mathematical improvement by computing the following ratio

$$I(d^2) = \frac{\sum_{i=1}^{M} [(b_i - x_i)^2 + (\xi_i - \alpha_i)^2]}{\sum_{i=1}^{M} [(\hat{x}_i - x_i)^2 + (\hat{\alpha}_i - \alpha_i)^2]} \qquad (186)$$

where the x_i's and α_i's are, respectively, the true object values and arguments and M = N. Mathematical improvement is shown if $I(d^2)$ exceeds unity.

Of course there is some arbitrariness as to the choice of parameters in (186). For b_i, ξ_i, and M the values were taken as those actually used. The x_i and α_i were chosen such that the α_i were equally spaced in the true nonzero source interval. The $\hat{\alpha}_i$ and corresponding $\hat{x}_i$ were chosen such that the estimated object was reasonably represented in the estimated nonzero source interval. To clarify this last statement it was assumed, on the basis of a posteriori information, that the regions where the estimated object was negative or small positively were regions where the estimated object intensity was zero.

The presentation of $I(d^2)$ on each restoration will allow a more valid assessment of improvement over the image. Unless otherwise noted, the $I(d^2)$ shown on the restorations is the improvement of the last iteration over the image.

It is well to remark that there are two general categories of simulated restorations: those which assume essentially no prior size information and those which do assume some prior size information. In the former

case $I(d^2)$ will present a valid improvement criterion. However, for the latter case much of the $I(d^2)$ improvement may be due to the fact that the term $(\alpha_i - \alpha_i)^2$ is near zero while $(\xi_i - \alpha_i)^2$ is still large. That is, for the latter case it may well be that much of the $I(d^2)$ improvement is due to the assumed size information. Throughout the results which follow we have been careful to emphasize those cases in which prior size information is assumed so that the reader can more critically assess the restoration improvement.

Results of the Various Restoration Schemes

The solutions investigated by actual simulations are identified as follows:

Solution Description	Equation Number or Reference
MSE	(155)
Phillips-Twomey	(44)
Dynamic Programming--Area	(78)
Dynamic Programming--Prior Vector	(Bellman, 1965)
Matrix version of the area constraint	(66)

To identify which solution was used and the various parameters involved, the following information block (Figure 10) is placed on each restoration.

Run								
M	N	Quad	RI	R	λ	Area	Prior	Iter
σ_n^2	σ_x^2	P_{max}	P_{min}	SFR	MSF	σ_A^2	$Pmax_A$	
$I(d^2)$								

Figure 10. Information block.

When the symbol NA appears in place of any of the quantities in Figure 10, it means that the quantity does not apply for the case in question.

Run is an identifying number. The first row of quantities under Run specifies the measuring and predicting scheme used and the solution used. The next row indicates the image and A matrix accuracy and the a priori variance σ_x^2 when σ_n^2 is specified. The last row indicates the mathematical measure of improvement. The quantities M, N, λ, R, σ_n^2, P_{max}, P_{min}, SFR, MSF, σ_A^2, P_{max_A}, and $I(d^2)$ have been explained. Further explanation of these parameters, as applicable, and those not previously mentioned is presented below.

The following explanation refers to the parameters Area, Prior, and Iter, which may be used to identify the type of solution used. To avoid any possible confusion, the solution description is also recorded in each of the figure captions. When Area is greater than zero it is the true area under the object and indicates that the area constraint was used. When Prior is greater than zero then the prior vector constraint was used. If Prior = 2.0 then the image was used as the initial prior vector. When Prior and Area are both zero then the Phillips-Twomey smoothing matrix was used. The quantity Iter is the maximum number of iterations used in the iteration procedure. In the cases where Iter > 0 and the Phillips-Twomey smoothing constraint or the matrix version of the area constraint

were used, this is an indication that only the initial prior vector was determined by these methods. Further iteration used the MSE solution (155).

The variance σ_x^2 specifies the amount of pseudo-priori information assumed in the smoothing ratio σ_n^2/σ_x^2. Figure 11 shows σ_n^2 vs. σ_x^2 for various σ_n^2/σ_x^2 ratios. The crosshatched region is that used most frequently in this paper.

The ratio

$$\text{RI} = \frac{\text{Measuring interval width}}{\text{Predicting interval width}} \tag{187}$$

relates pertinent information about the measuring and predicting scheme. The proper choice of this ratio, as will be demonstrated, can significantly alter the MSE in the restoration process. In order to differentiate between the cases when one prediction or two prediction intervals were used for the same level of diffraction R, the prediction interval width in the latter case is taken as the sum of the two intervals.

Three different quadrature weights were investigated. When Quad = 1.0 the most extensively used, and most powerful, Gauss quadrature is indicated; Quad = 2.0 corresponds to the use of Simpson's quadrature, and Quad = 3.0 refers to the case when the weights were unity.

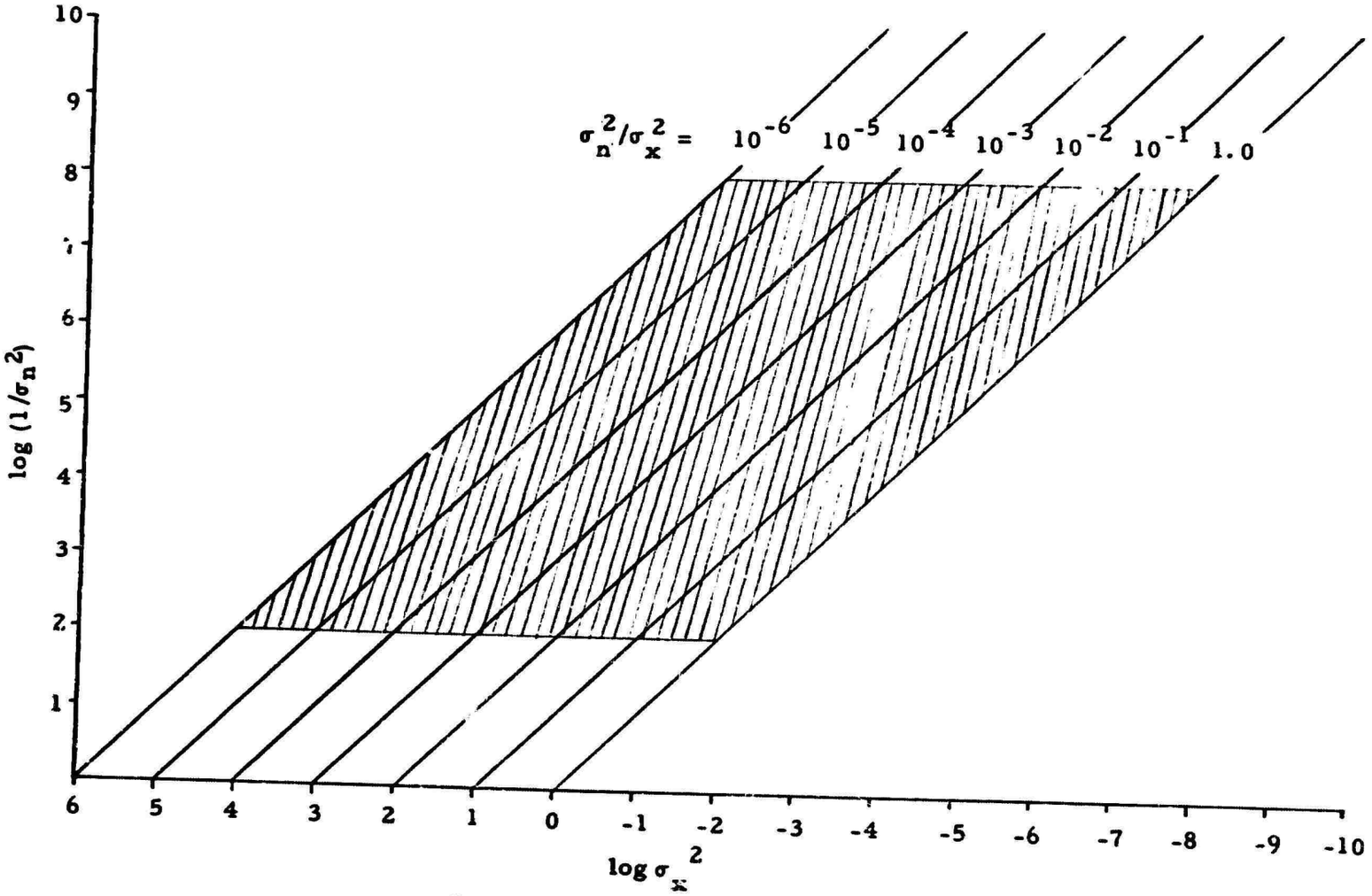

Figure 11. The log $(1/\sigma_n^2)$ vs log σ_x^2 for various ratios of σ_n^2/σ_x^2.

Effect of prior information in the inverse operator

As an introduction to the computational work a sequence of figures is shown illustrating the importance of using finite σ_x^2 in the inverse operator.

Figure 12 shows the restoration of a uniform pulse by operating on a low noise image (MSF = 5.0) with the straightforward matrix inverse (32). The estimated object is in error because more measurements are needed in the restoration procedure to adequately describe the object. That is, the restoration is quadrature-error limited.

To reduce this error the number of points in the [-1, 1] interval was increased from 5 to 7. Figure 13 shows the results, and the improvement is noticeable. Now the question to consider in further improving the restoration is the upper limit of the number of points one can use in the interval [-1, 1]. The answer to this question is that the number of points for reasonable restoration depends upon the system noise. In fact the MSE (159) depends upon the number of points one uses and increases rapidly as this number increases. If the system noise σ_n^2 is not small enough to overcome this increase in dimensionality, then poor restoration results. To illustrate this, the noise level of Figure 13 was increased. The noise variance was raised to 2.5×10^{-3} which resulted in a SFR of 1.4 - 1.7, while the previous image accuracy was at least 5 significant figures.

Run = 1								
M = 5	N = 5	Quad = 2	R1 = 1	R = 1	λ = 0.0	Area = 0	Prior = 0	Iter = 0
$\sigma_n^2 \approx 0$	NA	NA	NA	NA	MSF = 5	$\sigma_A^2 \approx 0$	NA	
I (d^2) = 2.21								

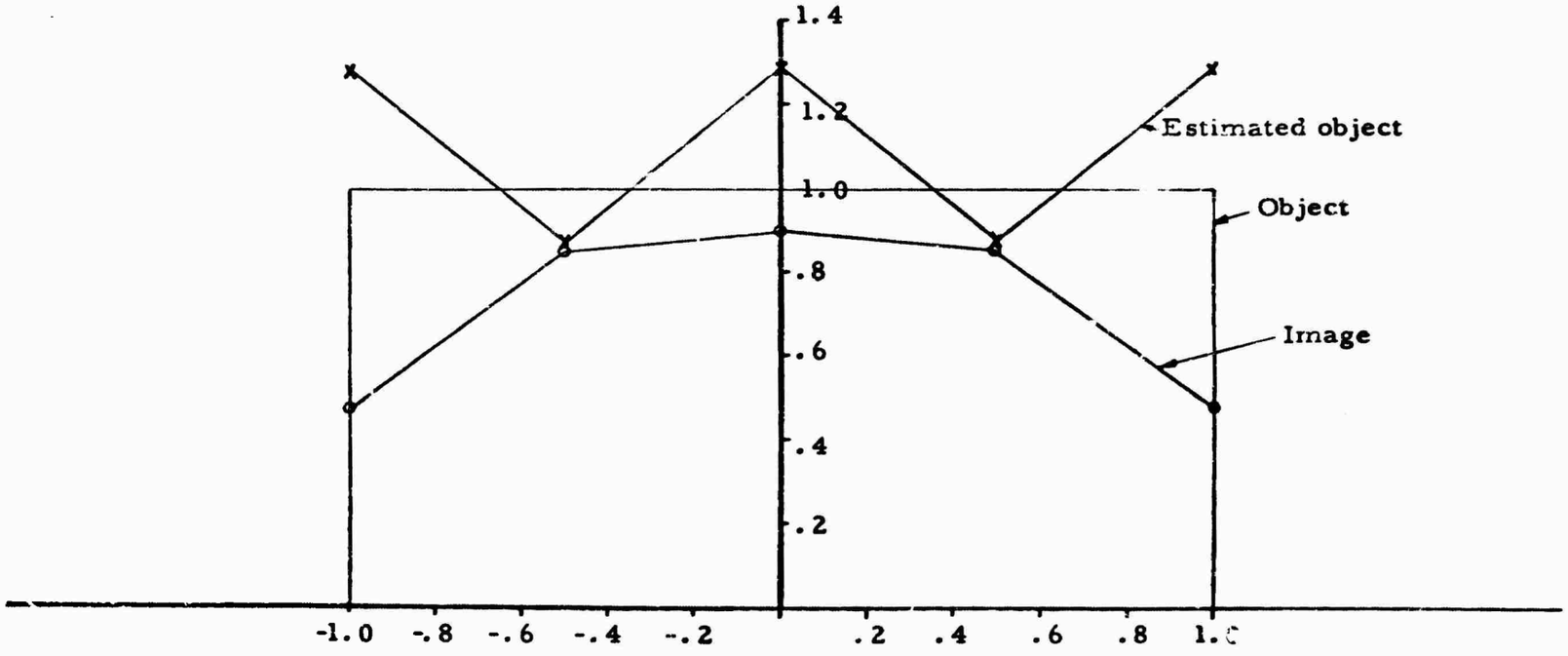

Figure 12. Restoration of a uniform pulse (R = 1.0) using (32).

Run = 4.0								
M = 7	N = 7	Quad = 2	R1 = 1	R = 1.	λ = 0.0	Area = 0	Prior = 2	Iter = 0
σ_n^2 = 0	NA	NA	NA	NA	MSF = 4	σ_n^2 = 0	NA	
I (d^2) = 190.78								

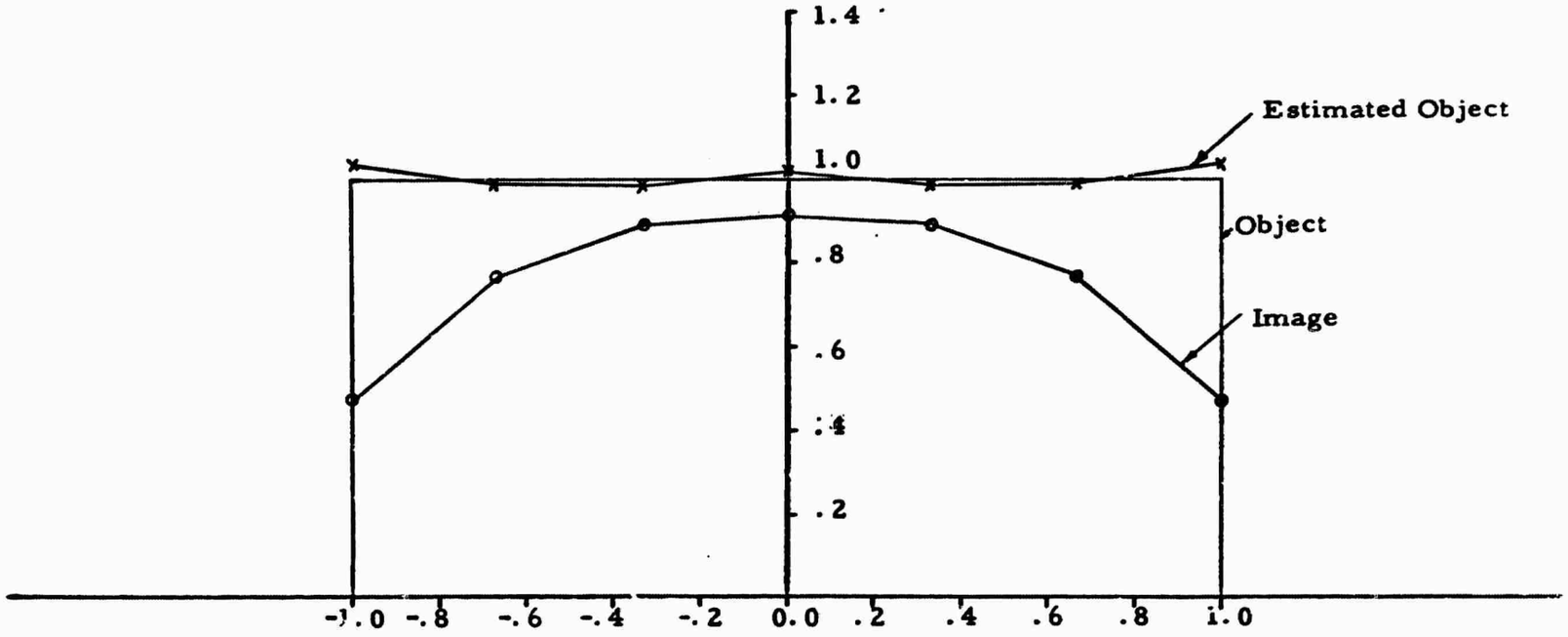

Figure 13. Restoration of a uniform pulse (R = 1.0) using (32).

Operation on the noisy image, in the same manner as before (32), produced the oscillatory object estimate shown in Table 1 below. This estimated object is meaningless. The image noise is now the limiting factor. A similar result in terms of an oscillatory solution was obtained by increasing the dimensionality to 9 and using the "no noise" image of MSF = 5.0.

The assumption of finite prior information can alleviate the oscillation. Consider Figure 14 where we have assumed $\sigma_x^2 = 2.5 \times 10^{-2}$ and used the Phillips-Twomey solution. The reduced SE in Figure 14 over the values shown in Table 1 confirms the usefulness of a priori information in the inverse operator.

Table 1. Estimated object and arguments for infinite σ_x^2.

$\hat{x}$	$\hat{\alpha}$
-300	-1.0
+200	-.67
-700	-.33
+400	0
-700	.33
+200	.67
-300	1.0

Run = 4.0								
M = 7	N = 7	Quad = 2	R1 = 1.0	R2 = 1.0	$\lambda = 10^{-3}$	Area = 0	Prior = 0	Iter = 0
$\sigma_n^2 = 2.5 \times 10^{-3}$	$\sigma_x^2 = 2.5 \times 10^{-2}$	Pmax=9.4	Pmin=18.1	SFR=1.4 -1.7	MSF=4	$\sigma_A^2 = 0$	NA	
$I(d^2) = 13.05$								

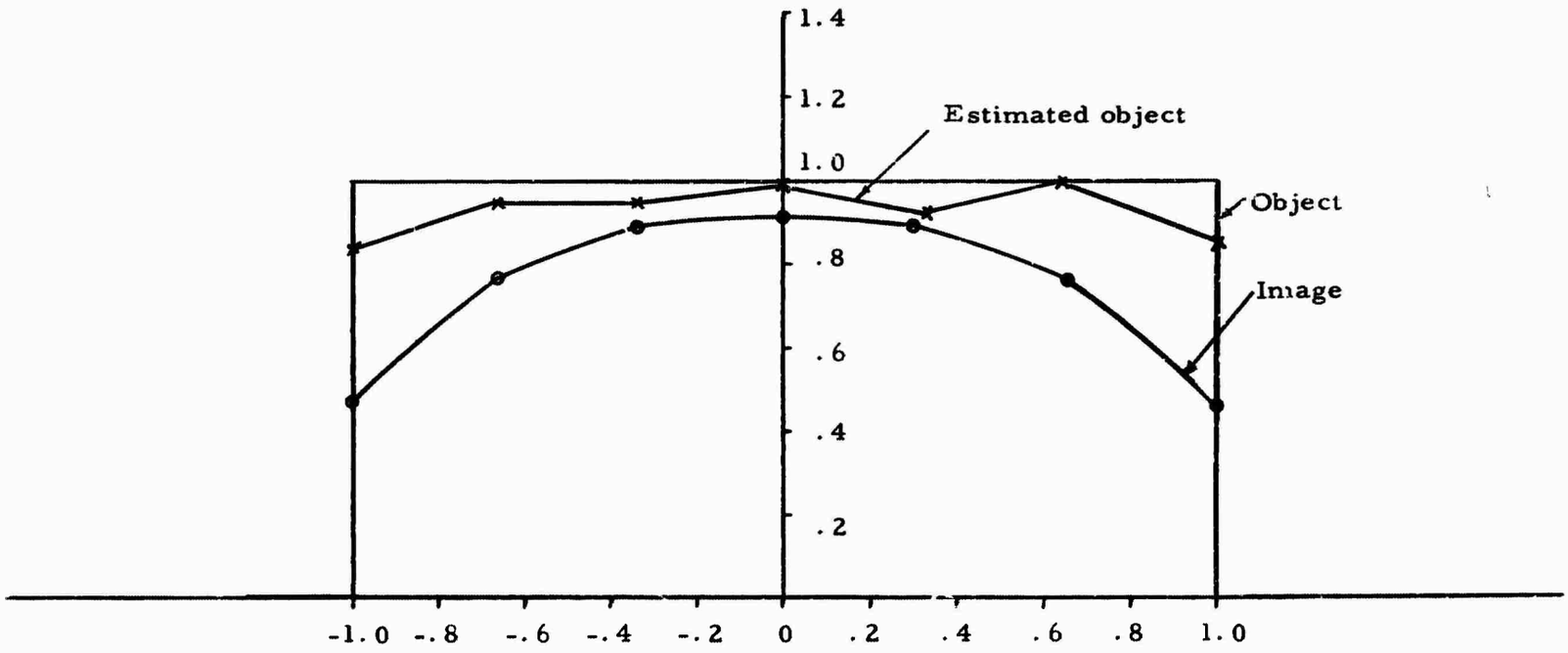

Figure 14. Restoration of a uniform pulse (R = 1.0) using (44).

The sequential estimation of a uniform pulse

Here we present results on the estimation of the size and shape of a uniform pulse which has been diffracted 5 times greater than the level corresponding to the Rayleigh resolution criterion for two point sources (R = 10.0). The importance of these figures is that they demonstrate the sequential estimation of an object which has been excessively diffracted, for the case when we have assumed virtually no a priori information about the true object.

Figures 15-20 present the series. The initial size in Figure 15 [-1.0, 1.0] was assumed larger than one would estimate on the basis of comparing widths of the point spread function and image. However, notice that the size is substantially reduced from [-1.0, 1.0] to [-.46, .46]. Using the updated size estimate the next figure shows further improvement in size along with shape improvement. This trend continues until Figure 19, which no longer indicates noticeable size improvement. No attempt was made to "push" the results to a limit by iterating further. However, the impending improvement in size and shape is apparent.

When the size is assumed smaller than [-.2, .2], Figure 20 shows that the end-points of x tend toward larger values than the true object. This suggests that the actual size limit of the restoration process has been exceeded, at least for the particular value of σ_x^2 used.

Run = 14.0								
M = 32	N = 32	Quad = 1	R 1 = 1.0	R = 10	λ = .01	Area = 0	Prior = 0	Iter = 5
$\sigma_n^2 \approx 0$	NA	NA	NA	NA		MSF = 5	$\sigma_A^2 \approx 0$	NA
$1(d^2) = 1.95$								

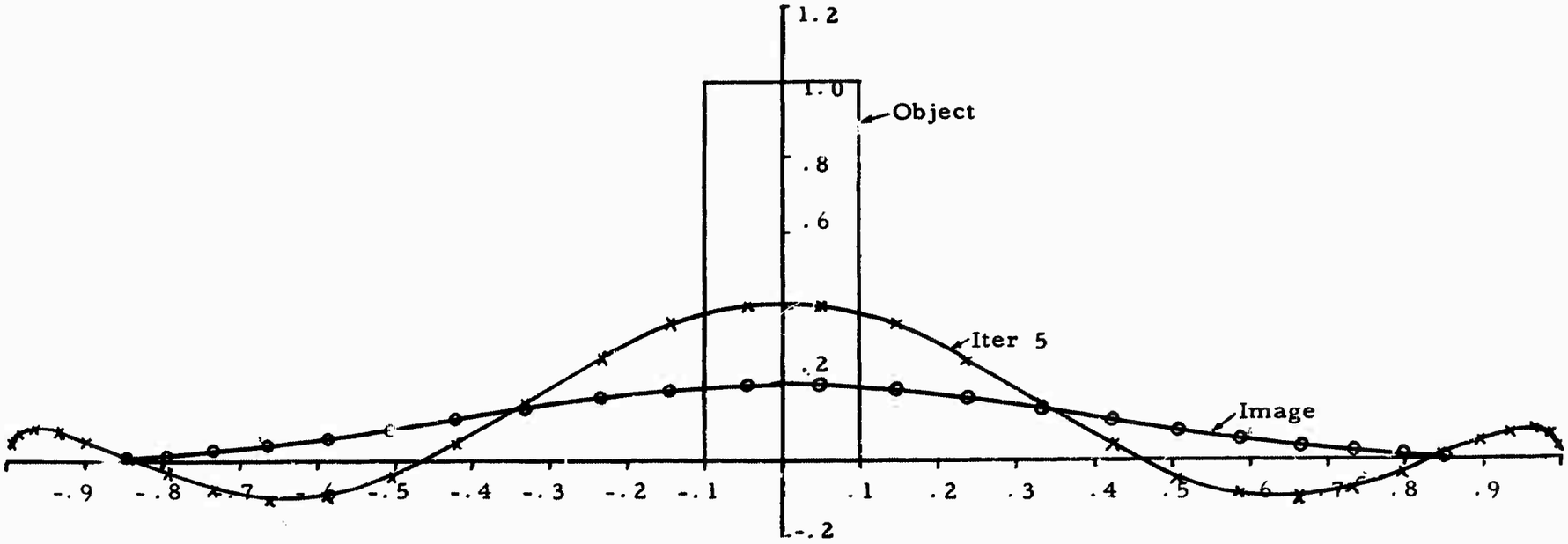

Figure 15. Restoration of a uniform pulse (R = 10.0) using (44)

Run = 14.4								
M = 48	N = 48	Quad = 1	R1=1.52	R = 10	λ = .01	Area = 0	Prior = 2	Iter = 10
$\sigma_n^2 \approx 0$	NA	NA	NA	NA	MSF=5	$\sigma_A^2 \approx 0$	NA	
$I(d^2) = 1.88$								

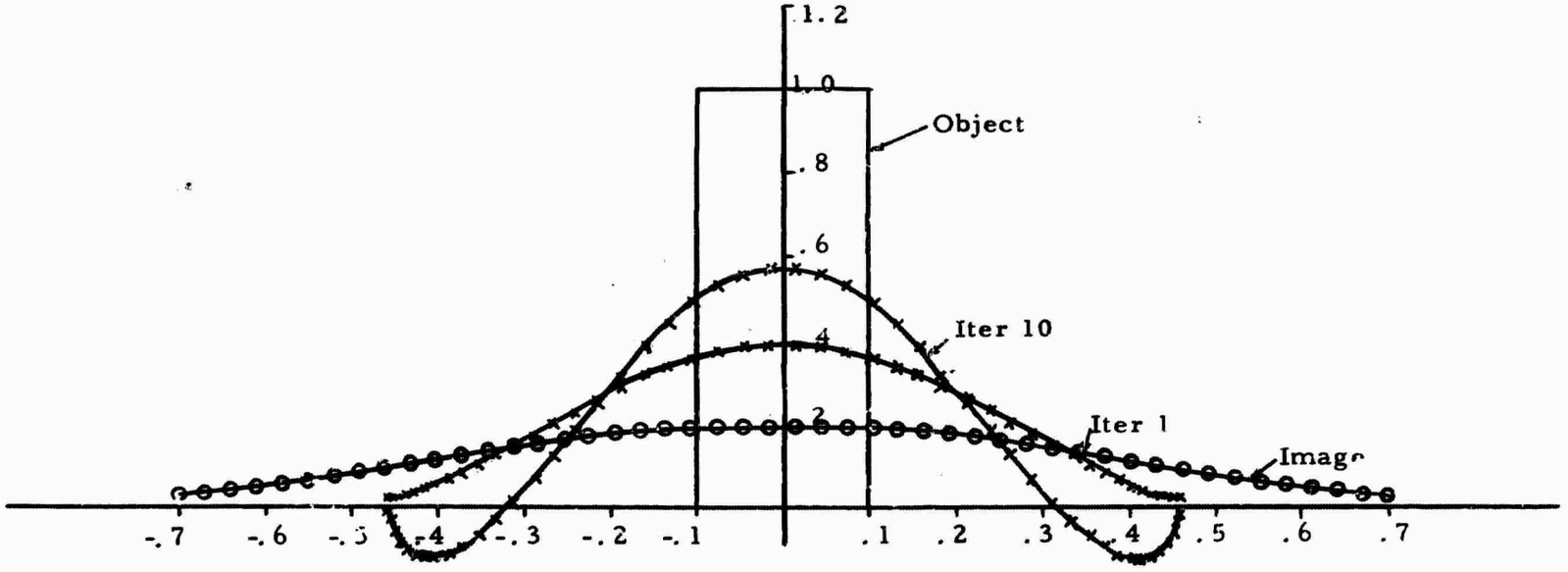

Figure 16. Restoration of a uniform pulse (R = 10.0) using (155) with $\bar{x}$ = b.

Run = 14.41								
M = 32	N = 32	Quad = 1	R1=2.26	R = 10	λ=.001	Area = 0	Prior = λ	Iter = 30
$\sigma_n^2 \approx 0$	NA	NA	NA	NA		MSF=5	$\sigma_A^2 \approx 0$	NA
$I(d^2) = 2.62$								

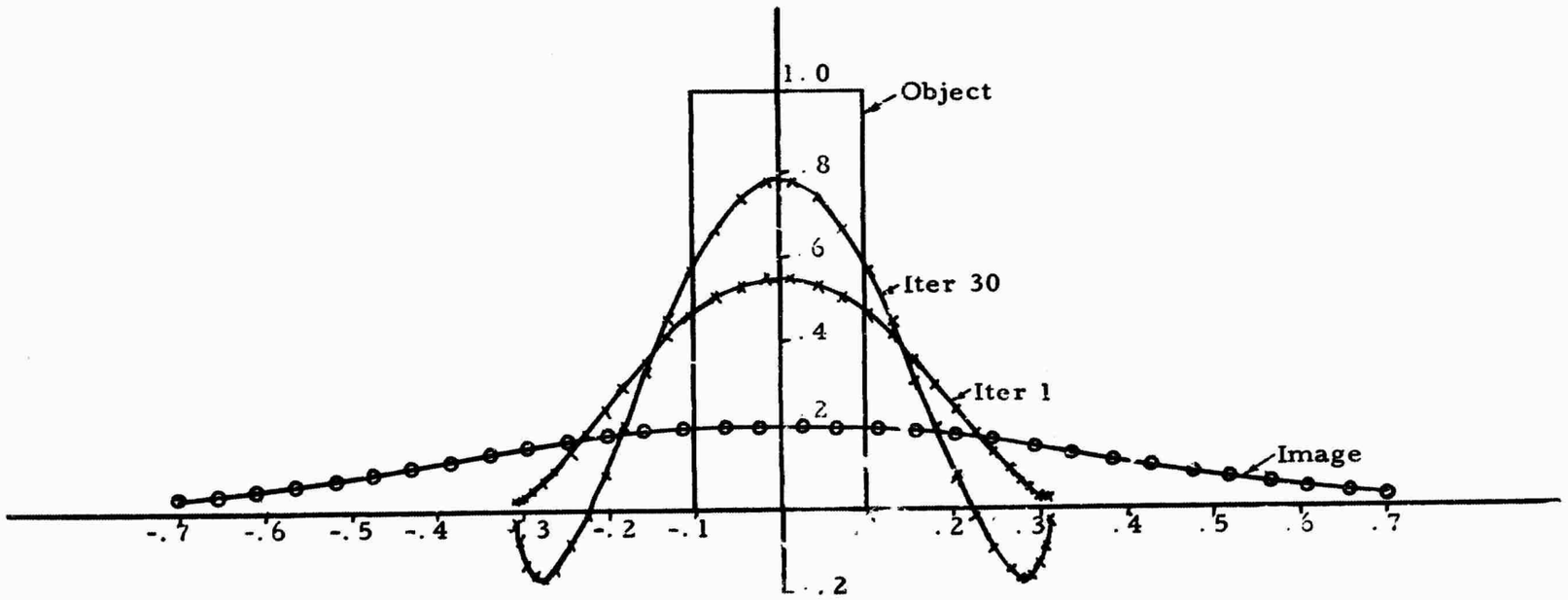

Figure 17. Restoration of a uniform pulse (R=10.0) using (155) with $\bar{x} = b$.

Run = 14.5								
M = 32	N = 32	Quad = 1	R1 = 2	R = 4	λ = .001	Area = 0	Prior = 2	Iter = 30
$\sigma_n^2 \approx 0$	NA	NA	NA	NA	MSF = 5	$\sigma_A^2 \approx 0$	NA	
I (d^2) = 2.45								

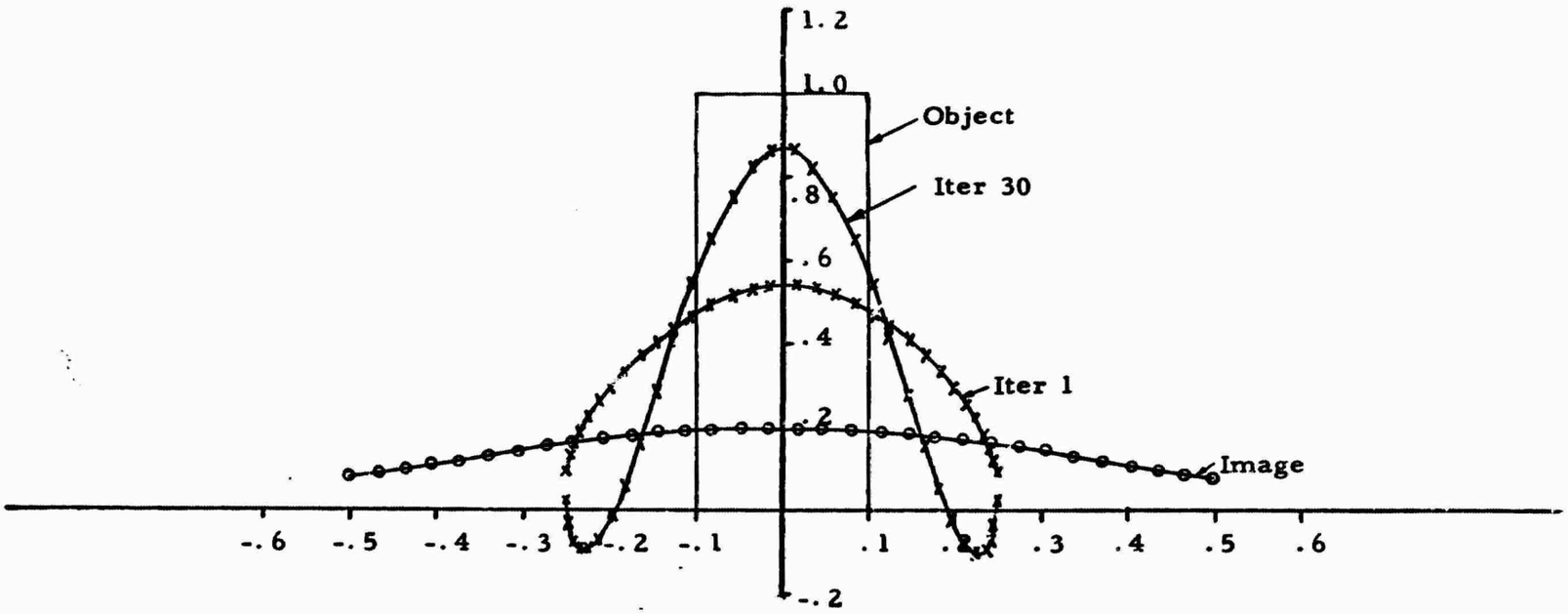

Figure 18. Restoration of a uniform pulse (R = 10.0) using (155) with $\bar{x}$ = b.

Run = 14.51								
M=32	N=32	Quad = 1	R1=3.7	R = 10	λ = .001	Area = 0	Prior = 2	Iter = 30
$\sigma_n^2 \approx 0$	NA	NA	NA	NA		MSF= 5	$\sigma_A^2 \approx 0$	NA
$I(d^2) = 2.63$								

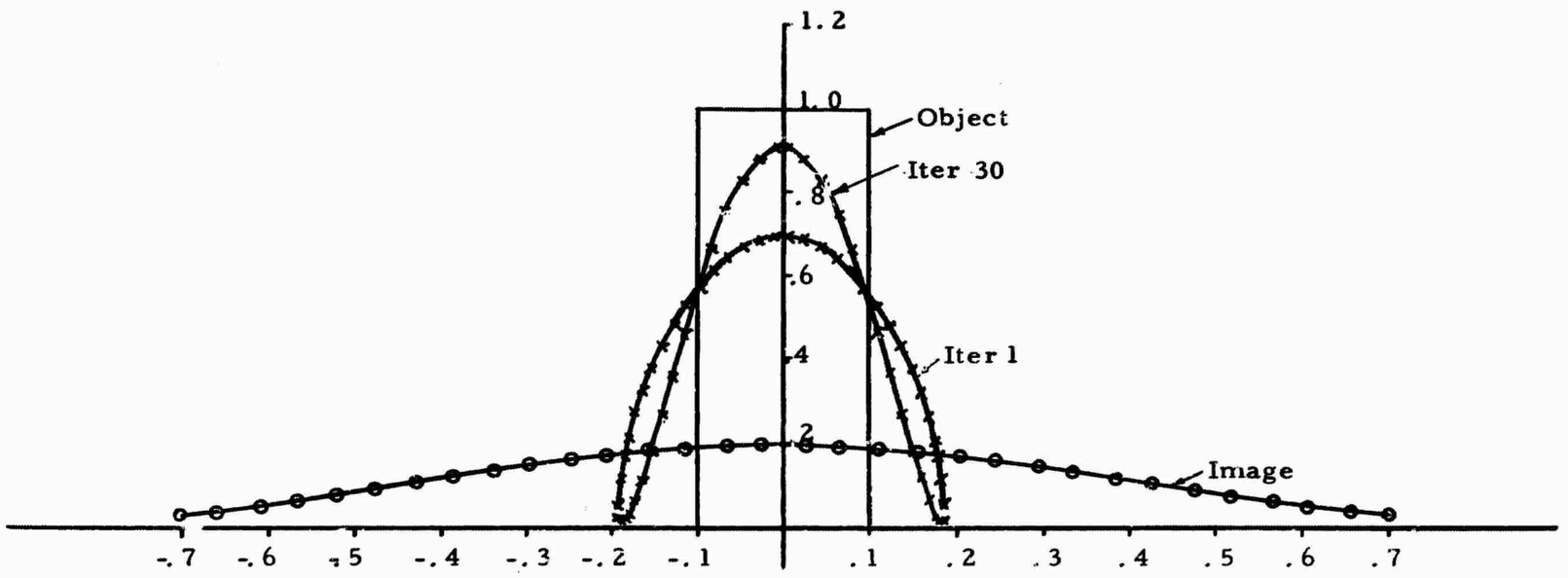

Figure 19. Restoration of a uniform pulse (R = 10.0) using (155) with $\bar{x}$ = b.

Run = 14.6								
M = 32	N = 32	Quad = 1	R1 = 4.2	R = 10	λ = .001	Area = 0	Prior = 2	Iter = 20
$\sigma_n^2 \approx 0$	NA	NA	NA	NA		MSF=5	$\sigma_A^2 \approx 0$	NA
I (d^2) = 7.10								

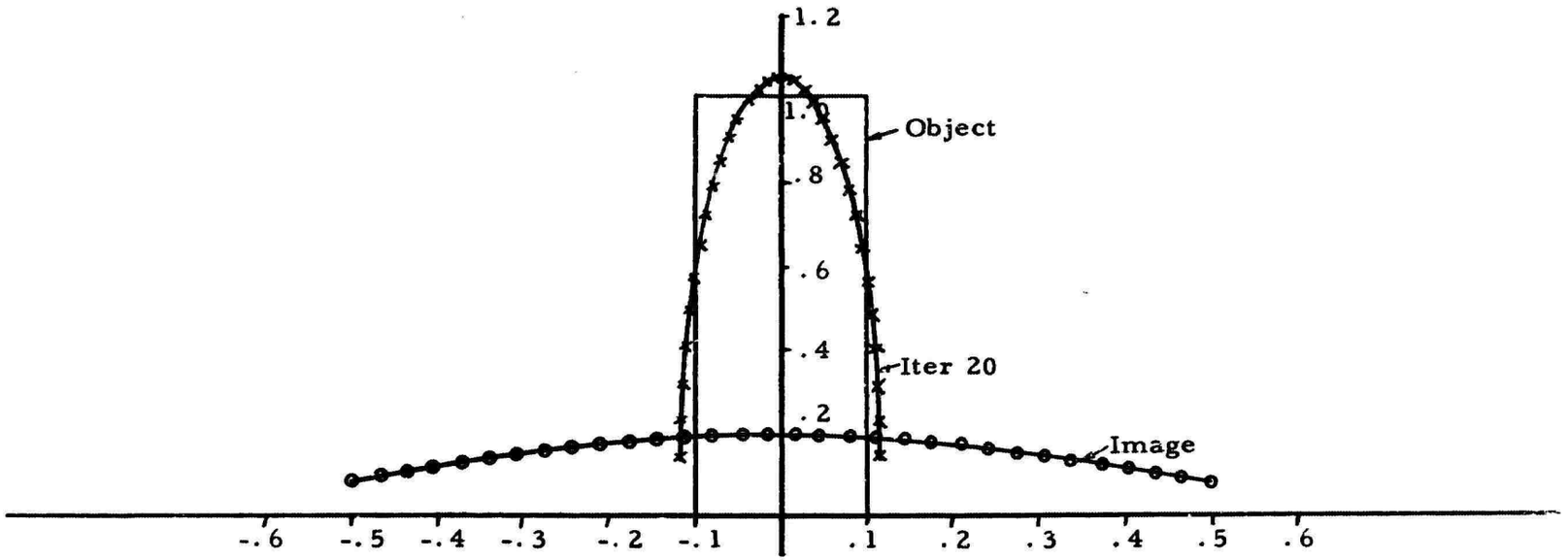

Figure 20. Restoration of a uniform pulse (R = 10.0) using (155) with $\bar{x}$ = b.

The sequential estimation of a smooth pulse

Results are presented here for the restoration of smooth cosine pulses. These restorations may be compared visually with those presented in Figures 12-20 to assess the loss incurred when the objects are not smooth but have sharp corners. As in the last series (Figures 15-20), these figures demonstrate restoration when no a priori object information is used.

Figure 21 shows that the source cos ($\pi\alpha/2$) having undergone a diffraction such that R = 1.0 can be very closely restored. Since the diffraction was small only one size estimate was necessary.

The estimation of the source cos ($2\pi\alpha$) is shown in Figures 22-26. Both the diffraction and image noise are greater than in Figure 22, but the final restoration Figure 26 is comparable to that of Figure 21. This suggests that σ_n^2 in Figure 21 could have been increased without significantly degrading the restoration.

Figures 25 and 26 are identical except for the prior weighting. Figure 26 ($\sigma_x^2 = 10^{-2}$) shows an increase in restored detail in both the first and 20th iteration. However, iteration 20 is asymmetrical. This feature is visually noticeable at the interval endpoints and occurs due to computational effects within the computer. A computational alternative is to choose a lower value of σ_x^2, begin with a more stable initial solution, and iterate longer. Here again, the solution eventually becomes unstable.

Run = 20.5								
M = 24	N = 24	Quad = 1	R_1 = 1	R = 4	λ = .001	Area = 0	Prior = 2	Iter = 10
$\sigma_p^2 = 1 \times 10^{-6}$	$\sigma_x^2 = 1 \times 10^{-3}$	P_{max} = .1	P_{min} = 2.3	SFR = 1.6-2.9	MSF = 4	$\sigma_A^2 \approx 0$	NA	
I (n^2) = 5.92								

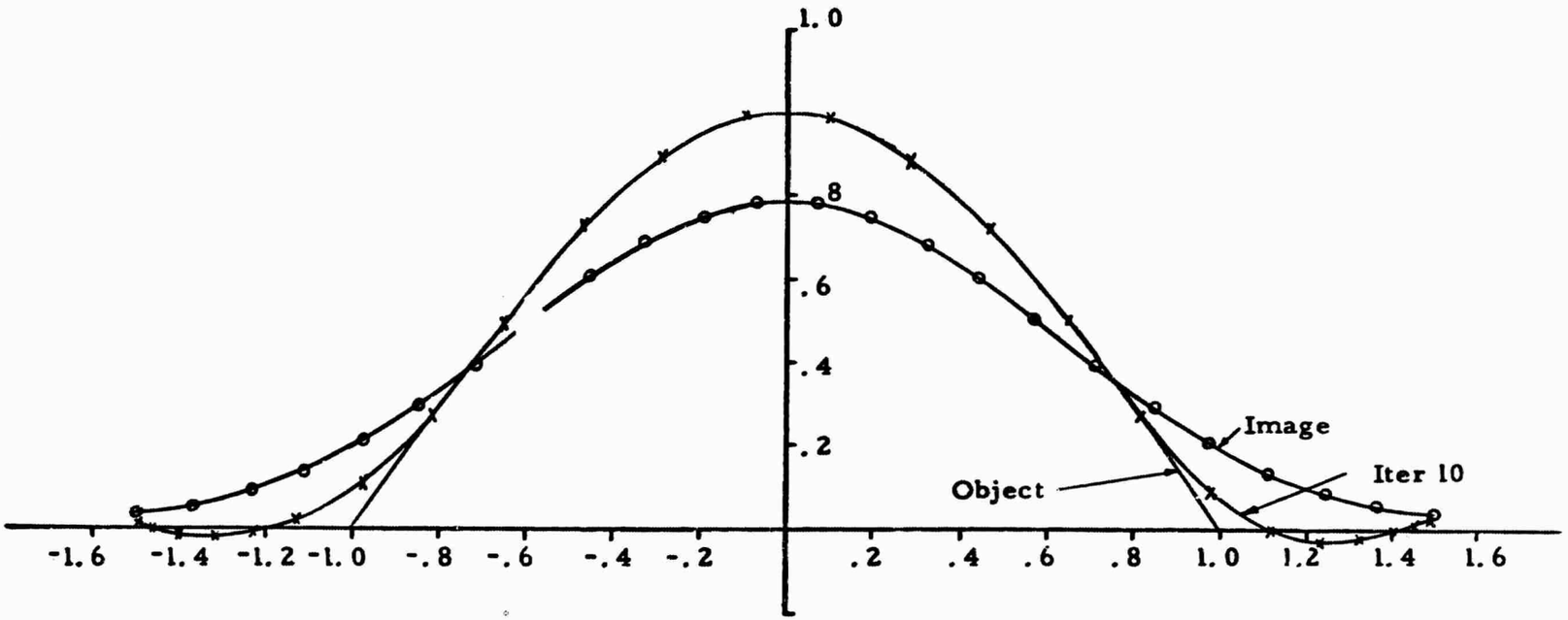

Figure 21. Restoration of the pulse cos ($\pi\alpha$ /2) using (155) with $\bar{x}$ = b.

Run = 20.00								
M = 24	N = 24	Quad = 1	R1 = .8	R = 4	λ = .0001	Area = 0	Prior = 2	Iter = 10
$\sigma_n^2 \approx 0$	NA	NA	NA	NA	MSE= 4	$\gamma_A^2 \approx 0$	NA	
I (d^2) = 4.57								

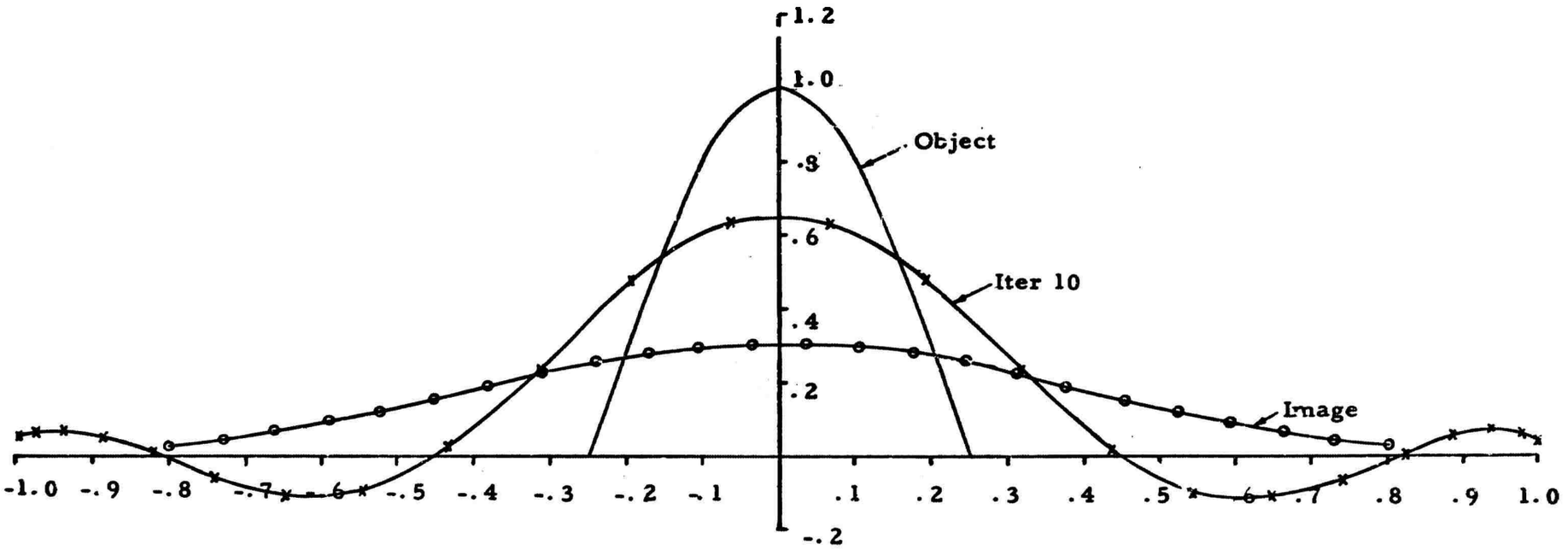

Figure 22. Restoration of the pulse cos (2 π α) using (155) with $\bar{x} = b$.

Run = 20.20								
M = 24	N = 24	Quad = 1	R1 = 1.4	R = 4	λ = .01	Area = 0	Prior = 2	Iter = 10
$\sigma_n^2 = 1 \times 10^{-5}$	$\sigma_x^2 = 1 \times 10^{-3}$	$P_{max} = .84$	$P_{min} = 5.1$	SFR = 1.3-2.1	MSF = 4	$\sigma_a^2 = 0$	NA	
$I(d^2) = 11.33$								

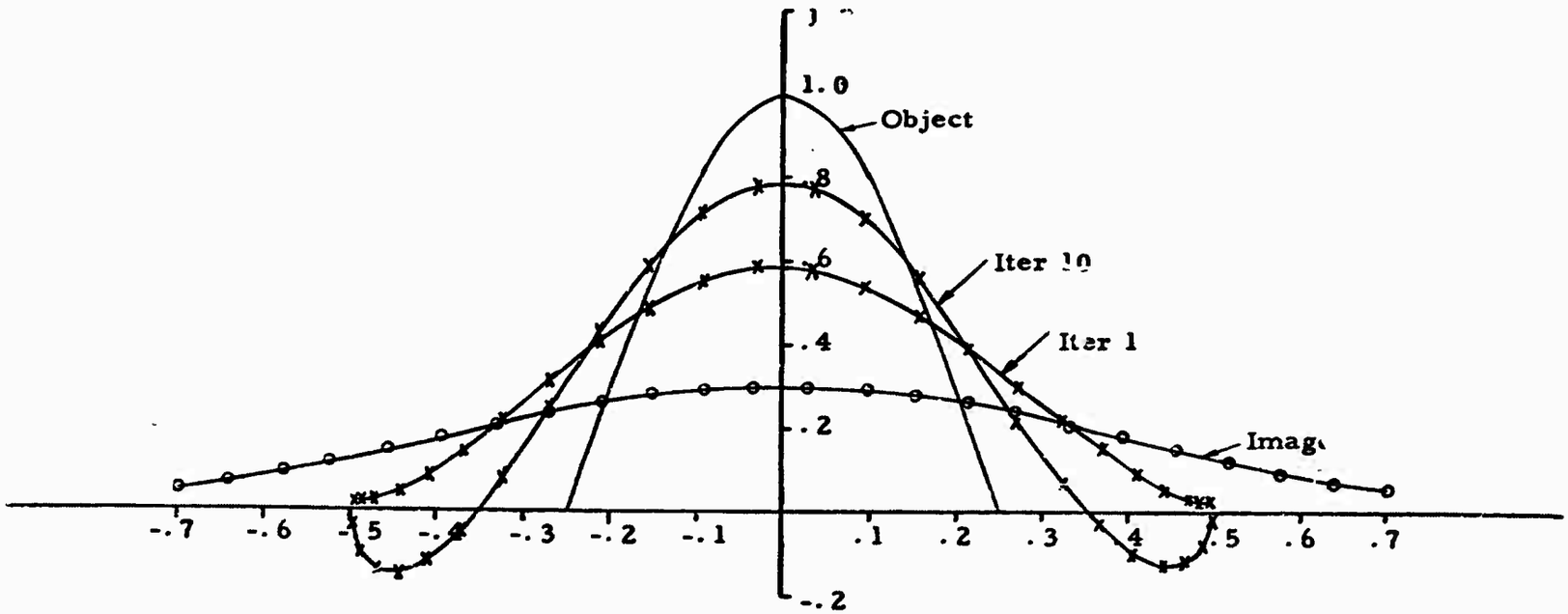

Figure 23. Restoration of the pulse cos (2 π α) using (155) with $\bar{x} = b$.

Run = 24.00								
M=24	N=24	Quad = 1.	R1=2	R = 4	λ =01	Area = 0	Prior = 2	Iter = 20
$\sigma_n^2 = 1 \times 10^{-5}$	$\sigma_x^2 = 1 \times 10^{-3}$	P_{max}=.84	P_{min} = 5.1	SFR=1.3-2.1		MSF=4	$\sigma_A^2 = 0$	NA
$I(d^2)$ = 51.11								

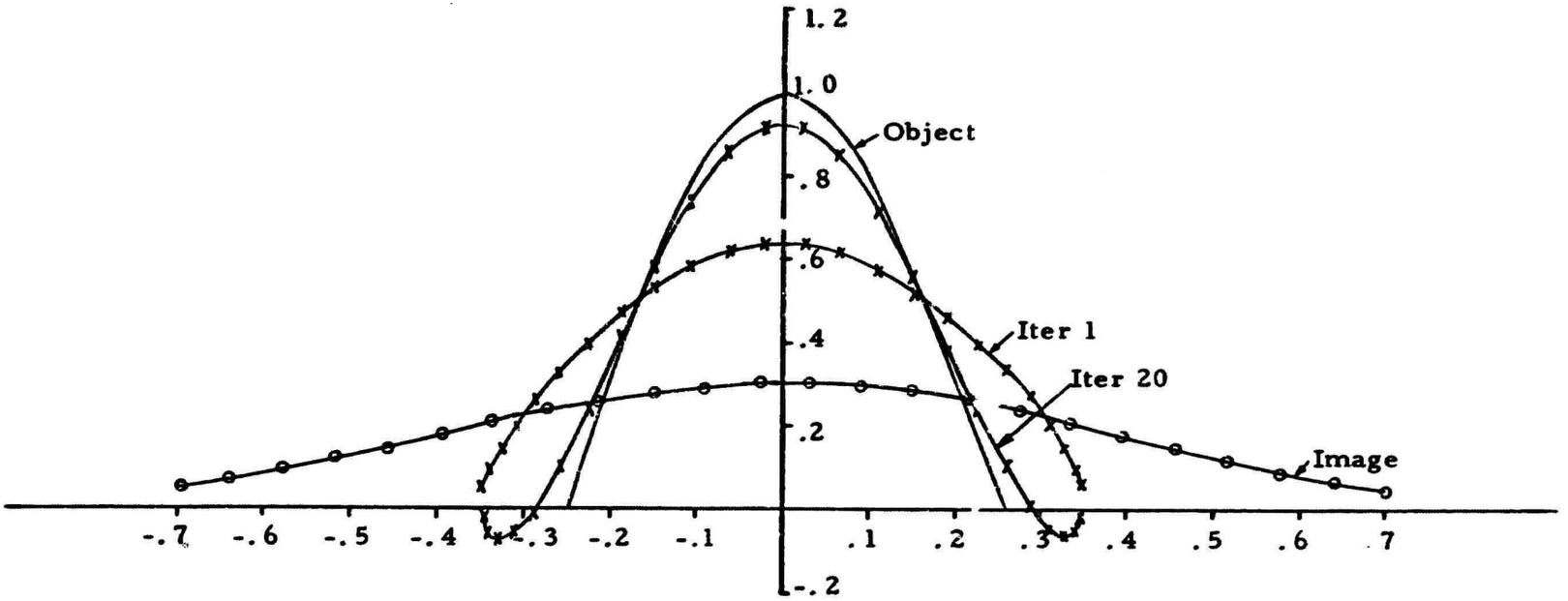

Figure 24. Restoration of the pulse $\cos(2\pi\alpha)$ using (155) with $\tilde{x} = b$.

Run = 24.10								
M=24	N=24	Quad = 1	R1=2.6	R=4	λ =.01	Area = 0	Prior 2	Iter = 20
σ_n^2=1x10^{-5}	σ_x^2 = 1x10^{-3}	Pmax=8.4	Pmin=5.1	SFR=1.3-2.1		MSF=4	σ_A ≈ 0	NA
I(d^2) = 64.30								

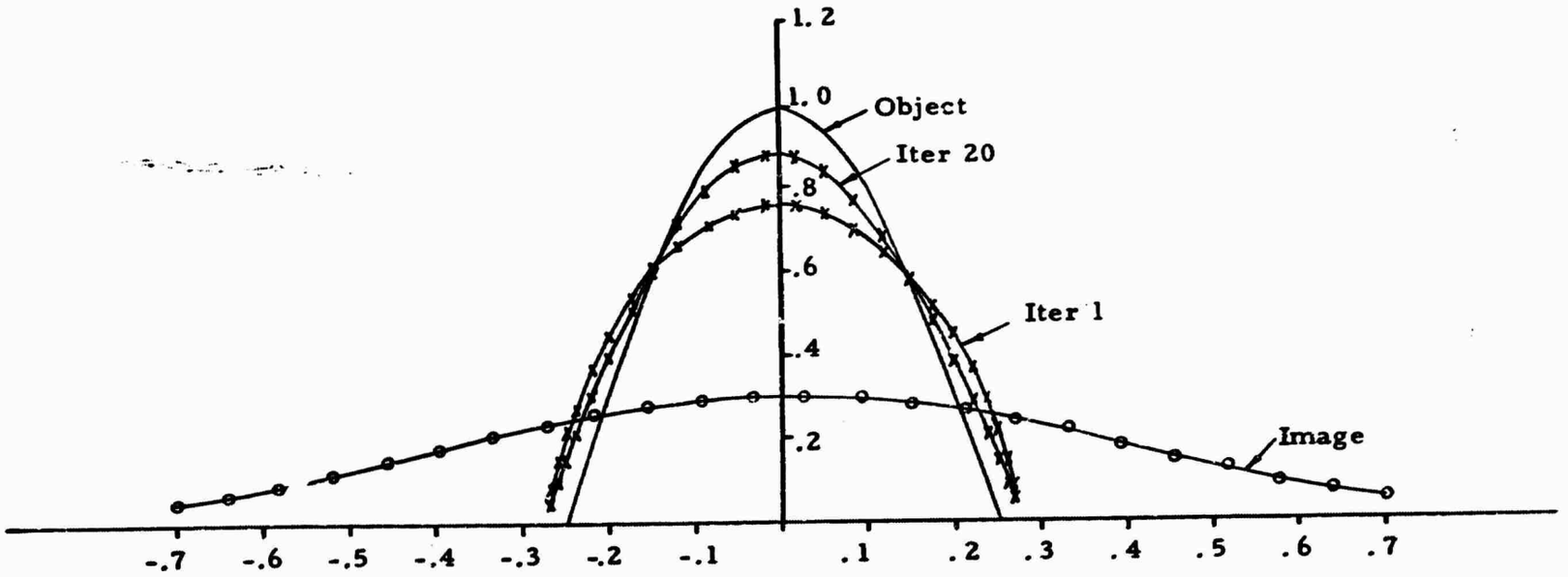

Figure 25. Restoration of the pulse cos (2πα) using (155) with $\bar{x} = b$.

Run = 24.10									
M=24	N = 24	Quad = 1	R_1 =2.6	R=4	λ = .001	Area=0		Prior = 2	Iter = 20
$\sigma_n^2=1x10^{-5}$	$\sigma_x^2=1x10^{-2}$	P_{max}=.84	P_{min}=5.1		SFR=1.3-2.1		MSF=4	σ_A^2 = 0	NA
$I(d^2)$ = 66.48									

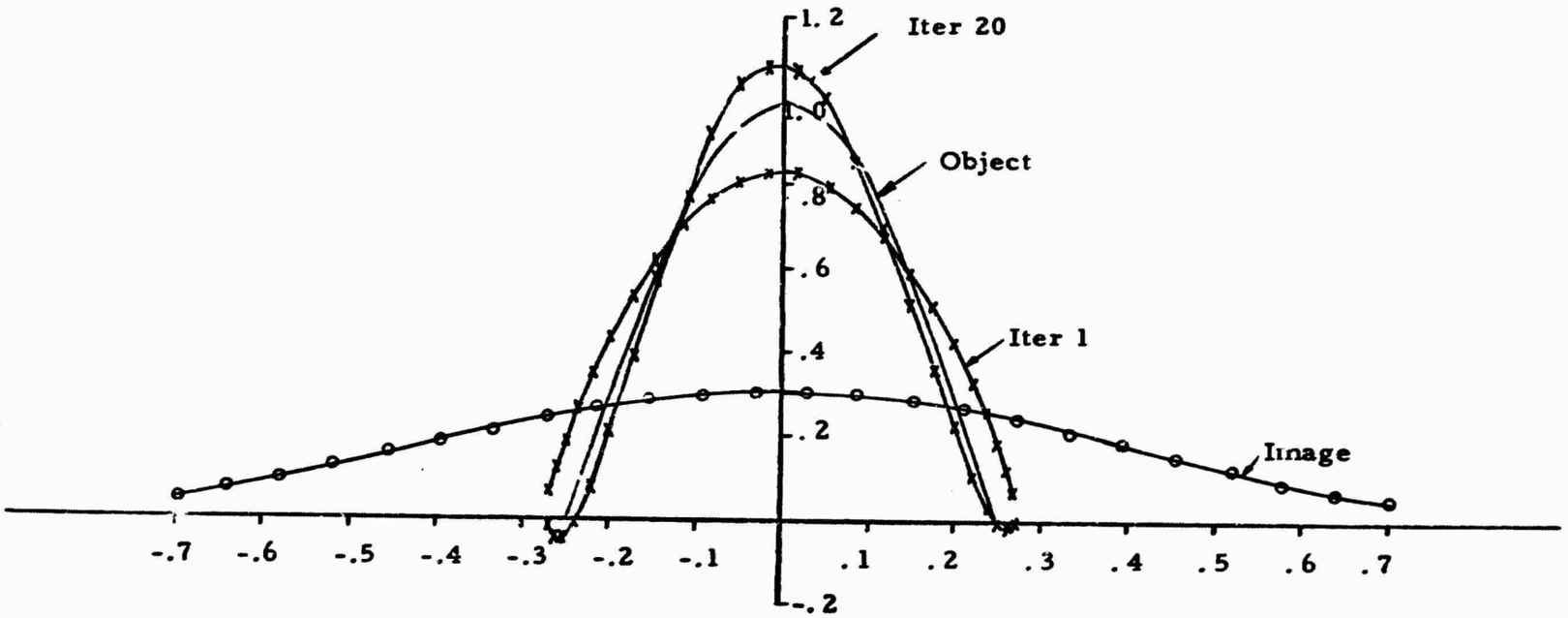

Figure 26. Restoration of the pulse cos ($2\pi\alpha$) using (155) with $\bar{x}$ = b.

However, we have the option of picking any solution between iteration 1 and 20 of Figure 26. Iteration 6 was very close to the true source, and one could ascertain by visually observing the plotted estimates that, as the number of iterations increased beyond 6, the computational inaccuracy began to increase.

The sequential estimation of split uniform pulses

In this series (Figures 27-32), we can assess the importance of prior size information. It is demonstrated that the lack of size knowledge can seriously affect the ability of the restoration process to restore object detail. It will also be shown that we can alleviate this problem somewhat by considering various size estimates and using an alternate size estimation scheme in which we initially choose a very small size estimate instead of a large one as used previously.

Figures 27-28 illustrate an effort to determine the outside size of the object. Notice that iteration 20 for $\sigma_x^2 = 10^{-2}$ is badly distorted, but for $\sigma_x^2 = 10^{-3}$ the 20th iteration is stable. These figures suggest that the noise level is sufficiently high to limit outside size determination, at least for the values of σ_x^2 and the number of iterations considered.

Figures 29-30 show considerable visual improvement over Figures 27-28. This improvement is attributable to the assumption of the true outside size in Figures 29-30. A comparison of Figures 29 and 30 allows a visual assessment of the effect in changing σ_n^2, ($\sigma_n^2 = 10^{-5}$ for Figure 29 but is reduced to 10^{-6} in Figure 30).

Run = 12.7								
M = 32	N = 32	Quad = 1	R1=1.75	R = 4	λ = .001	Area = 0	Prior = 2	Iter = 20
$\sigma_n^2 = 1x10^{-5}$	$\sigma_x^2 = 1x10^{-2}$	Pmax=3.2	Pmin=12.1	SFR=1-1.5		MSF=5	$\sigma_A^2 \approx 0$	NA
$I(d^2) = 1.19$								

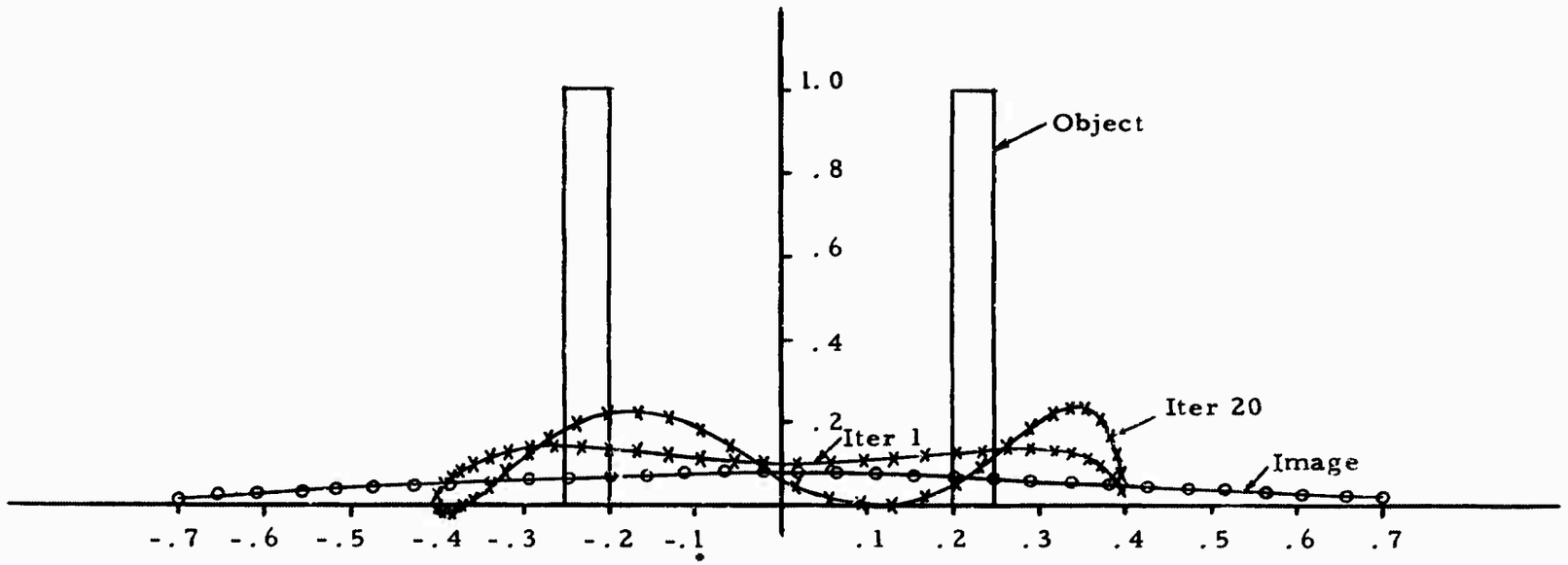

Figure 27. Restoration of split uniform pulses using (155) with $\bar{x} = b$.

Run = 12.7								
M = 32	M = 32	Quad = 1	R1 = 1.75	R = 4	λ = .01	Area = 0	Prior = 2	Iter = 20
$\sigma_n^2 = 1 \times 10^{-5}$	$\sigma_x^2 = 1 \times 10^{-3}$	$P_{max} = 3.2$	$P_{min} = 12.1$	SFR = 1 - 1.5	MSF = 5	$\sigma_A^2 \approx 0$	NA	
$\Sigma(d^2) = 1.20$								

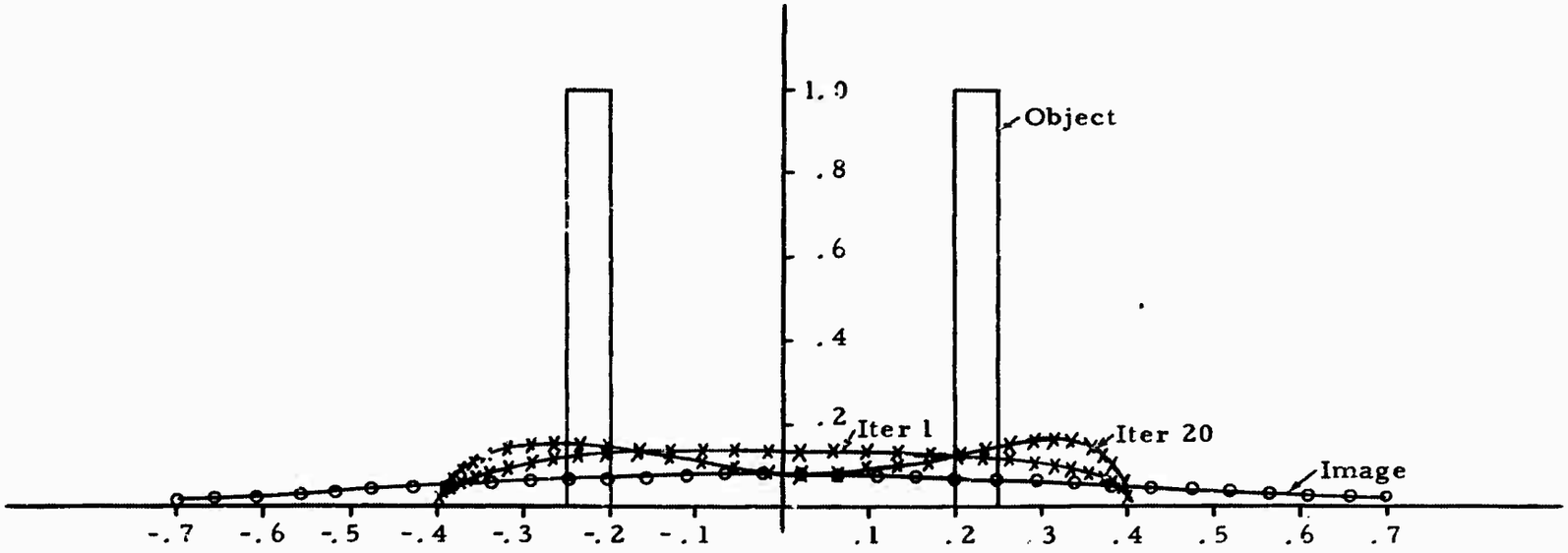

Figure 28. Restoration of split uniform pulses using (155) with $\bar{x} = b$.

Run = 12.1								
M = 24	N = 24	Quad=1	R1=2.7	R = 4	λ = .001	Area = 0	Prior = 2	Iter = 30
$\sigma_n^2 = 1x10^{-5}$	$\sigma_x^2 = 1x10^{-2}$	P_{max}=3.06	P_{min}=18.3	SFR=.7-1.5	MSF=5	$\sigma_A^2 \approx 0$	NA	
$I(d^2)$=1.93								

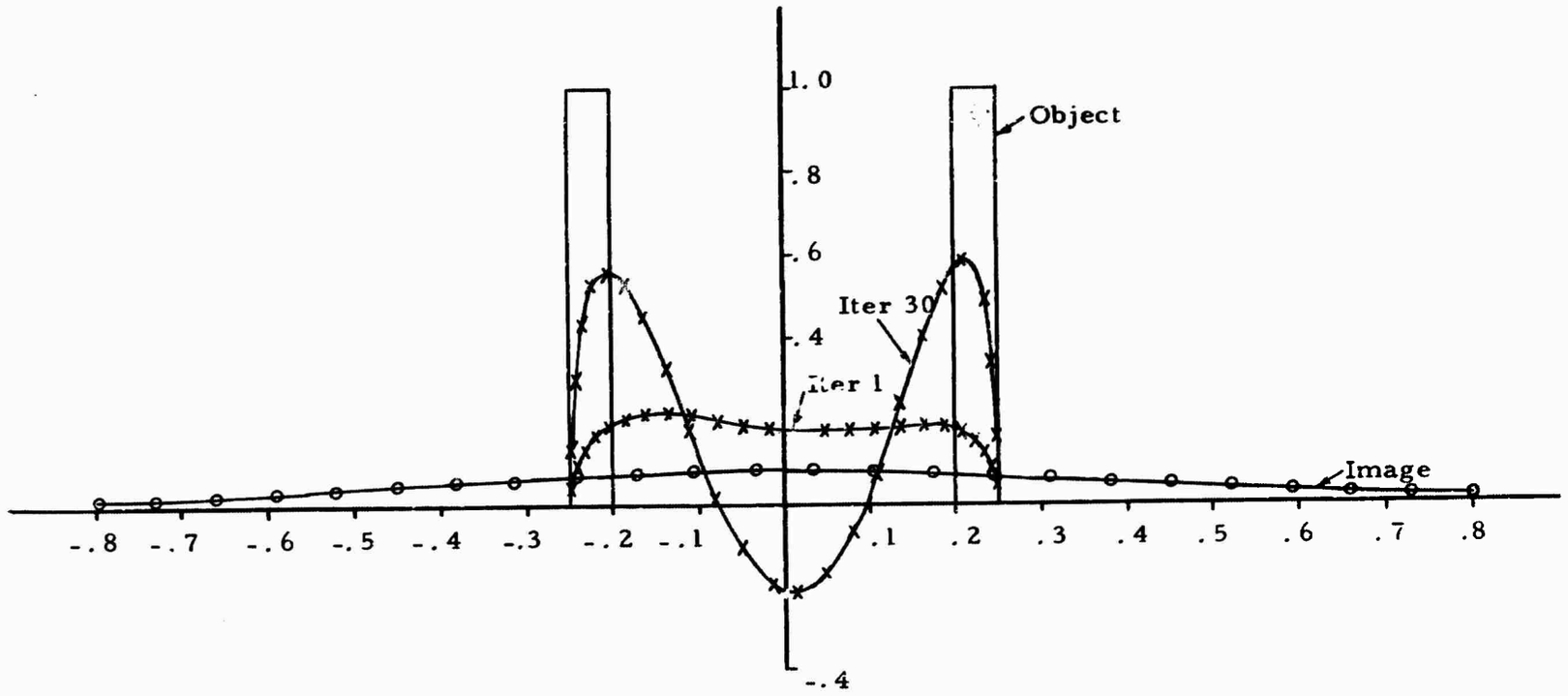

Figure 29. Restoration of split uniform pulses using (155) with $\bar{x} = b$.

Run = 12.2								
M=24	N=24	Quad = 1	R 1=3.2	R = 4	λ =.001	Area = 0	Prior = 2	Iter = 30
$\sigma_n^2 = 1x10^{-6}$	$\sigma_x^2 = 1x10^{-3}$	P_{max}=.97	P_{min}=5.6	SFR=1.2-20	MSF=5	$\sigma_A^2 \approx 0$	NA	
$I(d^2) = 2.20$								

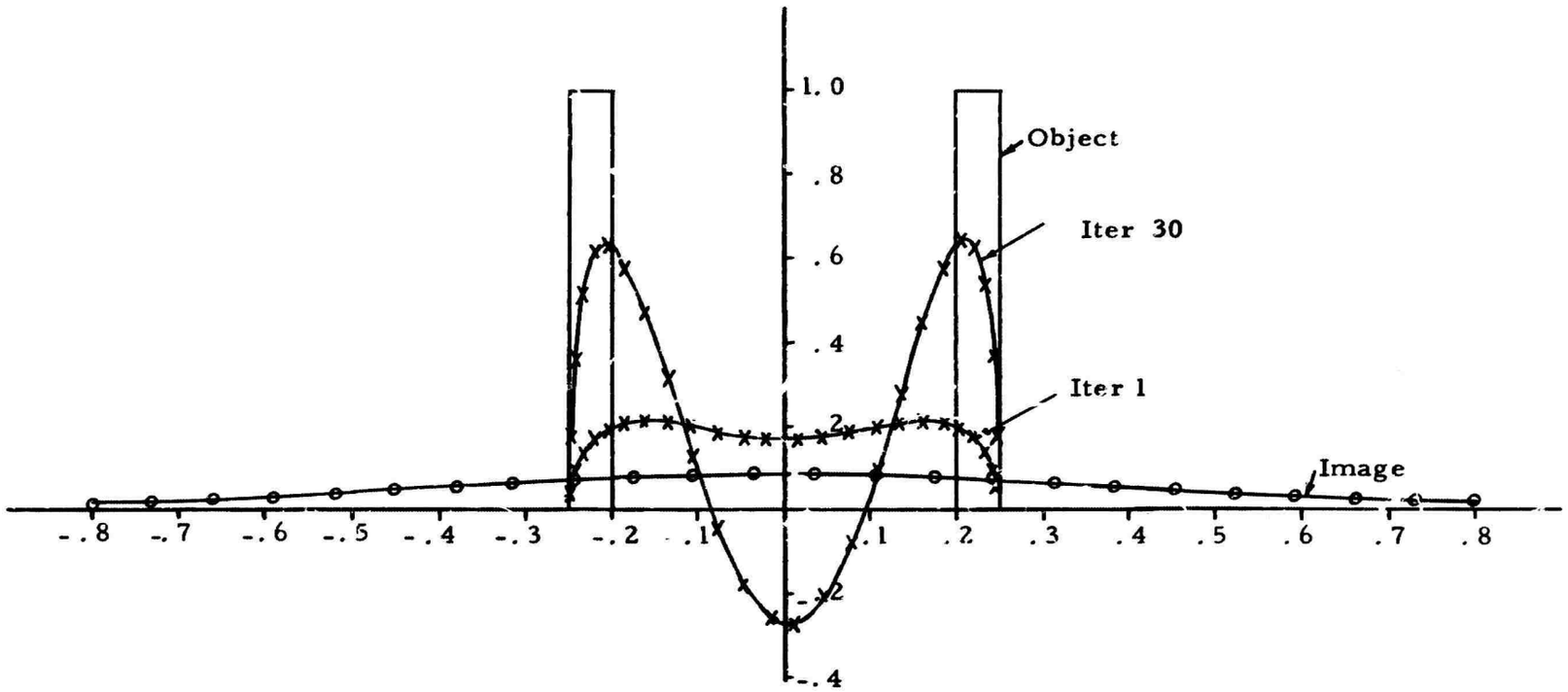

Figure 30. Restoration of split uniform pulses using (.55) with $\bar{x}$ =b.

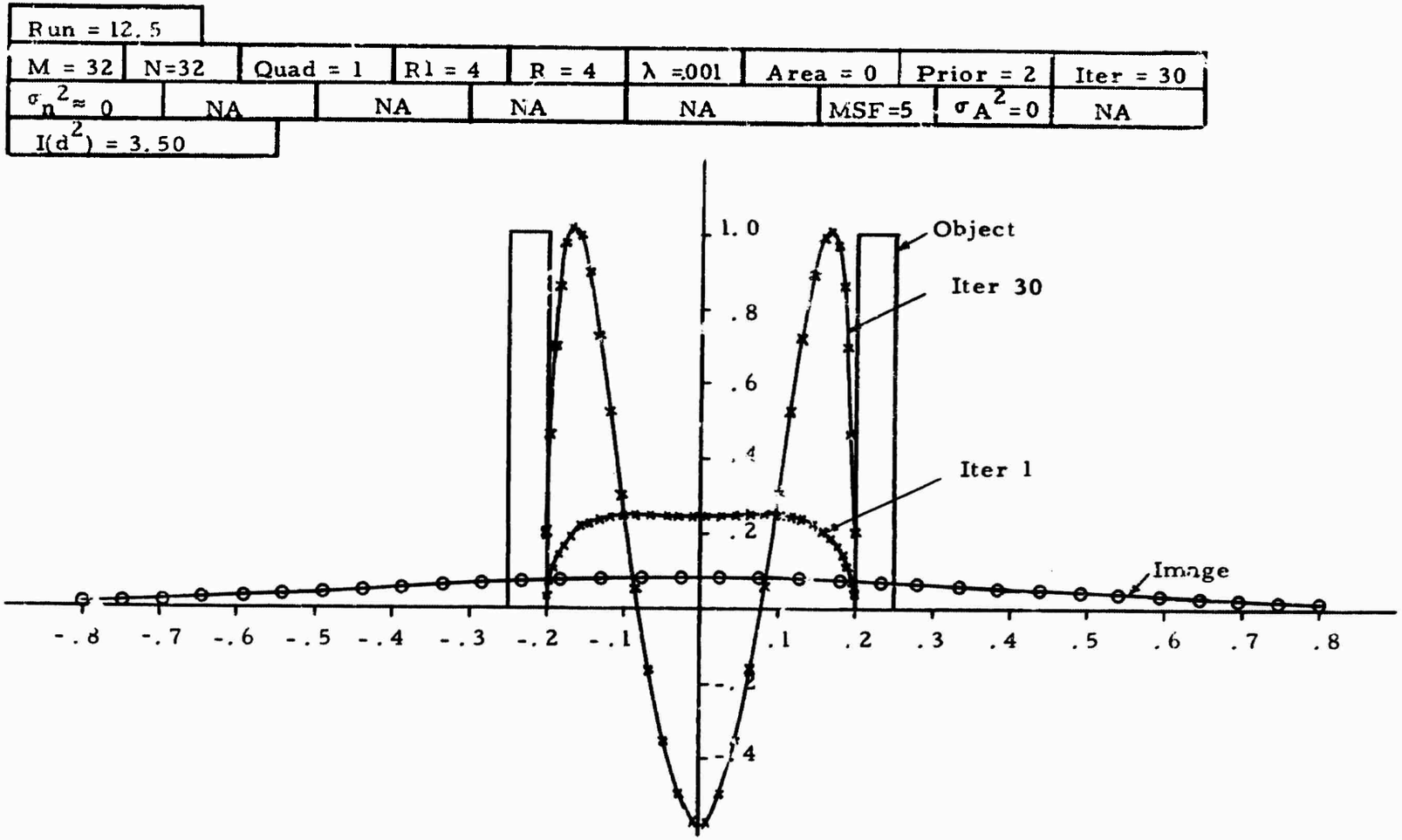

Figure 31. Restoration of split uniform pulses using (155) with $\bar{x} = b$.

Run = 12.94								
M=48	N = 48	Quad=1.0	R1=3.2	R=4	λ = .001	Area = 0	Prior = 2	Iter = 20
$\sigma_n^2 \approx 0$	NA	NA	NA	NA		MSF=5	$\sigma_A^2 \approx 0$	NA
$I(d^2) = 2.15$								

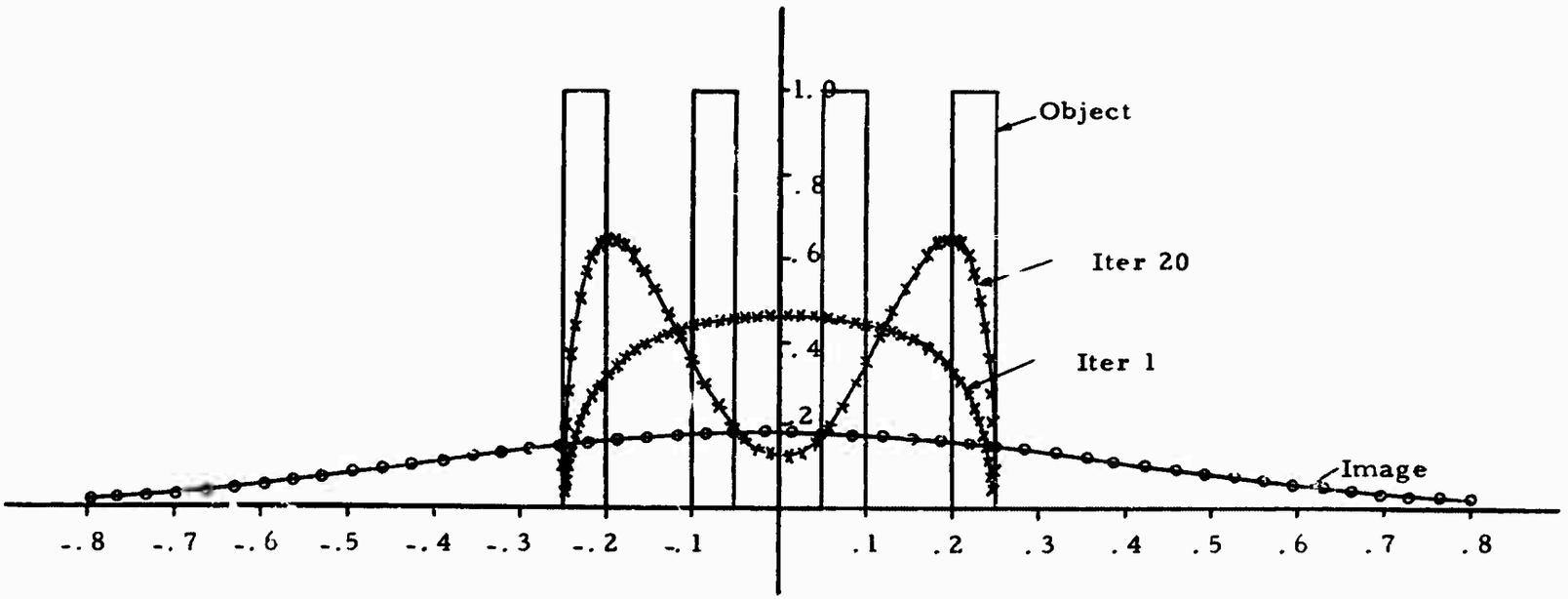

Figure 32. Restoration of split uniform pulses using (155) with $\bar{x} = b$.

Figure 31 presents a restoration in which the assumed outside size interval was smaller than the true value. Note the high values at the endpoints of x which, as before, suggest that the size estimate is too small. This figure also illustrates that the SE is not adequate in measuring improvement because it is evident that there is considerable gain in visual information, but the SE actually shows an improvement decrease. These results add additional insight to the size and shape estimation procedure. They indicate that we should, in any case, consider both large and small size estimates. Also indicated is an alternative size estimation procedure, which initially chooses a small prediction interval close to the image maximum.

Figure 32 shows a single run attempt at restoring 4 narrow closely spaced sources. The restoration does indicate two regions of enhanced object intensity which could be further investigated by using split prediction intervals, as demonstrated below.

The sequential estimation of two split smooth pulses

The last series showed results of estimating the outside size of a split object by assuming various size estimates. Here we demonstrate (Figures 33-37) the use of split prediction intervals in estimating the inside size of a split object. The last two figures (Figures 38 and 39) illustrate that we can resolve two narrow pulses which have been excessively diffracted.

Run=20.60								
M=24	N=24	Quad = 1	Rl=2.7	R = 4	λ =.001	Area = 0	Prior = 2	Iter = 50
$\sigma_n^2 \approx 0$	NA	NA	NA	NA		MSF=4	$\sigma_A^2 \approx 0$	NA
$I(d^2) = 2.50$								

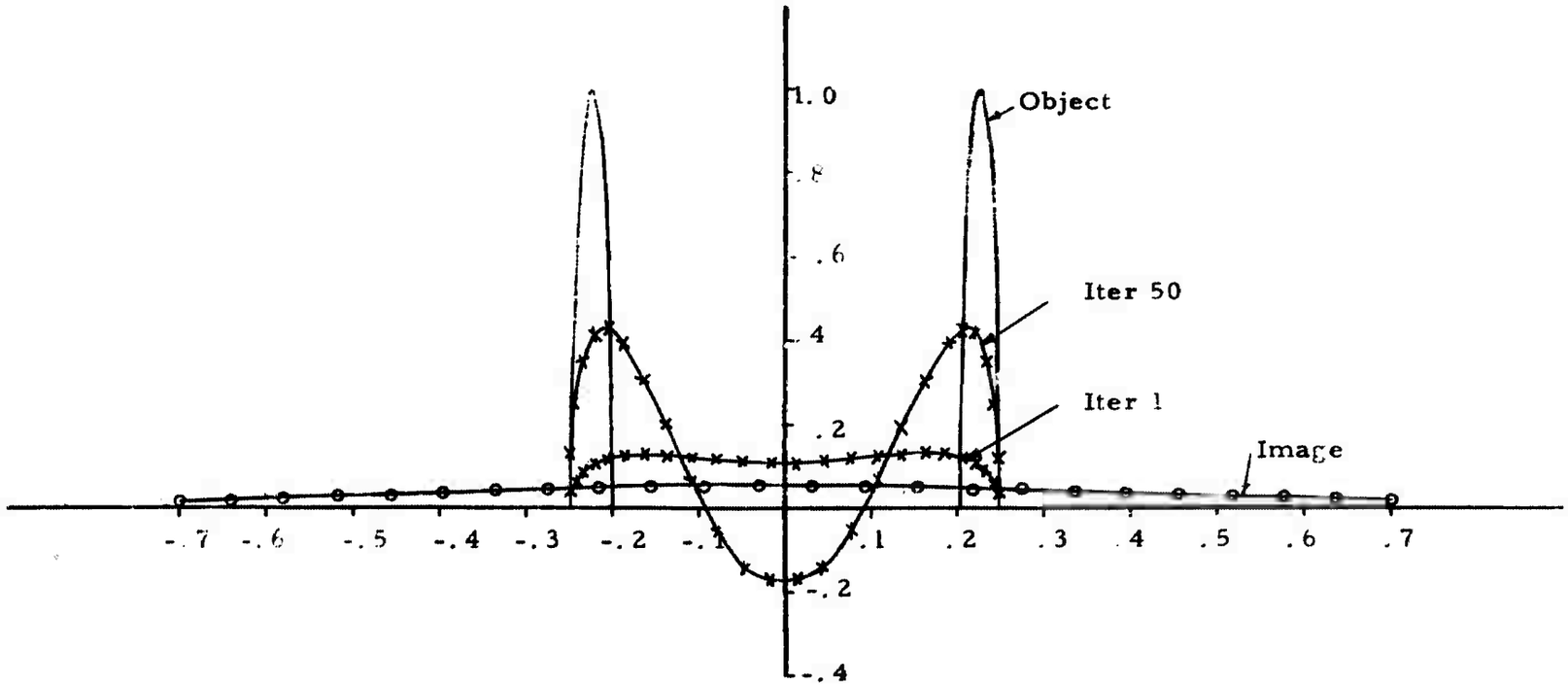

Figure 33. Restoration of the pulses cos $[20\pi(\alpha \pm .225)]$ using (155) with $\bar{x}$ = b.

Run=25.00								
M=22	N=22	Quad = 1	R1=4.1	R = 4	λ=.0001	Area = 0	Prior = 2	Iter = 20
σ_n^2 0	NA	NA	NA	NA	MSF=4	σ_A^2=0	NA	
$I(d^2)$ = 3.42								

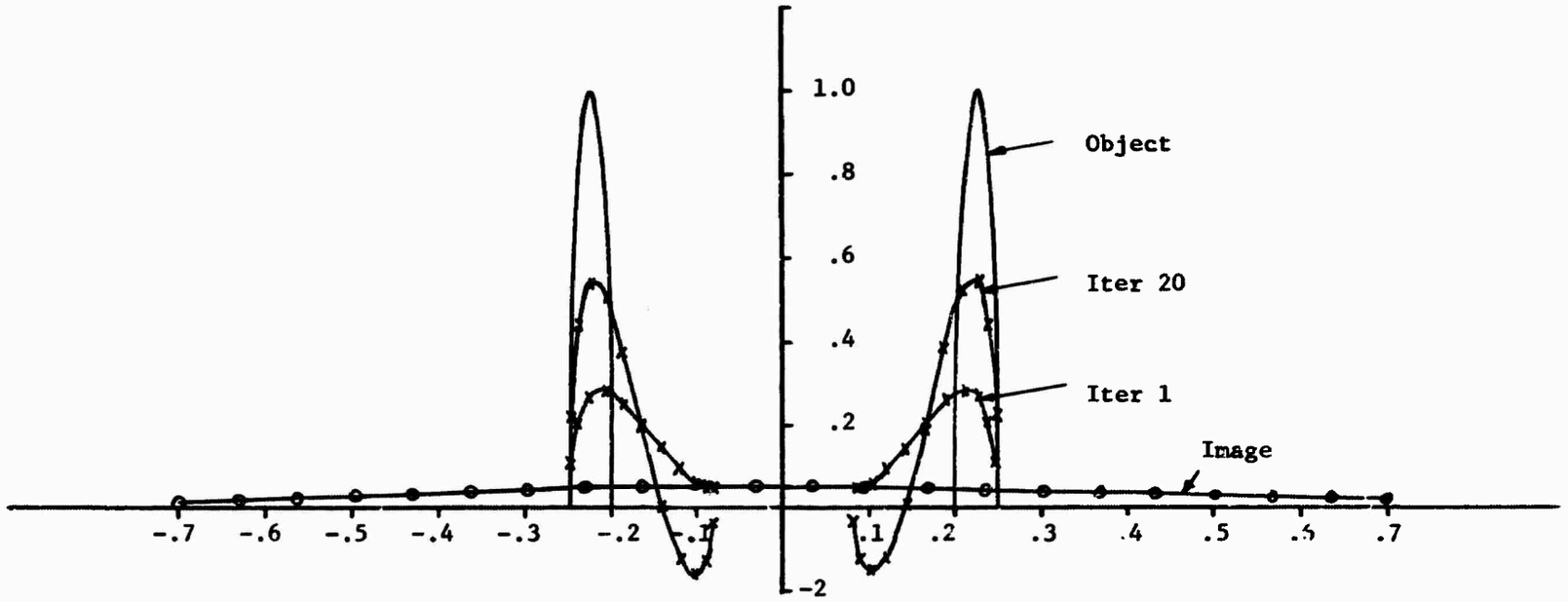

Figure 34. Restoration of the pulses cos [20 π (α ± .225)] using (155) with $\bar{x}$ = b.

Run = 25.10								
M = 22	N = 22	Quad = 1	Rl = 4.1	R = 4	λ=.0001	Area = 0	Prior = 2	Iter = 30
σ_n^2=2.5x10^{-7}	σ_x^2=2.5x10^{-3}	P_{max}= .75	P_{min}= 2.8	SFR = 1.9-2.6	MSF = 4	$\sigma_A^2 \approx 0$	NA	
J(d^2) = 2.80								

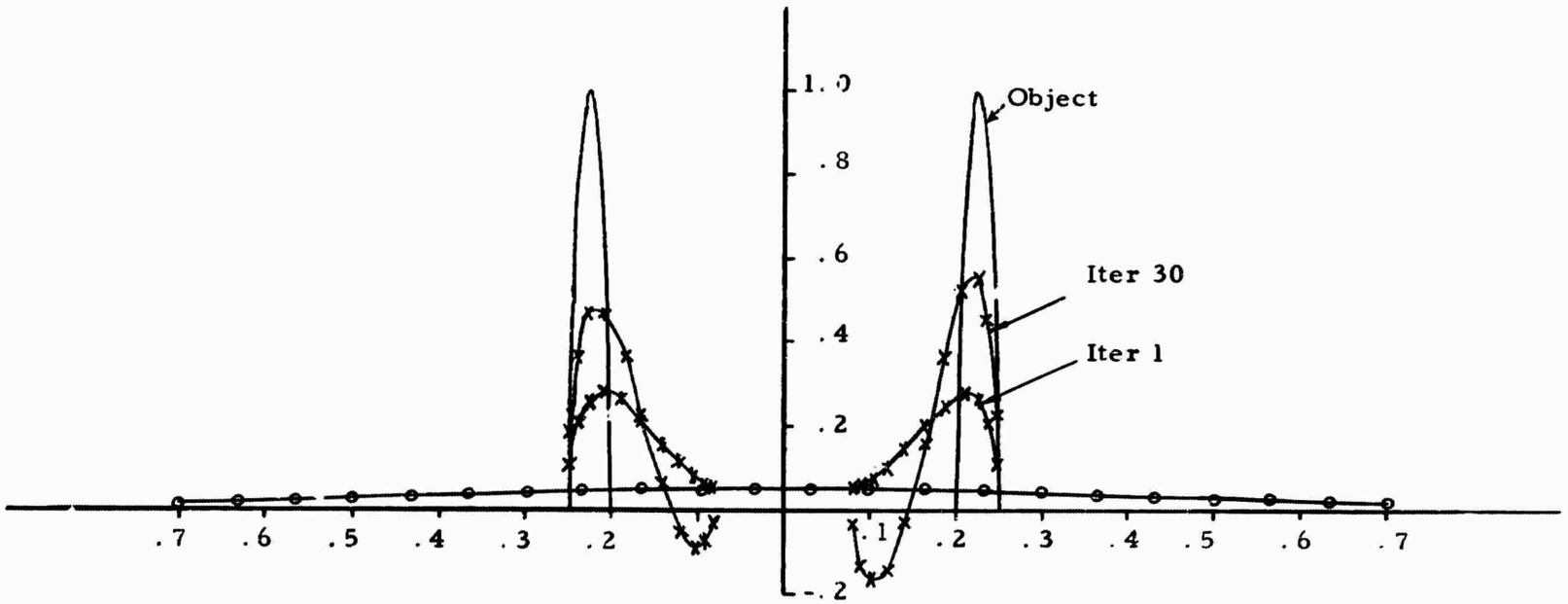

Figure 35. Restoration of the pulses cos [20π (α ± .225)] using (155) with $\tilde{x}$ = b.

Run = 25.2								
M = 22	N = 22	Quad = 1	R1 = 4.1	R = 4	λ = .00001	Area = 0	Prior = 2	Iter = 10
$\sigma_n^2 = 1 \times 10^{-8}$	$\sigma_x^2 = 1 \times 10^{-3}$	P_{max} = .15	P_{min} = .56	SFR = 2.2-2.8		MSF = 4	$\sigma_A^2 \approx 0$	NA
$I(d^2)$ = 4.29								

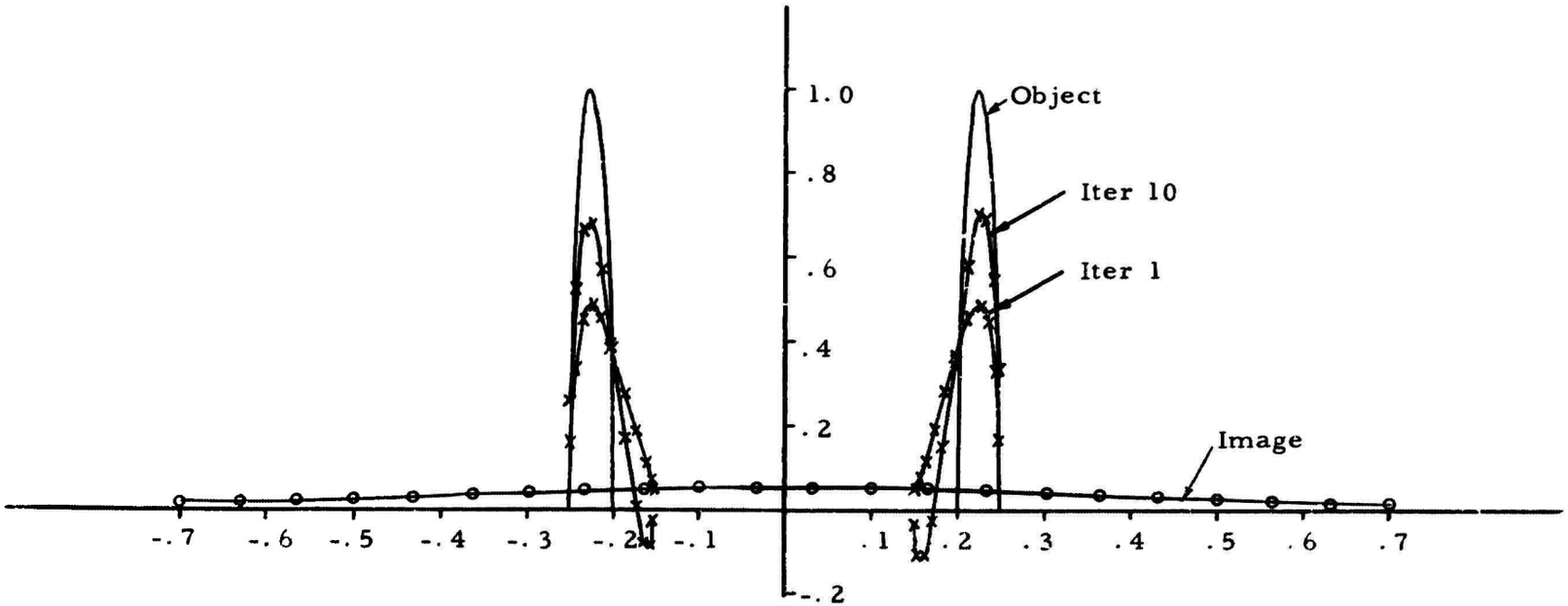

Figure 36. Restoration of the pulses $\cos [20 \pi (\alpha \pm .225)]$ using (155) with $\bar{x} = b$.

Run 25.3								
M = 22	N = 22	Quad = 1	R1 = 4.1	R = 4	λ = .01	Area = 0	Prior = 2	Iter = 80
$\sigma_n^2 = 1 \times 10^{-8}$	$\sigma_x^2 = 1 \times 10^{-6}$	P_{max} = .15	P_{min} = .56	SFR=2.2-2.8	MSF =4	$\sigma_A^2 \approx 0$	NA	
I (d^2) = 20.07								

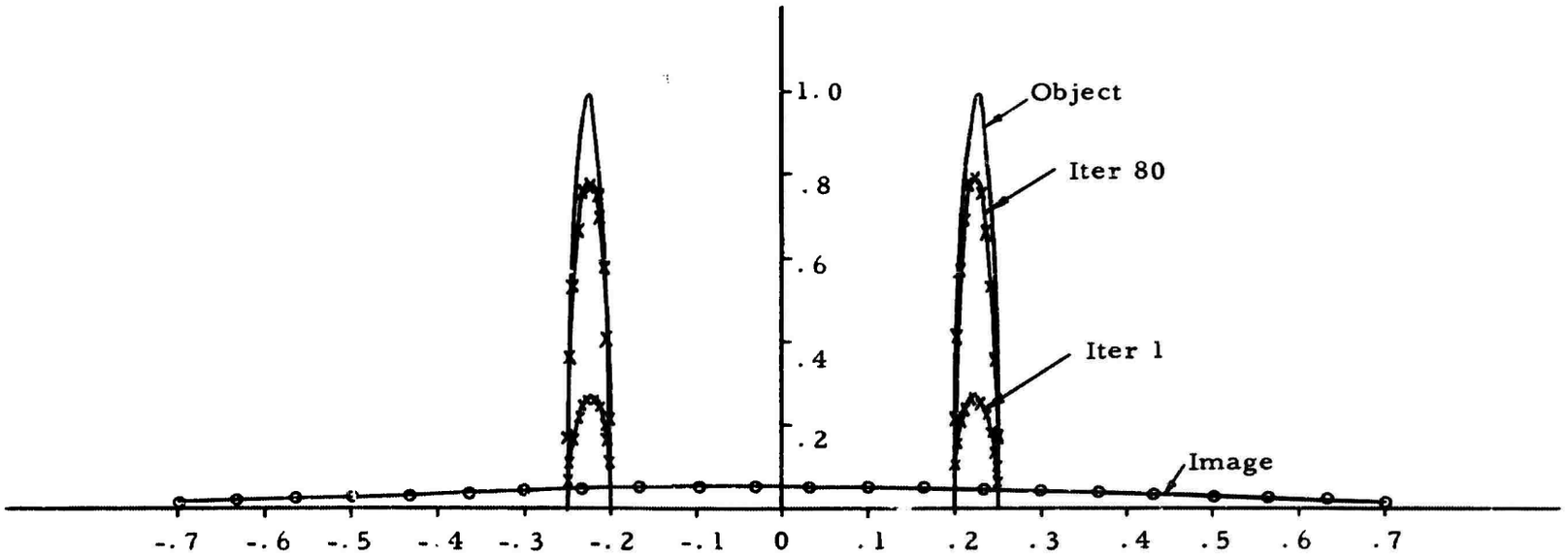

Figure 37. Restoration of the pulses cos [20 π ($\cdot \pm .225$)] using (155) with $\bar{x}$ = b.

In the series, Figures 33-37, on split prediction intervals, Figure 33 establishes the initial inside size. Figures 34-36 indicate further improvement. Notice again how size improvement generally implies shape improvement. Figure 37 shows the improvement when the true size is known. These figures illustrate that, even though the objects are diffracted to a level twice as great as the Rayleigh level, there may be a possibility of not only resolving the objects but of restoring detail as well.

As a follow up of the sequence just considered, one logically seeks the minimum separation between two objects which will still allow the two objects to be resolved. This minimum separation was not found analytically as a function of noise parameters, but Figures 38-39 show results obtained from the restoration of two split sources which are diffracted more than 5 times the usual Rayleigh level.

Figure 38 presents the results for a relatively high SNR case, and Figure 39 shows the impending cost for an increased amount of image noise. Both figures clearly show that the sources are resolved. Based on these results split prediction intervals can be used to further enhance and restore the "bright spots."

Before continuing it is well to reiterate that we have considered here a general restoration procedure which proposes to restore general objects in every detail. We have assumed that very little prior information is

Run = 12.81							
M=24	N=24	Quad=1	R1=?	R=10	$\lambda=10^{-6}$	Area =0	Prior = 2	Iter=15
$\sigma_n^2 \approx 0$	NA	NA	NA	NA	MSF=6	$FA \approx 0$	NA	
$I(d^2) = 1.64$							

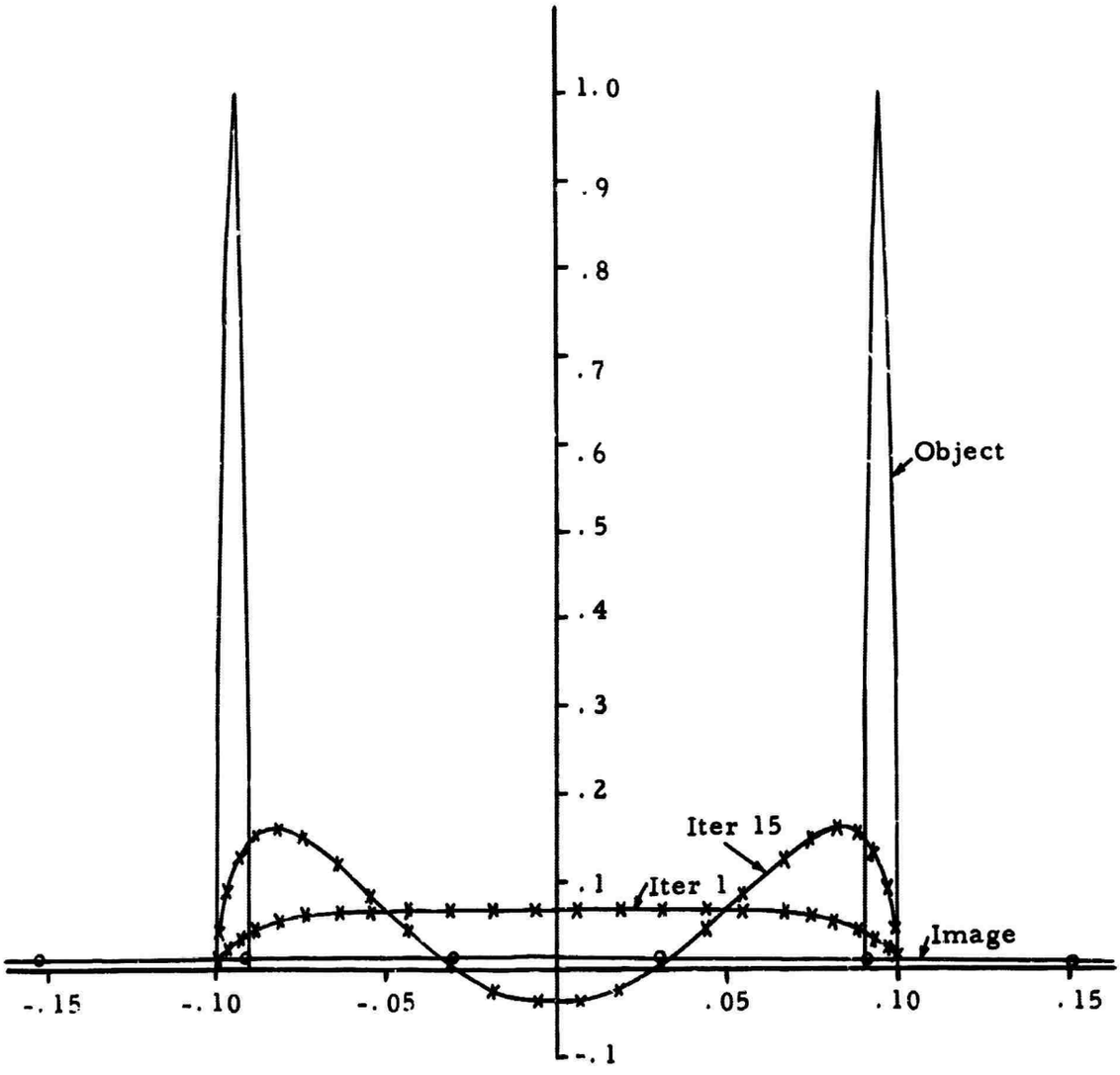

Figure 38. Restoration of the pulse cos $[100\ \pi\ (\alpha \pm .095)]$ using (155) with $\bar{x} = b$.

Run = 12.82								
M = 24	N = 24	Quad = 1	R1 = 5	R = 10	$\lambda = 10^{-5}$	Area = 0	Prior = 2	Iter = 15
$\sigma_n^2 = 1 \times 10^{-2}$	$\sigma_x^2 = 1 \times 10^{-3}$	$P_{max} = .66$	$P_{min} = 1.55$	SFR = 1.8-2.2	MSF = 6	$\sigma_A^2 \approx 0$	NA	
$I(d^2) = 1.43$								

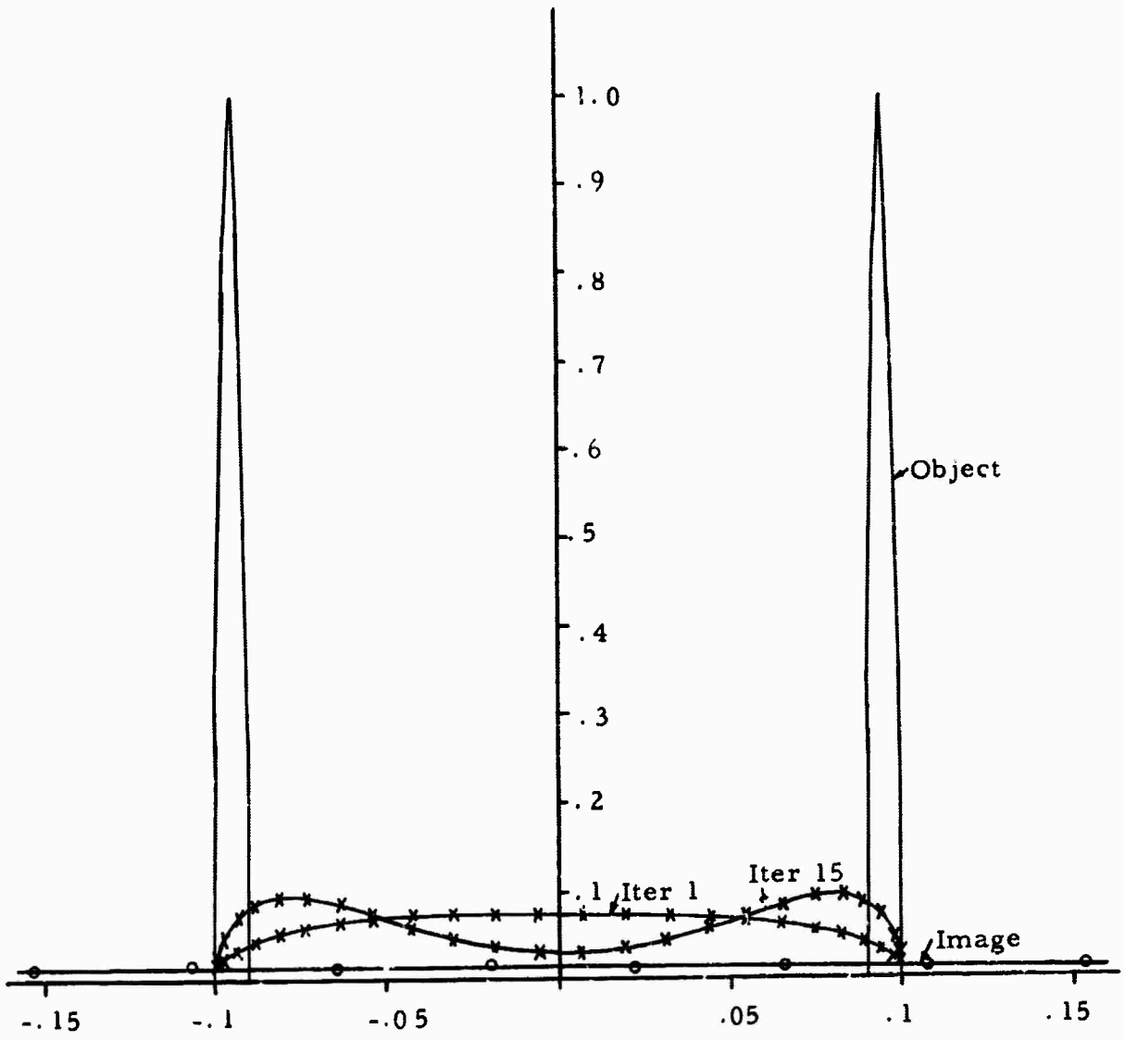

Figure 39. Restoration of the pulses cos [100 π (x ± .095)] using (155) with $\tilde{x} = b$.

available and have been careful to point out those cases when prior size information was used. If the problem were to find the minimum separation given that two point sources made up the object we could undoubtedly do better, in terms of both noise and diffraction, than the results shown in Figures 38-39.

Restoration using a noisy A matrix

Results showing the effect of various amounts of A matrix uncertainty coupled with noisy image measurements are shown in Figures 40-45. These results are likely the most practical of any shown thus far and clearly indicate the possibility of using the MSE restoration procedure in practice.

All of the figures have the same image noise ($\sigma_n^2 = 10^{-5}$, SFR = 1.3-2.1). Three A matrix noise levels are represented: $\sigma_A^2 = 10^{-5}$, $\sigma_A^2 = 10^{-4}$, and $\sigma_A^2 = 10^{-3}$. The first 3 figures use an initial size estimate of [-.5, .5] while the last 3 use an updated size estimate of [-.27, .27]. Again we note the marked general improvement in shape when the size estimate is closer to the true size.

The first two figures, Figures 40 and 41, have the same A matrix and image noise but illustrate the computational advantage of using a smaller σ_x^2 ($\sigma_x^2 = 10^{-4}$ in Figure 41 while $\sigma_x^2 = 10^{-3}$ in Figure 40) to obtain a more stable initial solution, which can be iterated longer and

Run=20, 21								
M=24	N=24	Quad=1	R1=1.4	R = 4	λ = .01	Area=0	Prior=2	Iter = 30
σ_n^2=1x10^{-5}	σ_x^2=.001	P_{max}=.84	P_{min}=5.04	SFR=1.3-2.1	MSF=4	σ_A^2=1x10^{-5}	$Pmax_A$=.27	
I(d^2) = 8.85								

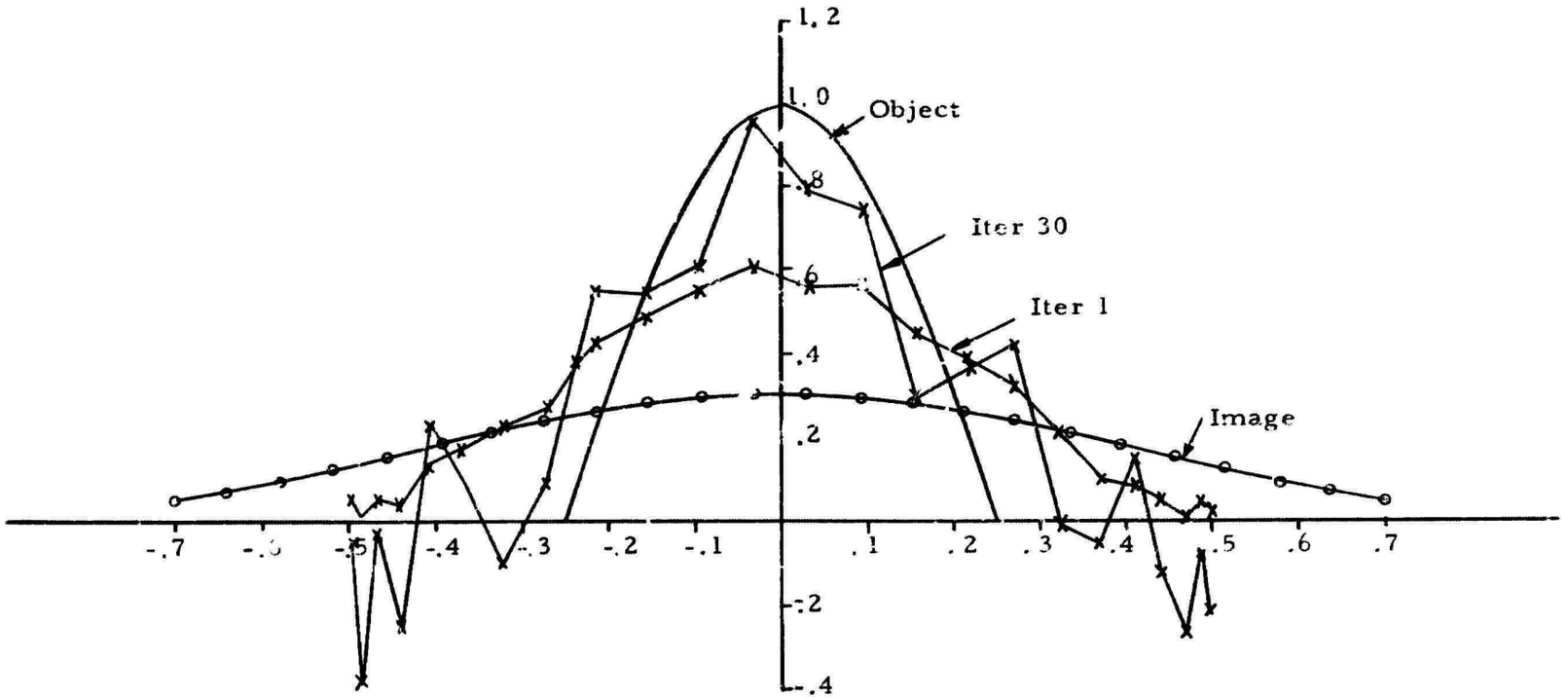

Figure 40. Restoration of the pulse cos ($2\pi\alpha$) using (155) with $\bar{x}$ = b.

Run = 20.21									
M = 24	N = 24	Quad = 1	RI=1.4	R = 4	λ = .1	Area = 0	Prior = 2	Iter = 75	
$\sigma_n^2 = 1 \times 10^{-5}$	$\sigma_x^2 = 1 \times 10^{-4}$	Pmax=.84	Pmin=5.04	SFR=1.3-2.1	MSF=4	$\sigma_A^2 = 1 \times 10^{-5}$	Pmax$_{A=.27}$		
$I(d^2)$ = 11.18									

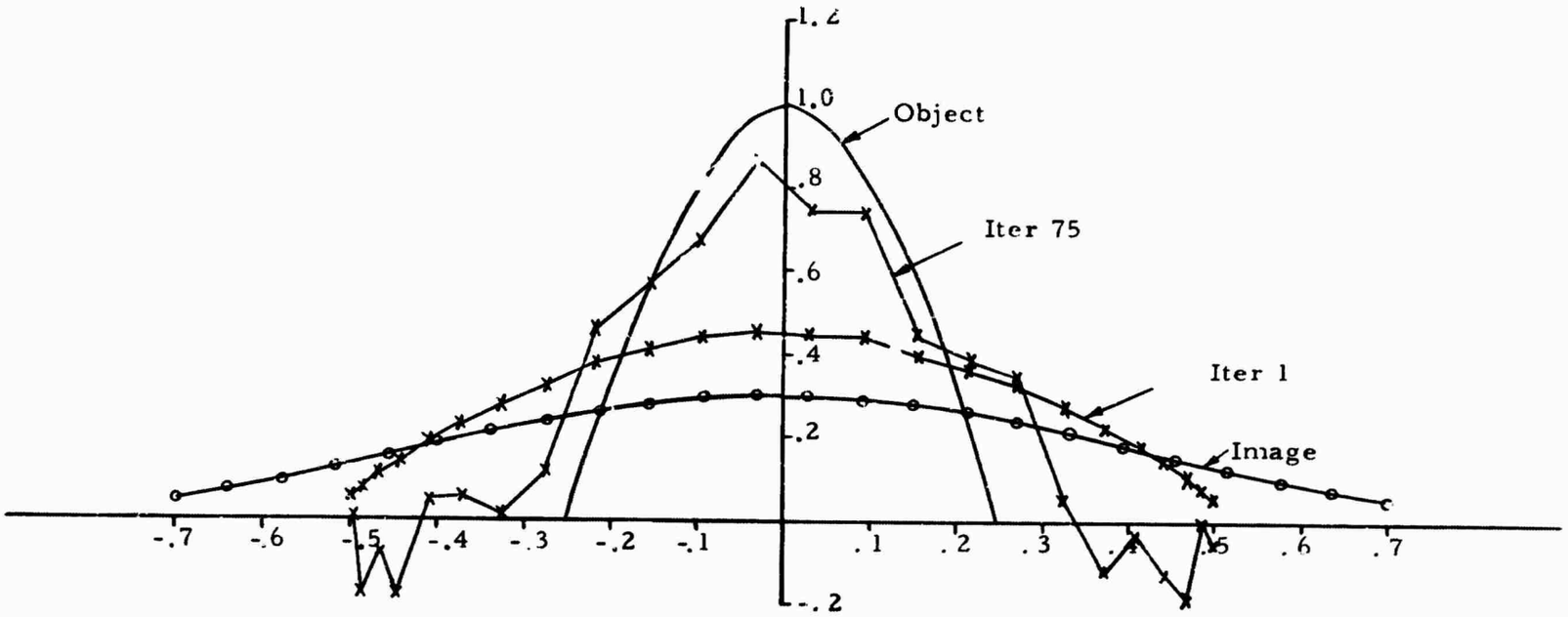

Figure 41. Restoration of the pulse cos $(2\pi\alpha)$ using (155) with $\bar{x}$ = b.

Run = 20.22								
M = 24	N = 24	Quad = 1	R1 = 1.4	R = 4	λ = .1	Area = 0	Prior = 2	Iter = 30
$\sigma_n^2 = 1 \times 10^{-5}$	$\sigma_x^2 = 1 \times 10^{-4}$	Pmax = .84	Pmin = 5.04	SFR = 1.3-2.1	MSF = 4	$\sigma_A^2 = 1 \times 10^{-4}$	Pmax$_A$ = .86	
I(d^2) = 3.69								

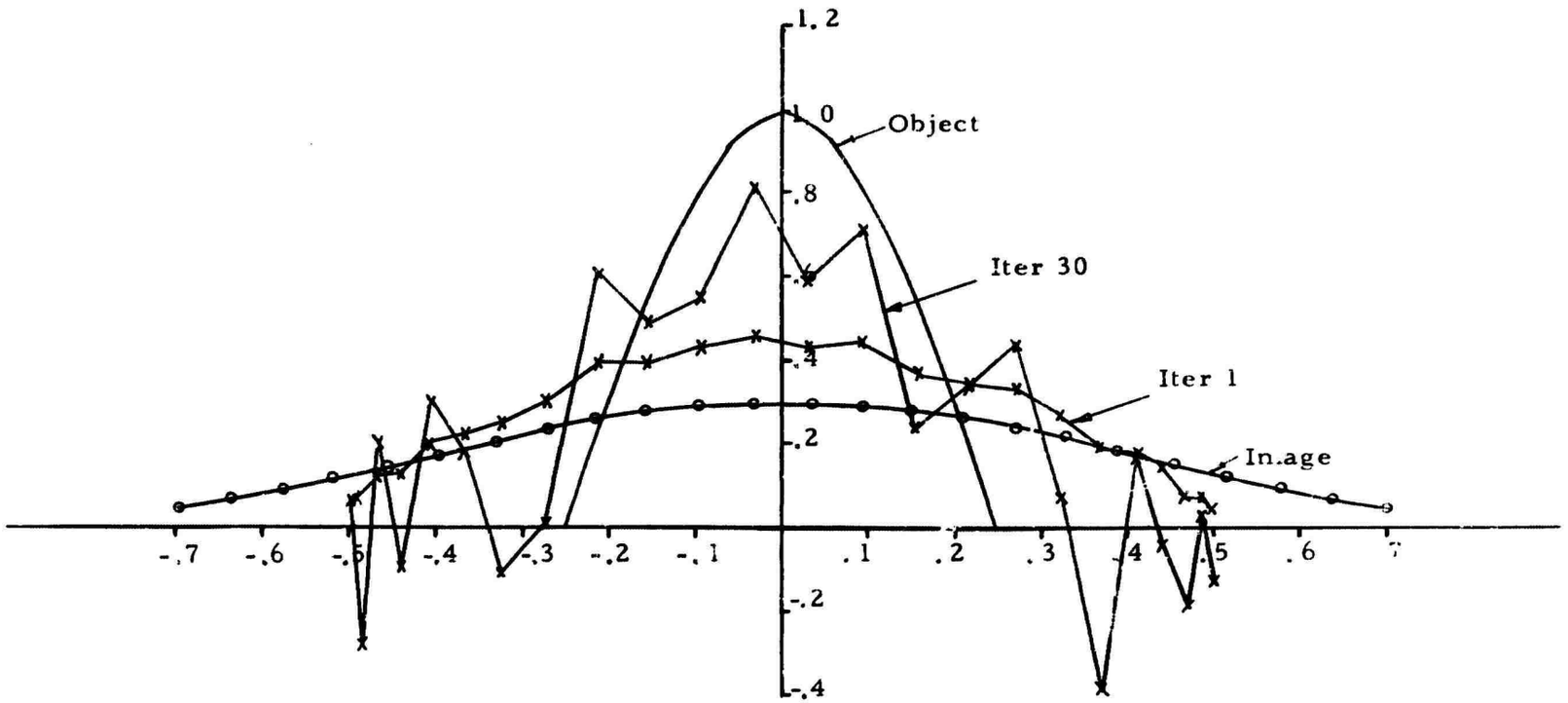

Figure 42. Restoration of the pulse cos($2\pi\alpha$) using (155) with $\bar{x}$ = b.

Run = 20.23									
M = 15	N = 15	Quad = 1	R1 = 2.6	R = 4	λ = .1	Area = 0	Prior = 2	Iter = 50	
$\sigma_n^2 = 1 \times 10^{-5}$	$\sigma_x^2 = 1 \times 10^{-4}$	Pmax = .71	Pmin = 4.3	SFR = 1.3-2.1	MSF=4	$\sigma_A^2 = 1 \times 10^{-5}$	P_{max_A} = .226		
$I(d^2)$ = 15.52									

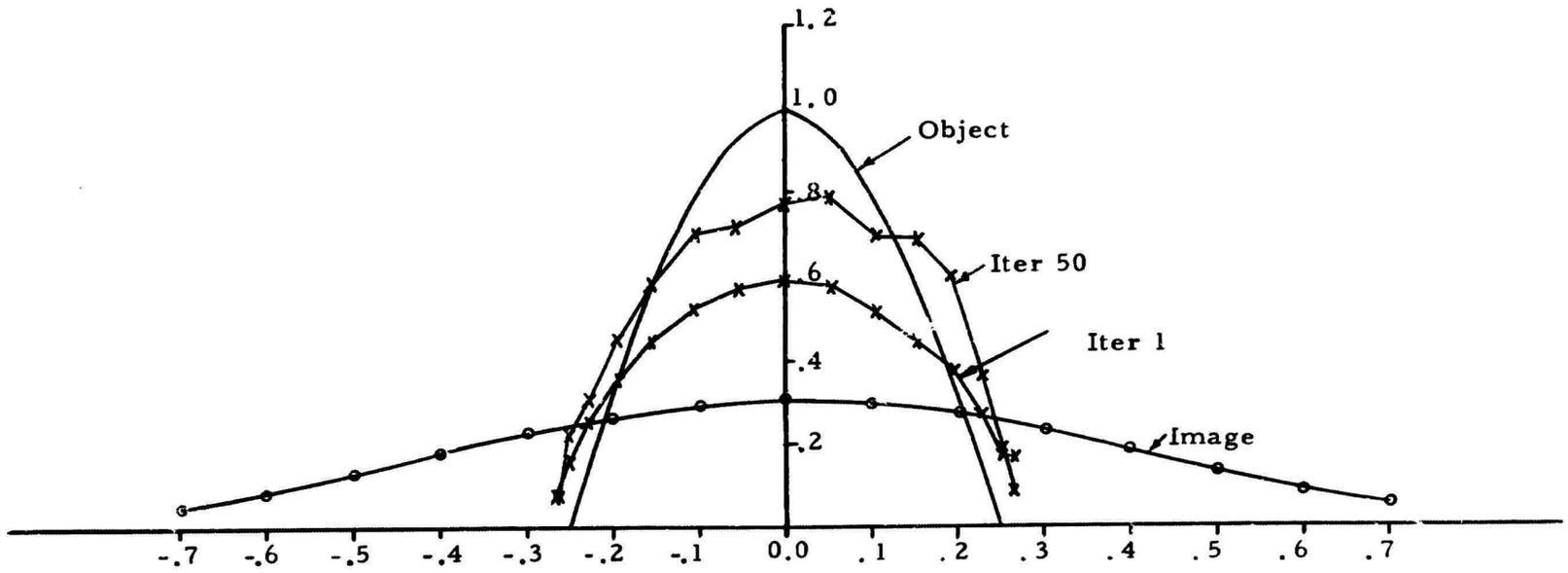

Figure 43. Restoration of the pulse $\cos(2\pi\alpha)$ using (155) with $\bar{x} = b$.

Run = 20.24								
M = 15	N = 15	Quad = 1	R1=2.6	R = 4	λ = .1	Area = 0	Prior = 2	Iter = 50
σ_n^2=1x10^{-5}	σ_x^2=1x10^{-4}	Pmax=.71	Pmin=4.3	SFR=1.3-2.1	MSF=4	σ_A^2=1x10^{-4}	PmaxA=.84	
I(d^2) = 3.17								

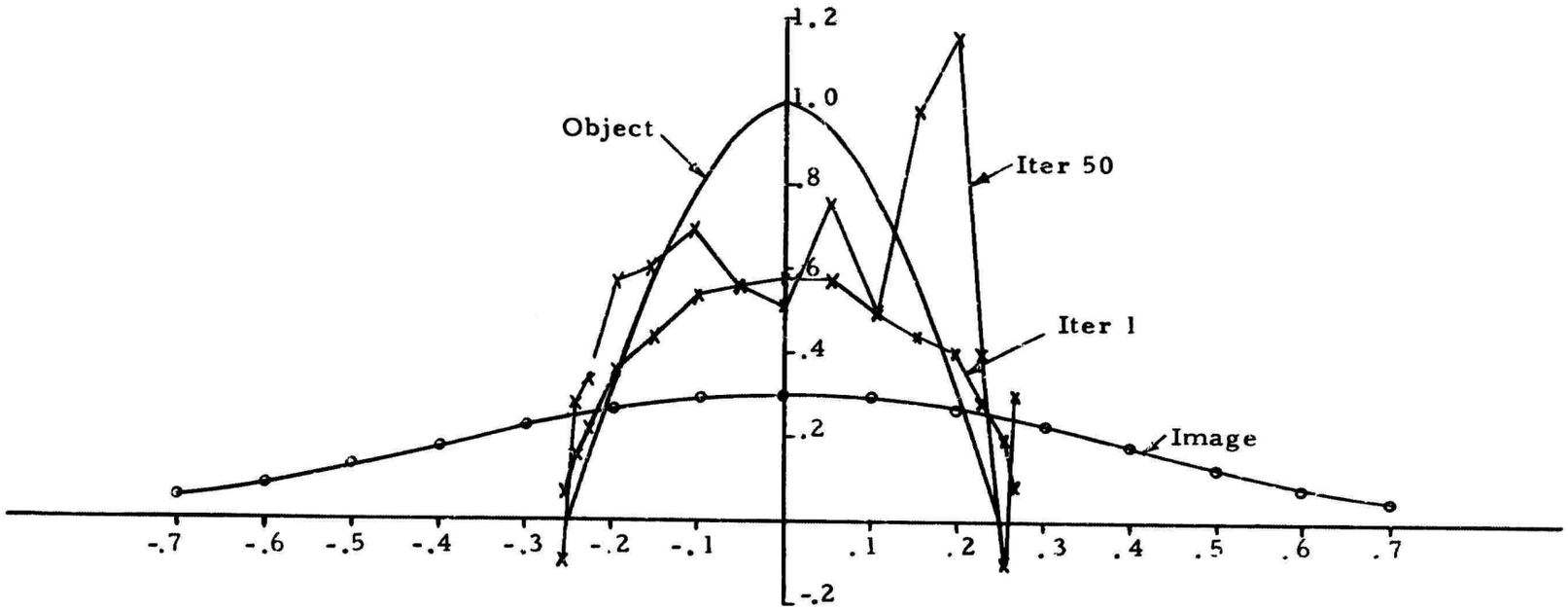

Figure 44. Restoration of the pulse $\cos(2\pi\alpha)$ using (155) with $\bar{x} = b$.

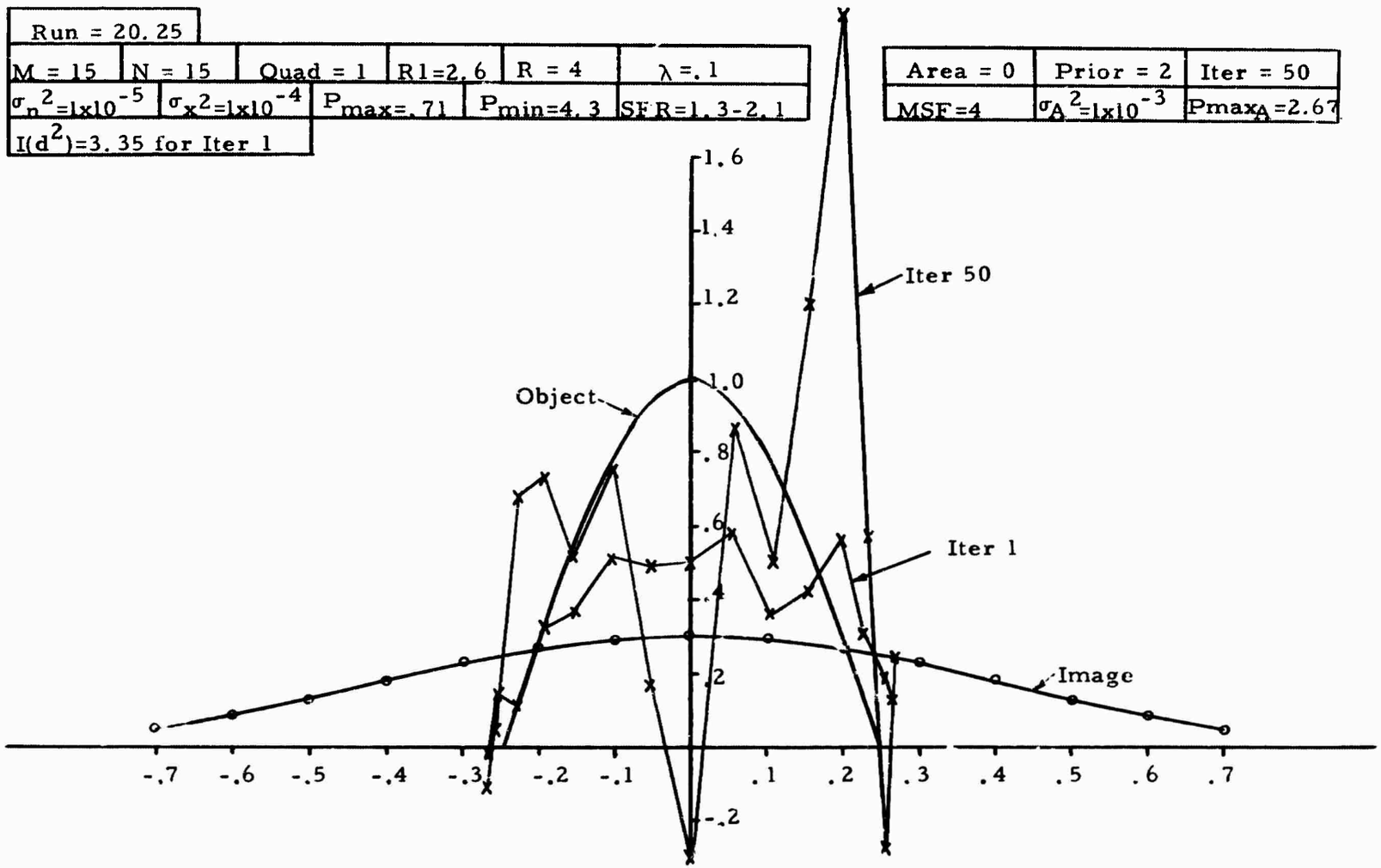

Figure 45. Restoration of the pulse $\cos(2\pi\alpha)$ using (155) with $\bar{x} = b$.

thus produce a better restoration. Figure 42 then shows the cost of increased noise on the improved solution of Figure 41.

Figures 43-44 show the same cost as Figures 41-42 when the noise is increased from $\sigma_A^2 = 10^{-5}$ to $\sigma_A^2 = 10^{-4}$, except that the size is now assumed to be closer to the true size. The improvement is noticeable in Figures 43-44 over Figures 41-42.

The last figure in this series, Figure 45, illustrates the further cost in restoration detail when the A matrix noise is further increased to $\sigma_A^2 = 10^{-3}$.

Image smoothing for the low SNR Case

Now we consider the question of what happens when we encounter excessive image noise. Suppose the image we have at our disposal appears to be too noisy to improve. We consider here a possible approach to this problem by smoothing the original noisy image in accordance with prior information inherently available due to diffraction effects.

Figure 46 presents a restoration resulting from the operation on a very noisy image (SFR = .6-1.6). Iteration 1 is not stable for $\sigma_x^2 = 2.5 \times 10^{-1}$. In fact we must weight the prior noisy image so heavily in order to stabilize the solution that there is essentially no improvement. Apparently the noise level is so high that restoration efforts are futile. Of course, if additional samples of the image were available, some improvement may be possible.

Run = 15.3								
M = 50	N = 11	Quad = 2	R1 = 1.5	R = 1	λ = .01	Area = 0	Prior = 2	Iter = 20
$\sigma_n^2 = 2.5 \times 10^{-3}$	$\sigma_x^2 = 2.5 \times 10^{-1}$	$P_{max} = 4.4$	$P_{min} = 42.8$	SFR = .6-1.6	MSF = 5	$\sigma_A^2 \approx 0$	NA	
Not computed								

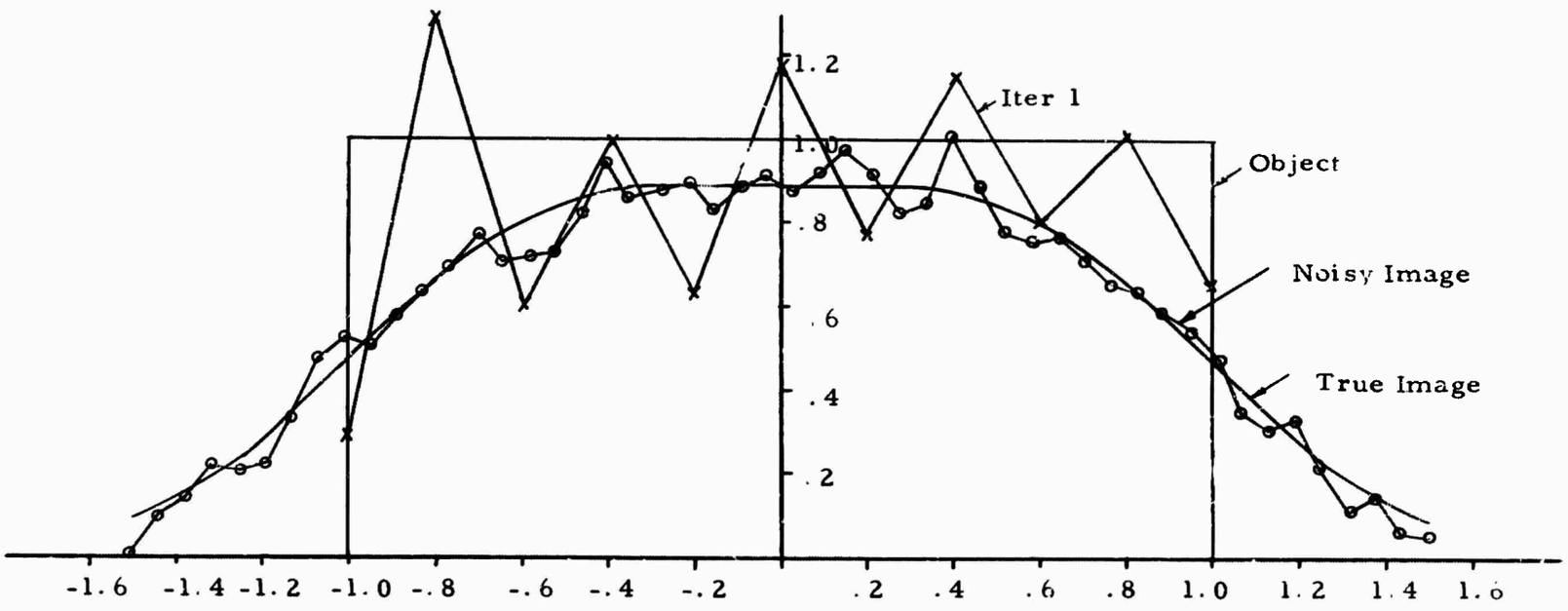

Figure 46. Restoration of a uniform pulse (R= 1.0) using 155) with $\bar{x}$ = b.

On examining the noisy image it is obvious that the spatial extent of the noise fine structure is small compared with the width of the point spread function. This suggests that the noise fine structure in this case can be smoothed, since it could not logically be preserved by the imaging system. A resulting smoothed image is shown in Figure 47. (No sophistication was used in the smoothing process.)

Figure 48 shows the results of using a 9x9 system and operating on the smoothed image. The improvement is noticeable. These results illustrate that even in a low SNR situation in which the detector noise is appreciable some measures can be taken which may result in restoration improvement.

Quadrature effects

As previously discussed in the Imaging Equation section, the transfer from the continuous functions to the discrete version involves some error. The visual effect of this error is illustrated in Figures 49-51. The restoration shown in Figure 49 used unity weights, Simpson's quadrature was used in Figure 50, and Gauss quadrature was used in Figure 51.

Under the stress of both image and A matrix noise it is evident that the more powerful Gauss quadrature is superior. There may be some question about the small values of $\hat{x}$ in Figure 49; however, it was found that decreasing σ_x^2, in general, has a greater scale effect for $w_i = 1.0$ than when one of the other quadratures was used. (This scale change is

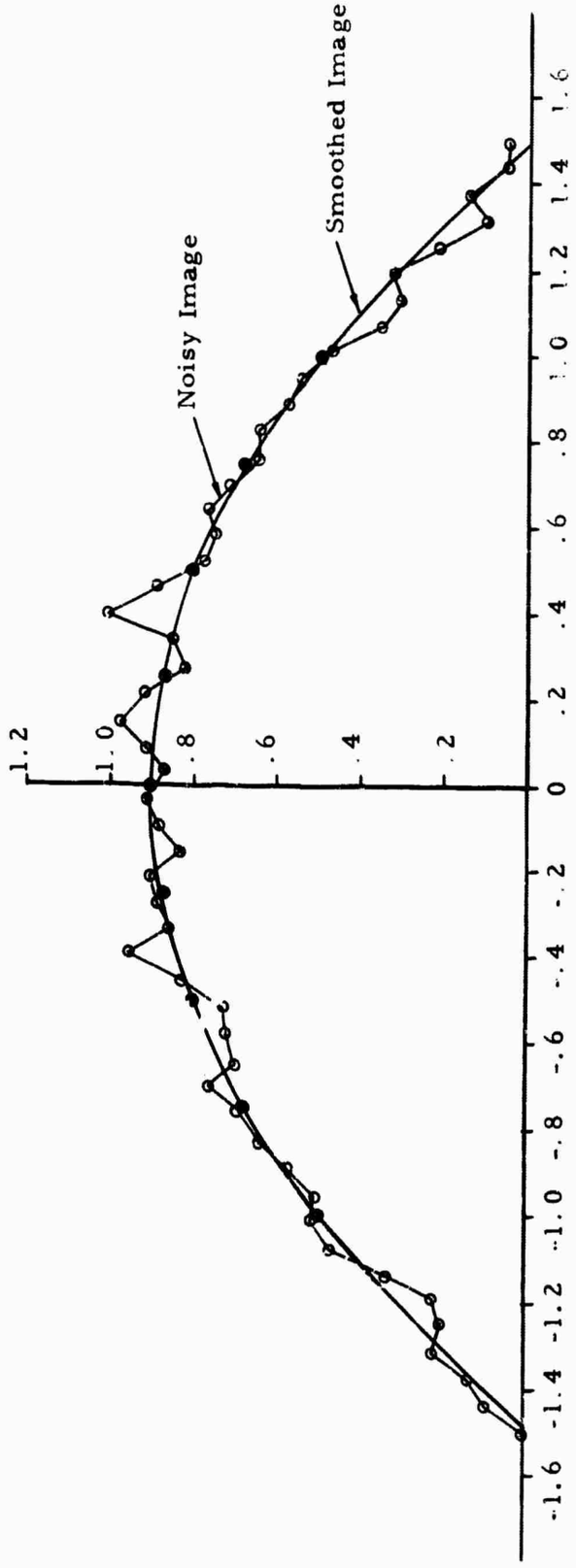

Figure 47. Comparison of the noisy image and the smoothed images.

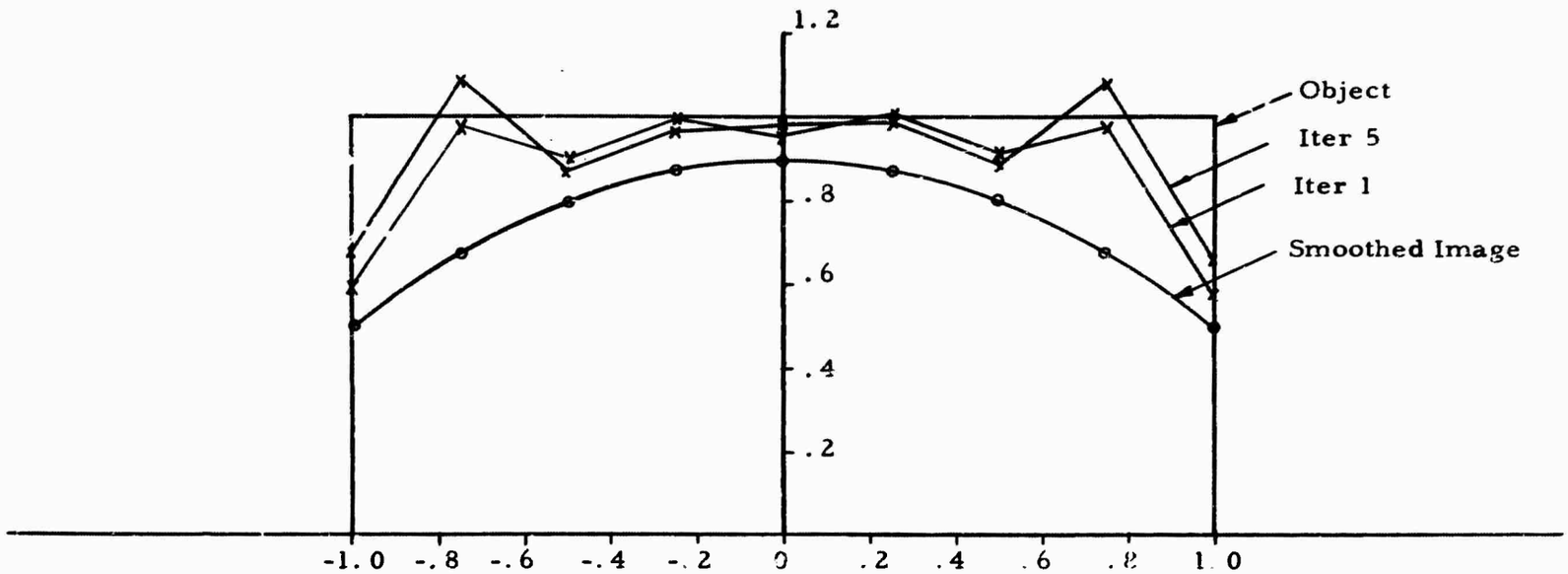

Figure 48. Restoration using the smoothed image of Figure 47. The MSE solution (155) was used with $\bar{x}$ equal to the smoothed image. The ratio $I(d^2) = 2.18$.

Run = 20.26								
M = 15	N = 15	Quad = 3	R1=2.6	R = 4	λ = .1	Area = 0	Prior = 2	Iter = 50
σ_n^2=1x10^{-5}	σ_x^2=1x10^{-4}	Pmax=.71	P_{min}=4.34	SFR=1.3-2.1	MSF=4	σ_A^2=1x10^{-5}	P_{max}=.27	
$\Sigma(d^2)$ = .72								

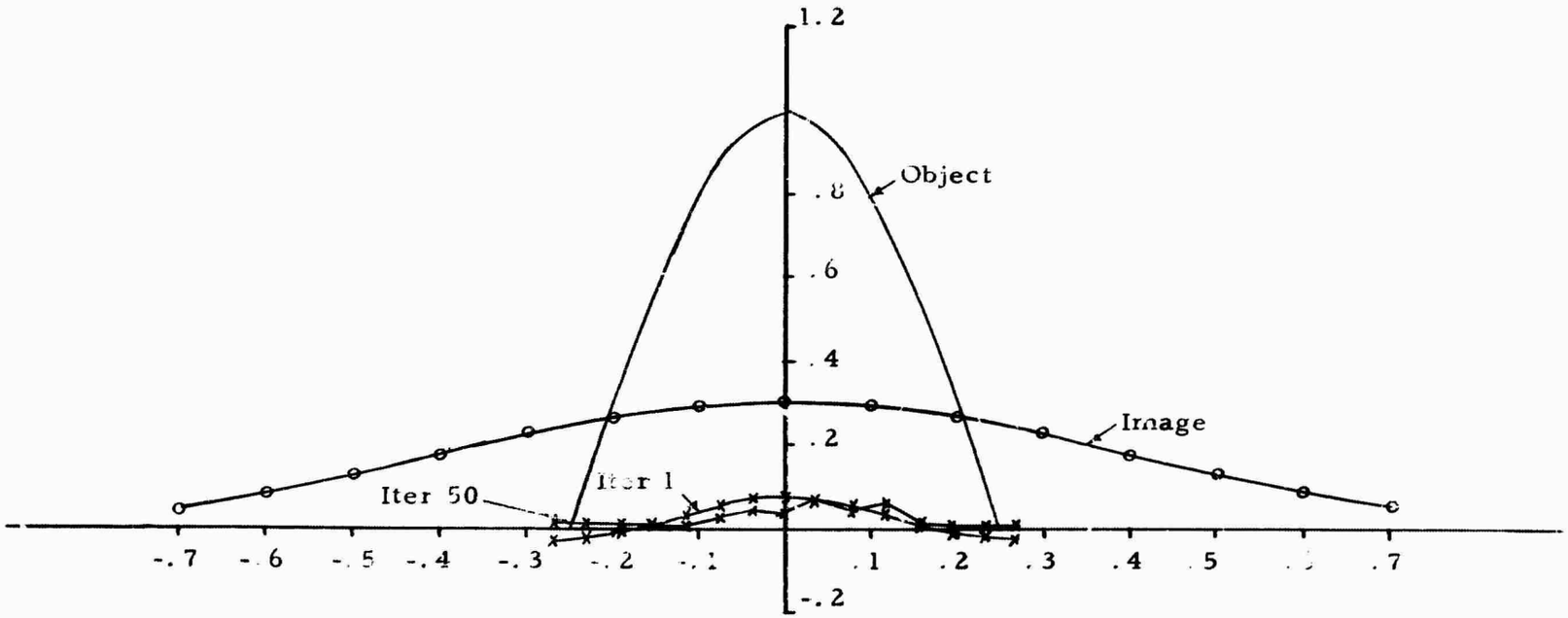

Figure 49. Restoration of the pulse cos $(2\pi\alpha)$ using unity weights in (155) with $\bar{x}$ = b.

Run = 20.27								
M = 15	N = 15	Quad = 2	R1=2.6	R = 4	λ = .1	Area = 0	Prior = 2	Iter = 50
σ_n^2=1x10^{-5}	σ_x^2=1x10^{-4}	P_{max}=.71	P_{min}=4.3	SFR = 1.3-2.1	MSF=4	σ_A^2=1x10^{-5}	$Pmax_A$=.266	
I(d^2) = 4.83								

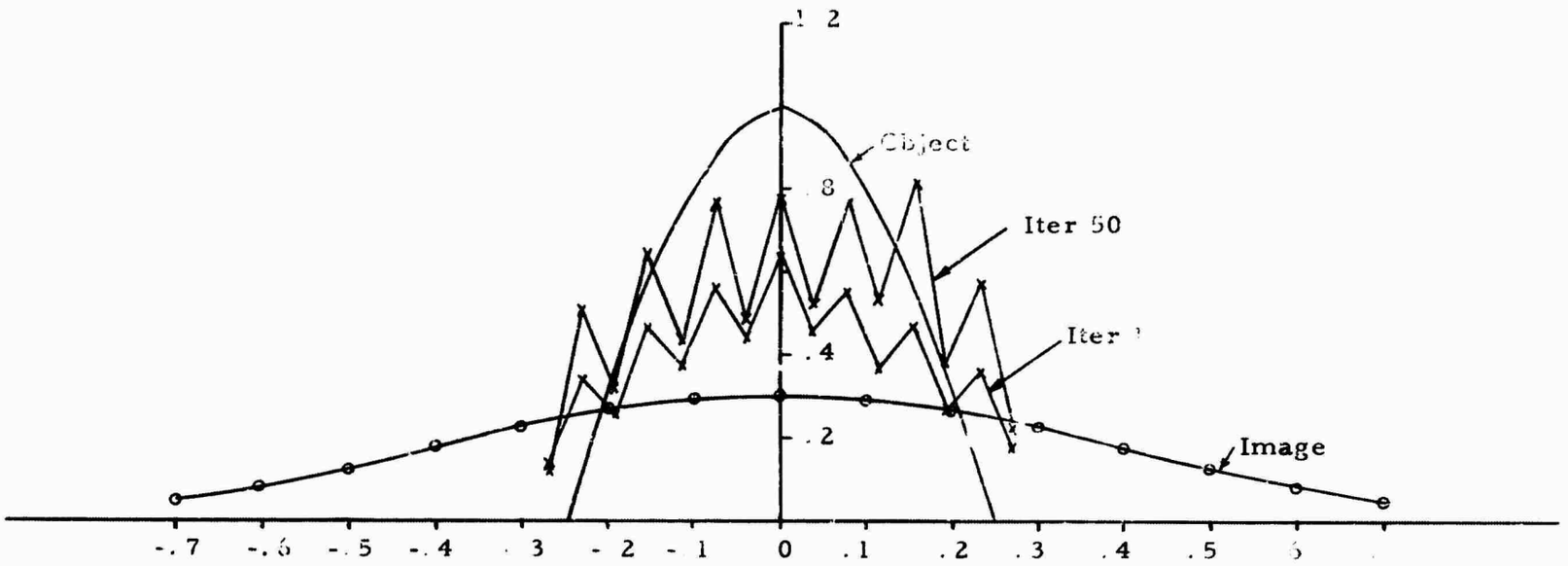

Figure 50. Restoration of the pulse cos $(2\pi x)$ using Simpson's quadrature in (155) with $\bar{x} = b$.

Run = 20.29								
M = 15	N = 15	Quad = 1	R1 = 2.6	R = 4	λ = .1	Area = 0	Prior = 2	Iter = 50
σ_n^2=1x10^{-5}	σ_x^2=1x10^{-4}	P_{max}=.711	P_{min}=4.34	SFR = 1.3-2.1	MSF=4	σ_A^2=1x10^{-5}	$Pmax_A$=.71	
$I(d^2)$ = 3.16								

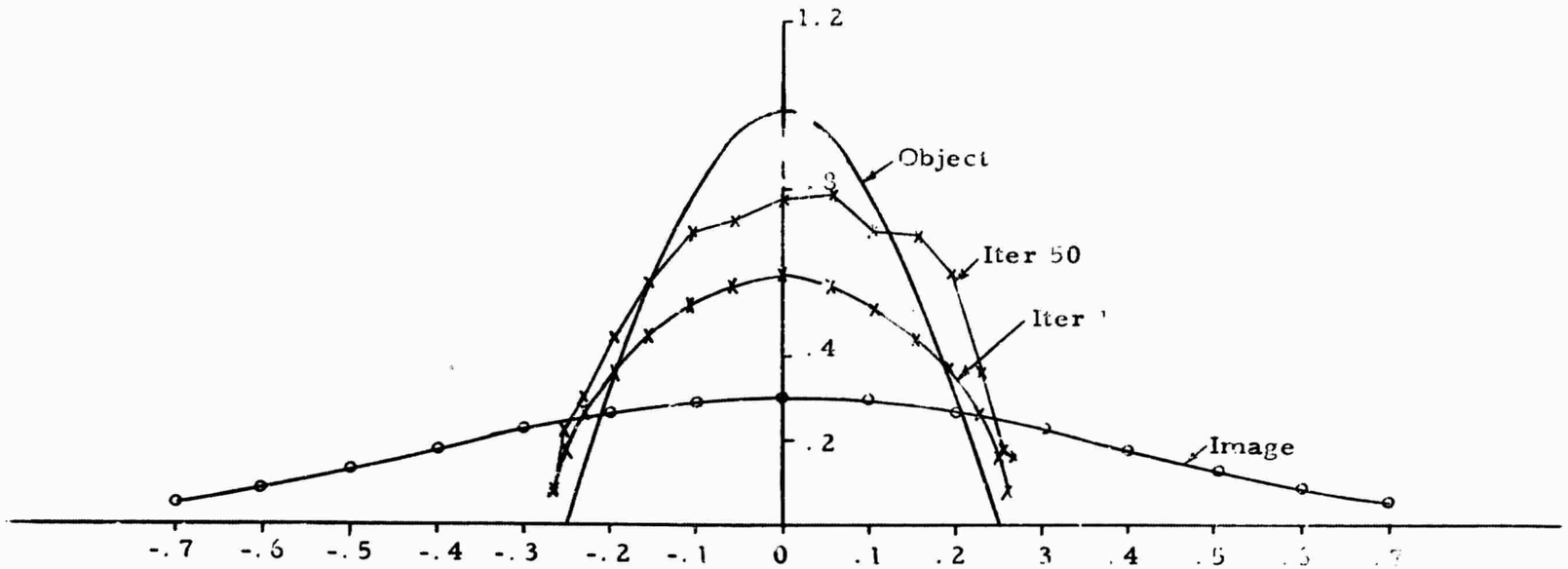

Figure 51. Restoration of the pulse cos (2πα) using Gauss quadrature in (155) with $\bar{x}$ = b.

further evident if we consider the problem of numerically integrating to find the area under a uniform pulse of unity height while using unity weights.)

The choice of quadrature and its relationship to the MSE is discussed later in the section entitled "MSE Variation--Choosing the A Matrix Parameters."

Comparison of the Phillips-Twomey and MSE Solutions

If we accept the use of successive approximations, then the essential difference between the Phillips-Twomey solution and the MSE solution is just the initial prior vector used. Figure 52 shows the asymptotic convergence of the SE vs. iterations for the two solutions. (The SE is increasing with increasing iterations since the 5x5 system used is approaching the quadrature error limited solution for $\lambda = 0$ shown in Figure 12.)

The figure indicates that the Phillips-Twomey solution has a lower SE than the MSE solution for iteration 1; however, after 2 iterations the two solutions differ only slightly in the rate of convergence. In fact, it is merely a matter of definition as to which solution reaches the asymptotic SE first. That is, iteration 1 for the MSE solution could have been defined as the zeroith solution, thus shifting the entire MSE solution curve 1 iteration to the left.

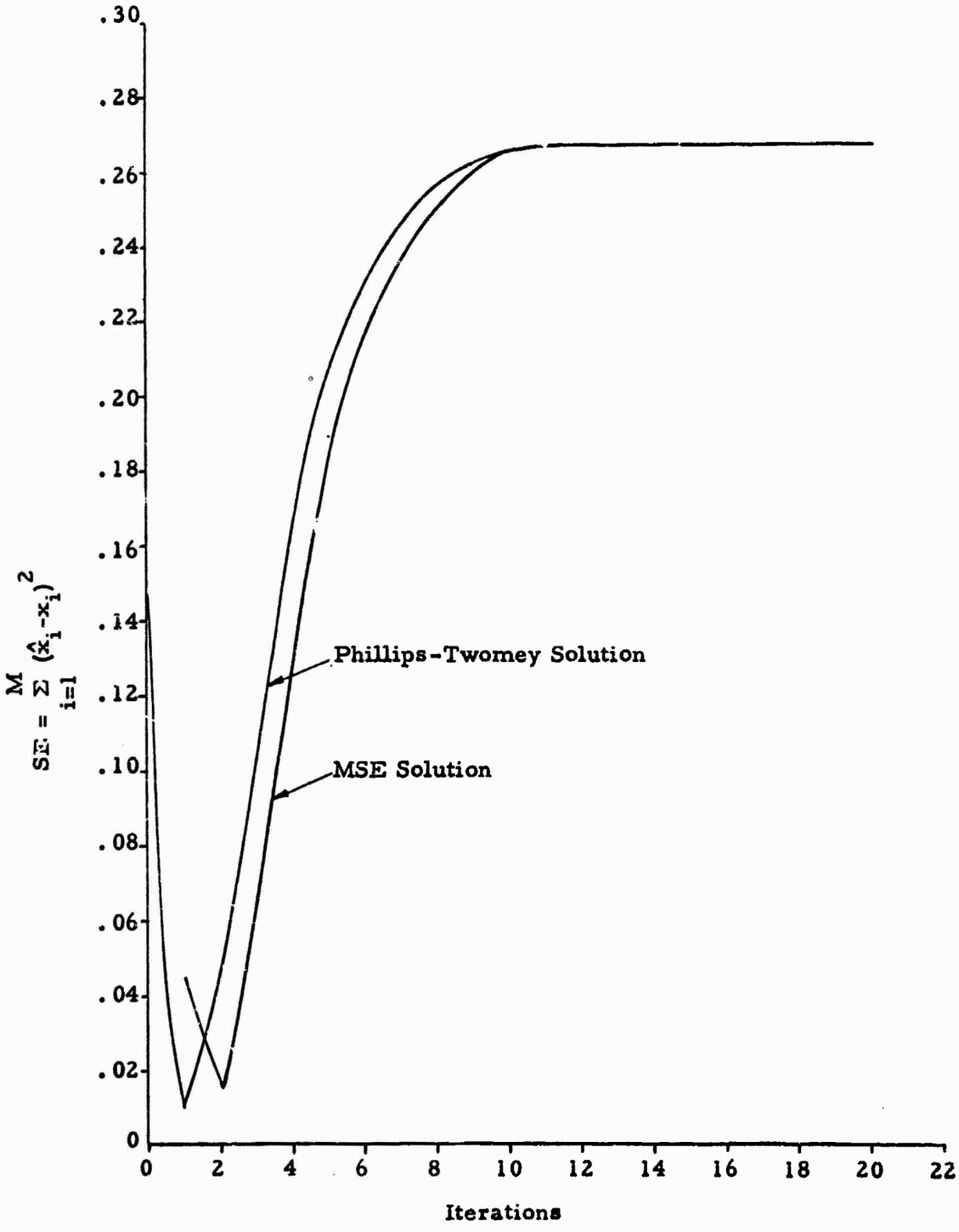

Figure 52. The comparison of the MSE and Phillips-Twomey solutions.

Dynamic programming restorations

Figures 53-56 illustrate the use of dynamic programming in restoring a uniform pulse source when R = 1.0. The prior vector solution (Bellman, 1965) was used for Figures 53-55. Notice that the solution for small prior weighting (Figure 53) is not symmetrical and is poorly behaved as one progresses toward the initial estimates on the left. This results because the entire A matrix is not used until the last scalar estimate x_M is found.

The solutions for the problem Bellman (1965) considered appear to be more regular for the initial x_i's than the above results show. However, the prior vector he assumed was closer to the true solution for the initial estimates. This would tend to reduce the initial irregular behavior.

Larger values of prior weighting improve the solution, as shown in Figures 54 and 55, and it appears that a value of λ between 1 and 10 would result in even greater improvement.

Results of the area constraint dynamic programming solution are depicted in Figure 56. These results, along with results from the matrix inverse solution using the area constraint, indicate that the area constraint is less stringent in controlling the characteristic oscillatory solutions even when prior weighting is used.

Based on the results obtained, it appears that the dynamic programming solutions investigated are less effective for optical restoration in one dimension than the MSE and Phillips-Twomey solutions.

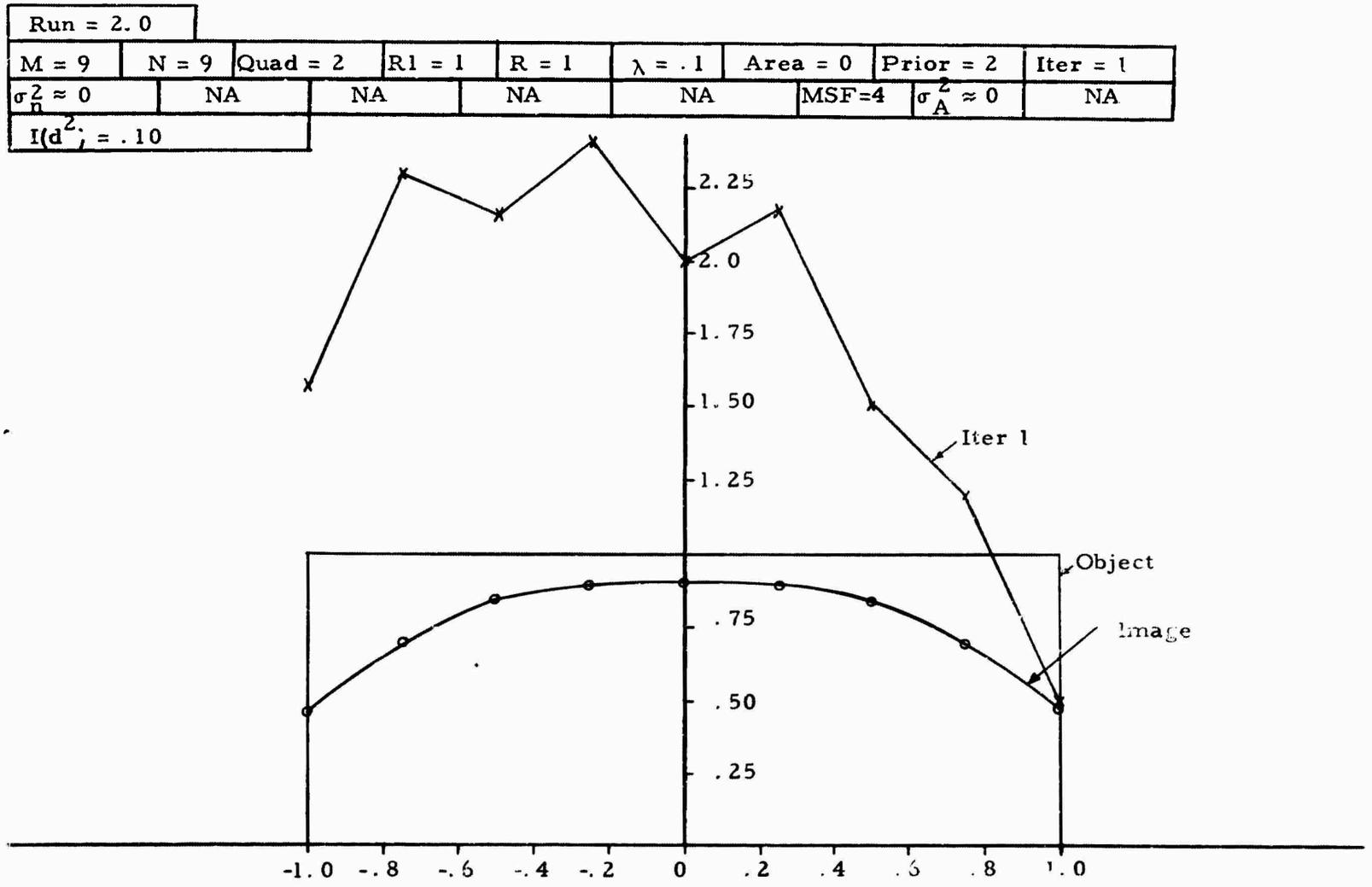

Figure 53. Restoration using the dynamic programming version of (47), (Bellman, 1965).

Run = 2.0								
M = 9	N = 9	Quad = 2	R1 = 1	R = 1	λ = 1.0	Area = 0	Prior = 2	Iter = 1
$\sigma_n^2 = 0$	NA	NA	NA	NA		MSF = 4	$\sigma_A^2 \approx 0$	NA
$I(d^2) = 1.12$								

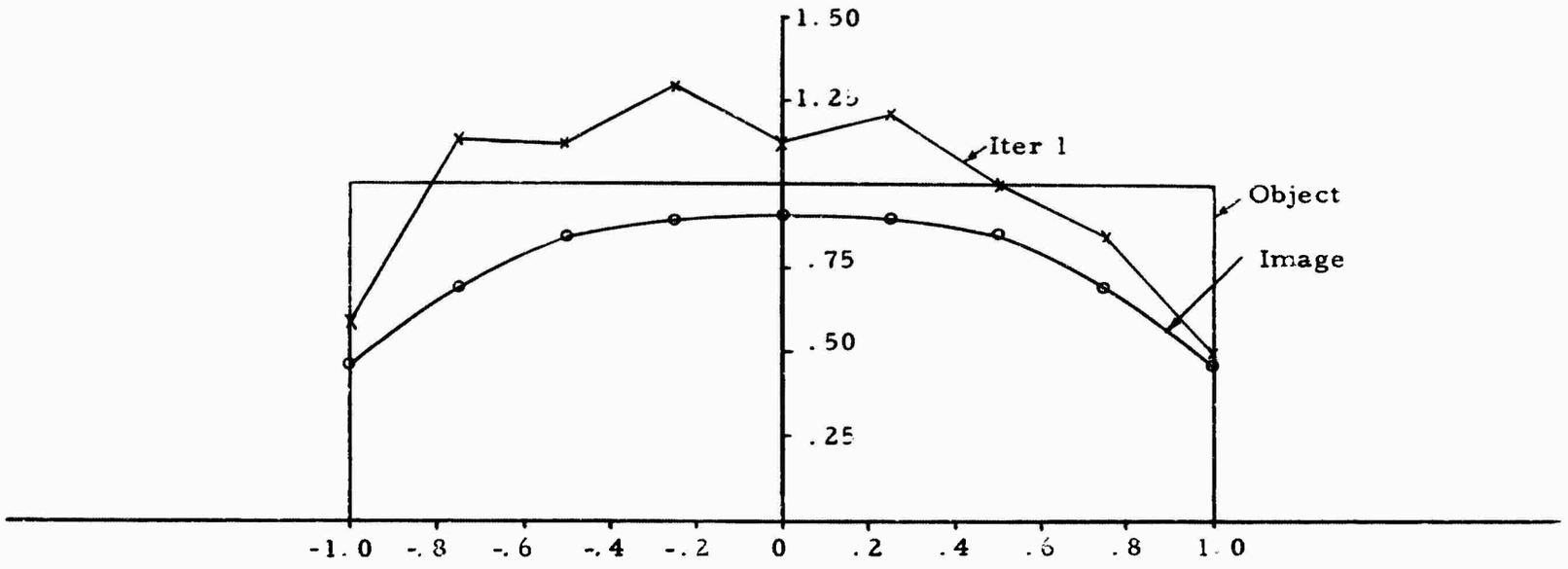

Figure 54. Restoration using the dynamic programming version of (47), (Bellman, 1965).

Run = 2.0								
M = 9	N = 9	Quad = 2	RI = 1	R = 1	λ = 10	Area = 0	Prior = 2	Iter = 1
$\sigma_n^2 \approx 0$	NA	NA	NA	NA		MSF=4	$\sigma_A^2 \approx 0$	NA
$I(d^2) = 1.13$								

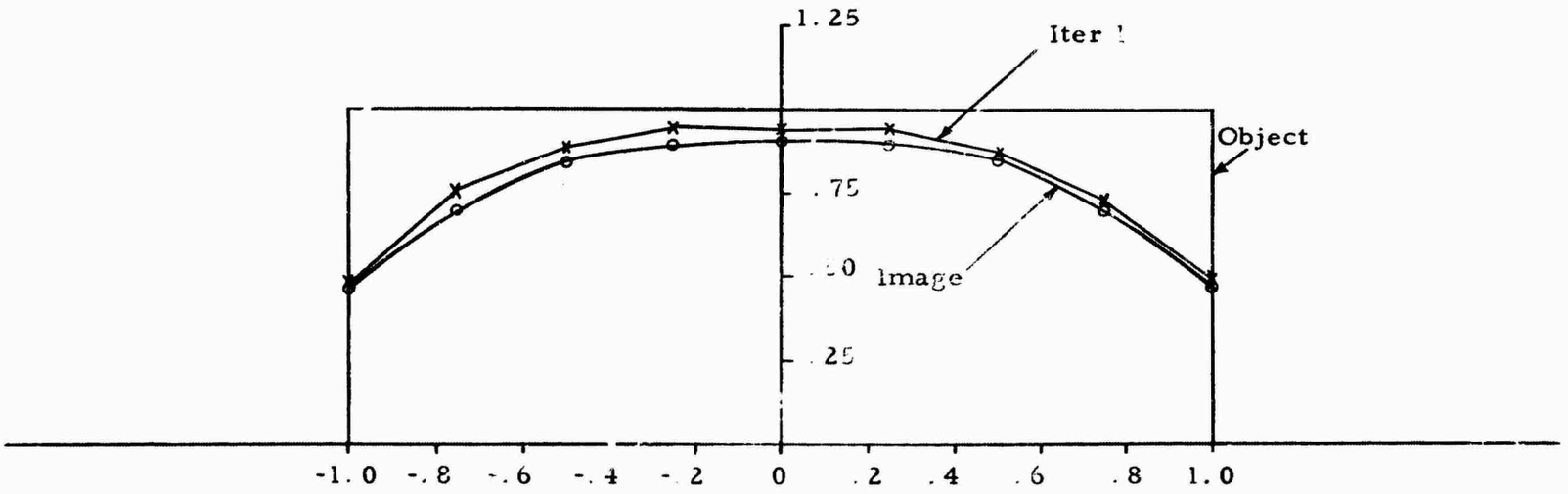

Figure 55. Restoration using the dynamic programming version of (47), (Bellman, 1965).

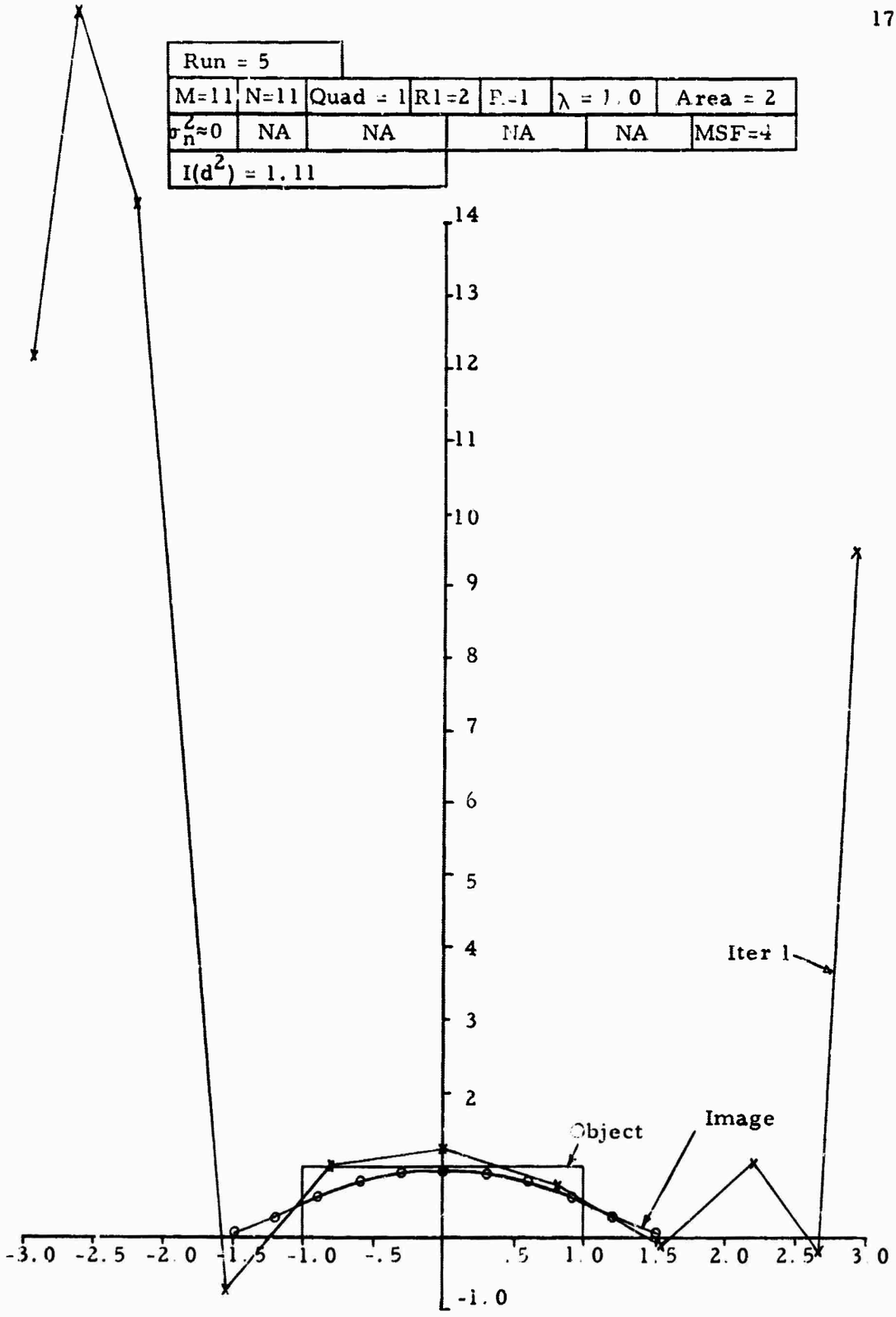

Figure 56. Restoration using the dynamic programming solution (78).

MSE VARIATION--CHOOSING THE A MATRIX PARAMETERS

This section considers the study of the MSE and the insight the MSE variation provides in enabling sound computational procedures to be determined. We first present further discussion which defines the A matrix parameters for the computational scheme we wish to consider. Next we present and discuss the MSE variation with and without a priori information. The final topic presents the procedures which have been developed and used in the preceding restorations for choosing the A matrix parameters.

The form of the MSE we consider here is that found by letting $K_x = \sigma_x^2 I$ and $K_n = \sigma_n^2 I$ in Equation (152). The resulting MSE is

$$MSE = \sigma_n^2 \, tr \left\{ [A'A + (\sigma_n^2/\sigma_x^2)I]^{-1} \right\} \qquad (188)$$

When the noise variance σ_n^2 is fixed, which is usually the case since it is determined by the experiment, there are essentially two ways to further minimize the MSE. We can either reduce σ_x^2 (assume more pseudo-prior information) or adjust the parameters that govern the A matrix.

Obviously one seeks the best measuring and predicting scheme which would entail optimal choices of M and N, the ξ_i's, the α_i's and the quadrature. This is indeed a difficult problem. In fact such optimal

choices would appear, intuitively, to depend upon the assumption of more prior object information (for example, concentrating on a certain class of well defined objects). In this paper we do not consider that such a priori information is available. Here we have sought to establish procedures for the general case. To accomplish this we have numerically studied the MSE as the various parameters were varied.

A Matrix Parameters

There is an uncountably infinite number of one-dimensional schemes we could consider. We have most extensively studied the scheme which uses M = N equally spaced measurement locations (ξ_i's) and M = N equally spaced prediction points (α_i's). The case when unity quadrature weights were used was most extensively studied, but results have also been obtained for Simpson's and Gauss quadratures. (The α_i's were spaced in accordance with the Gauss method when Gauss quadrature was used.)

The ratios R and R1 were used to relate the diffraction level to the measuring and predicting intervals being considered. In this case R is defined as the ratio of the point spread width to the prediction interval being considered.

We note then that the A matrix parameters are R, R1, M and the quadrature weights. Specification of these parameters determines the A matrix.

MSE vs. A Matrix Parameters for Large A Priori Uncertainty

Figures 57-59 show the MSE vs. R relationship for various values of R1 when $\sigma_x^2 = \infty$, and $w_i = 1.0$, $i = 1, \ldots, M$. The MSE was normalized with respect to noise in Figures 57-58 by considering MSE/σ_n^2. The MSE in Figure 59 was normalized with respect to dimensionality as well as noise by considering $MSE/\sigma_n^2 M$.

Notice that the MSE is lowest when we measure "over" the prediction points (R1 = 1.0) for small values of diffraction (the cases Austin (1966) treated when A = I), but when R > 3.8 then R1 = 2.0 is "best." Barnes (1966b) was able to prove that the A matrix is positive definite when R is finite and R1 = 1.0, which demonstrates that the ξ_i's and α_i's for any N = M can be chosen such that the MSE is finite. However, as the above results indicate, this does not preclude the choice of R1 other than unity to further reduce the MSE.

Another important and obvious feature can be deduced by noting the marked MSE increase for the curves of Figure 58 when compared with the curves of Figure 57. Figure 59 shows the absolute MSE increase for R1 = 1.0 when M increases from 2 to 3. This increase is most apparent for R > 1.0 and is attributable only to the addition of one more row and column in the A matrix. When we extrapolate to matrices of larger dimension (say 24 x 24 as used several times in the simulations) the MSE

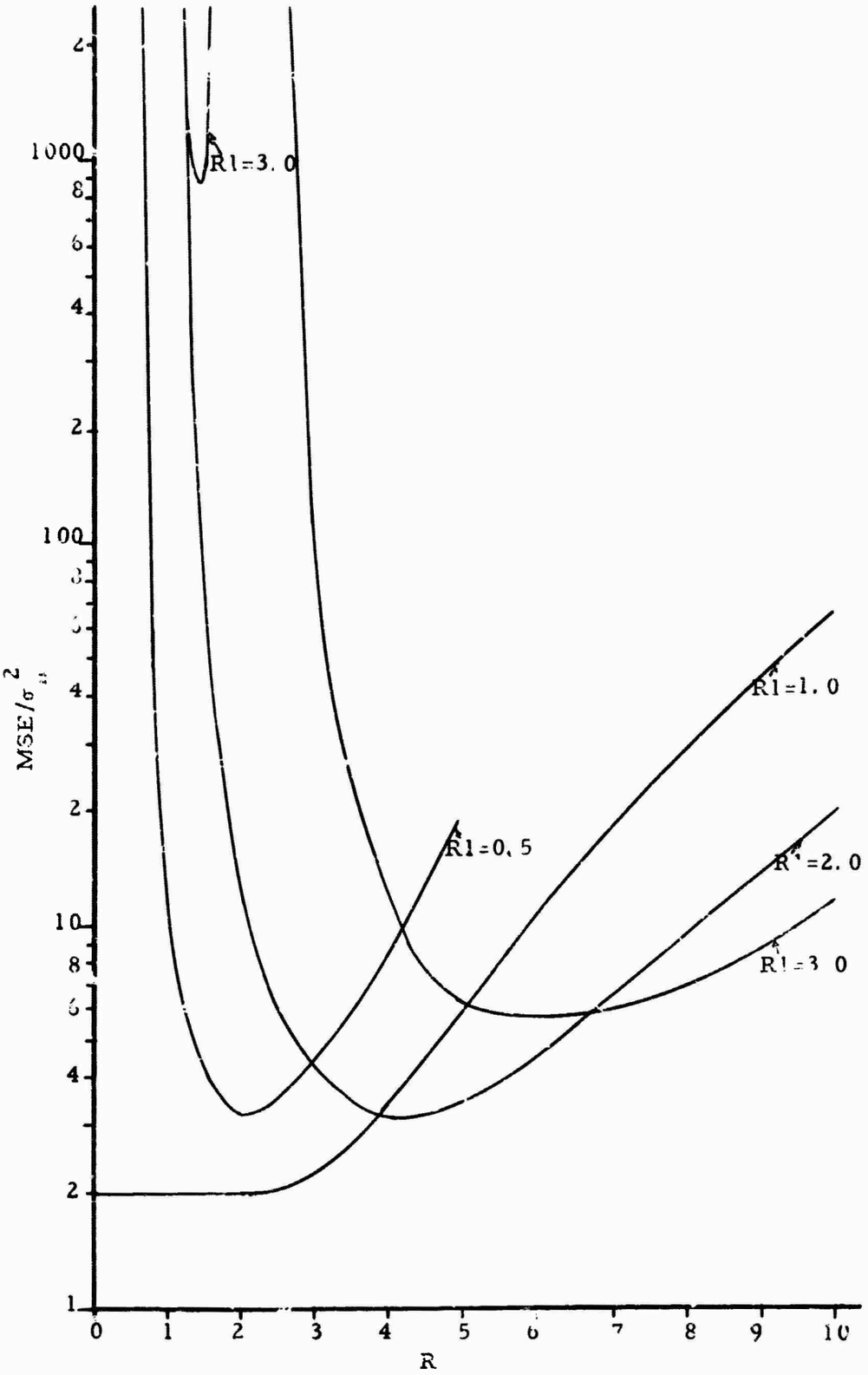

Figure 57. MSE/σ_n^2 vs. R for various values of R1 when $\sigma_x^2 = \infty$, M = 2, and $w_i = 1.0$.

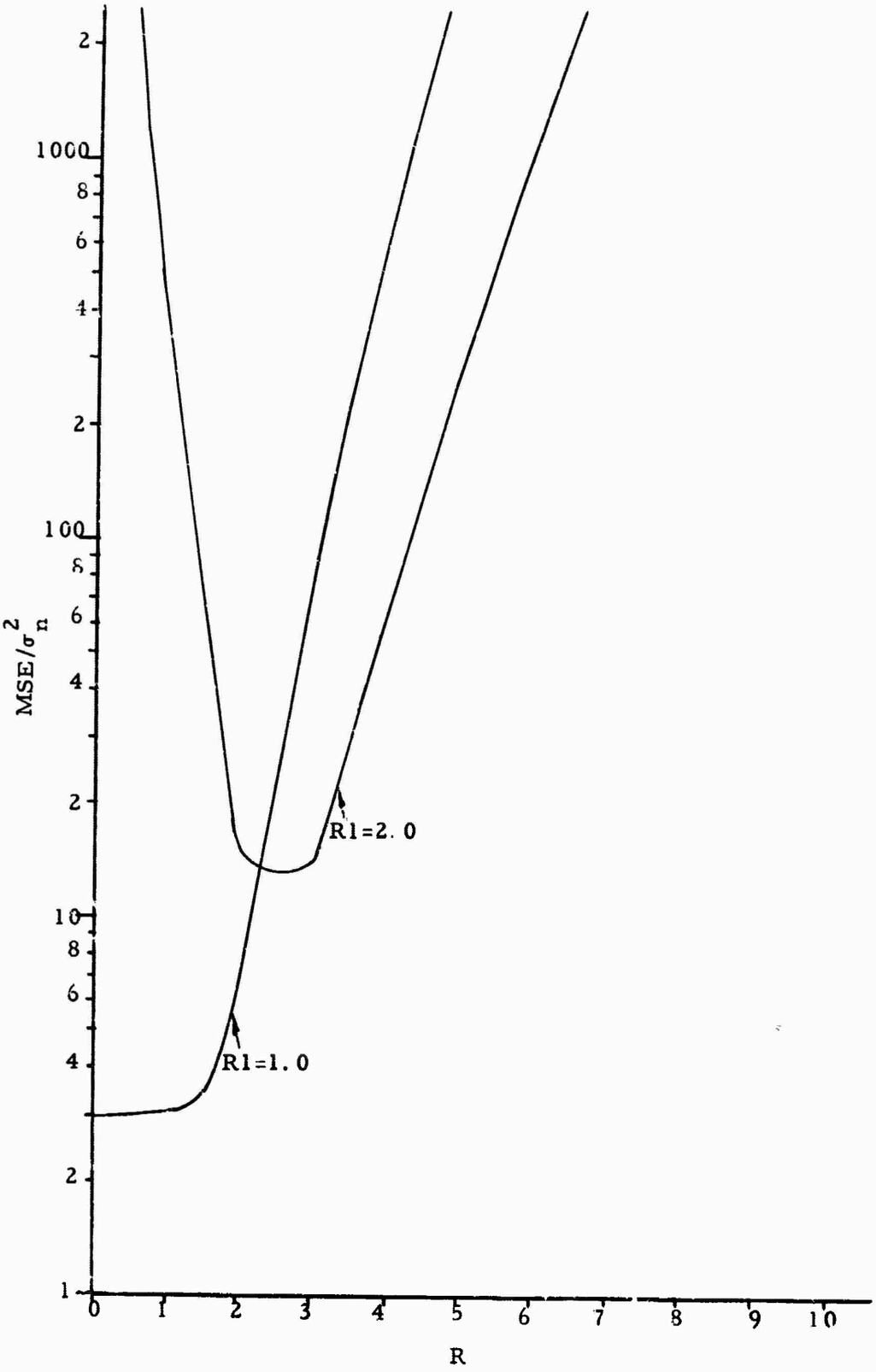

Figure 58. MSE/σ_n^2 vs. R for R1 = 1.0 and R1 = 2.0 when $\sigma_x^2 = \infty$, M = 3 and $w_i = 1.0$.

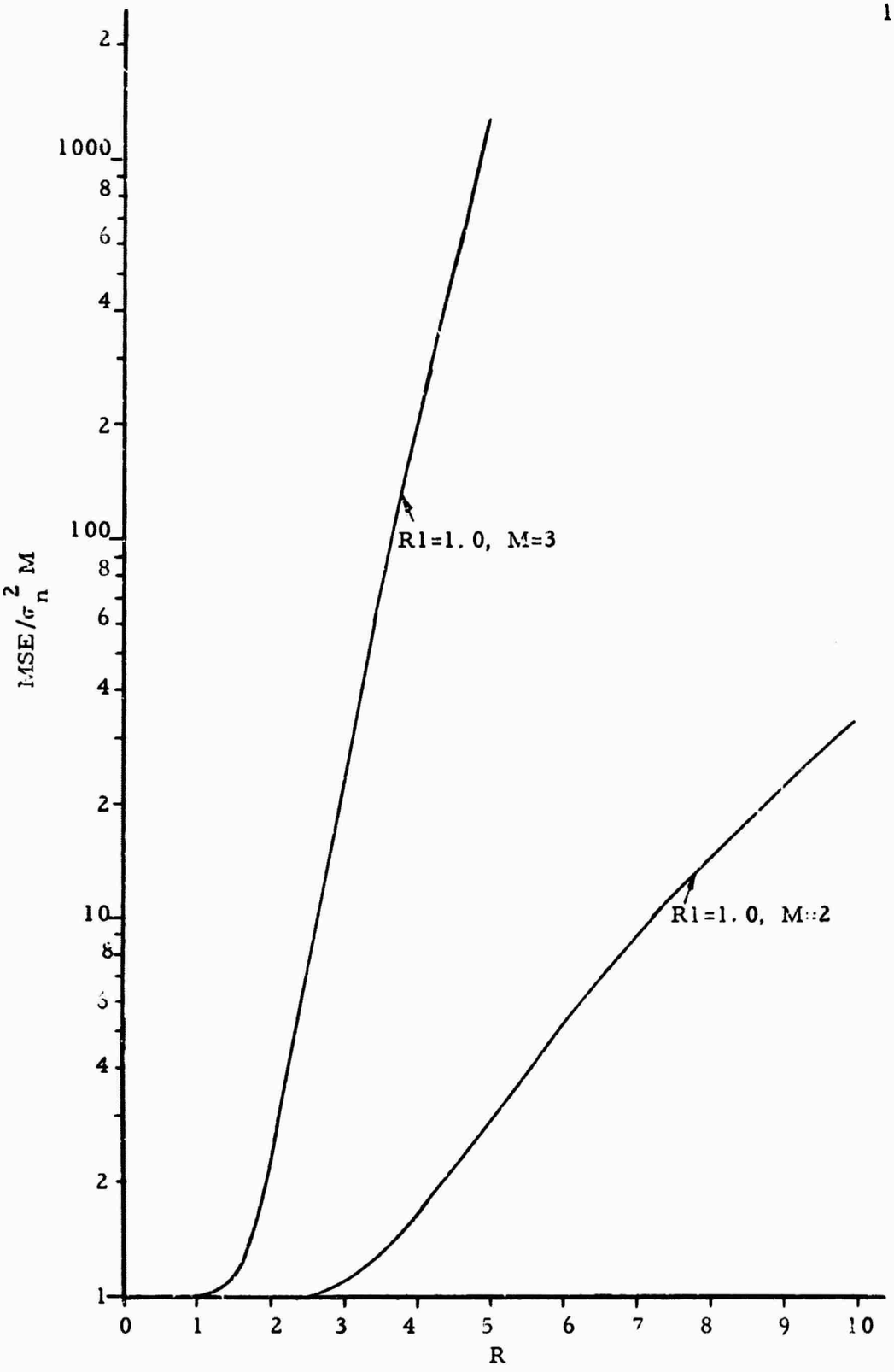

Figure 59. Comparison of $MSE/\sigma_n^2 M$ vs. R for R1 = 1.0 for the cases when M = 2 and M = 3.

for R > 1 and the values of R1 shown becomes astronomical. This result again emphasizes the amplifying effect that excessive diffraction has on the initial arbitrary noise level and essentially explains the reason for the oscillatory solutions when the A matrix dimensionality is large and σ_x^2 is infinite. The next figures we present, which are explained below, illustrate how finite σ_x^2 alleviates this effect.

MSE vs. A Matrix Parameters for Finite A Priori Information

Figure 60 illustrates the MSE behavior for various parameters as the prior information is varied. The general behavior agrees with intuition since as prior information increases (σ_x^2 decreases) we infer the possession of more knowledge about the object and the MSE decreases accordingly.

Figure 61 shows the measuring and predicting intervals for the top four curves shown in Figure 60. After viewing Figure 61 refer again to Figure 60. Now consider the asymptotic MSE for the top four curves. These top four curves show the advantage of choosing a measuring interval which is wider than the prediction interval (R1 is large) since the MSE decreases by at least a factor of 10^7 while the diffraction level, weights, and M remain fixed. The bottom two curves show a similar reduction for a smaller diffraction level.

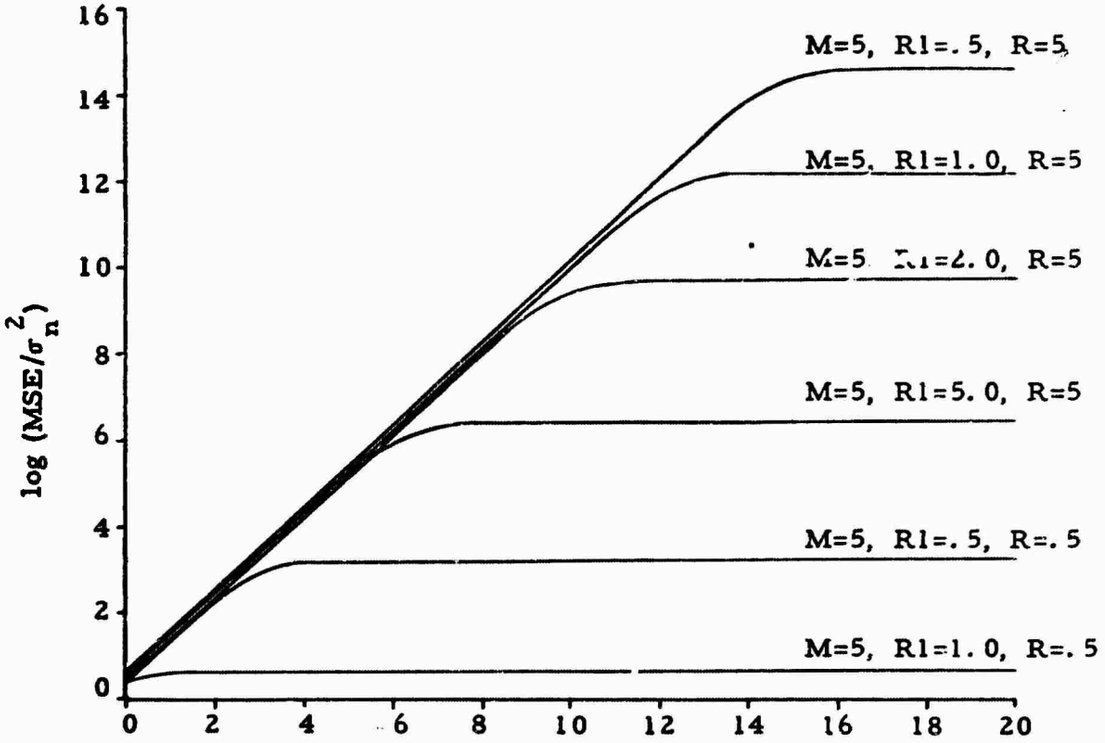

Figure 60. The log (MSE/σ_n^2) vs. log (σ_x^2/σ_n^2) for various A matrix parameter values when $w_i = 1.0$.

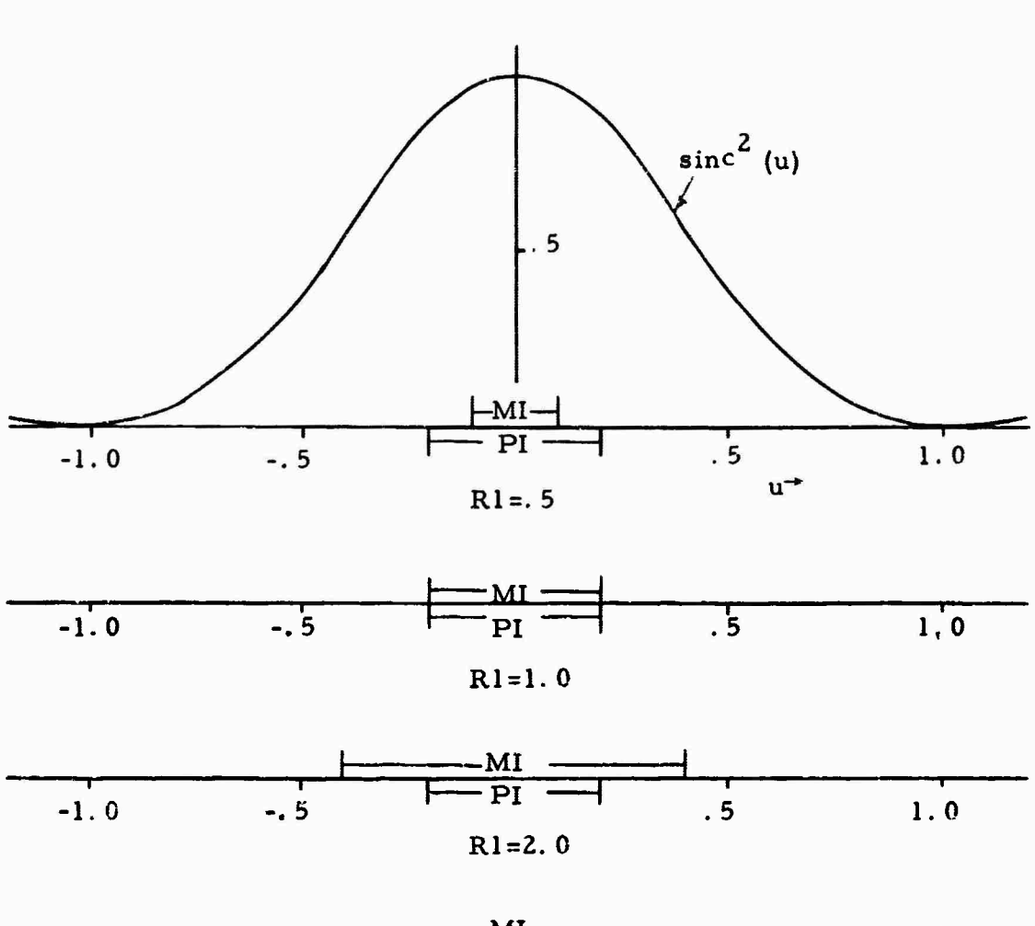

Figure 61. Measuring and predicting intervals superimposed with the point spread function for the top four curves in Figure 60. The symbols MI and PI represent, respectively, the measuring and predicting intervals.

Figure 60 exhibits results for the M = 5 case. Figure 62 presents curves, in the same manner as Figure 60, for the simulated results shown in the previous section. From Figure 62 it is evident that the prior weighting necessary to produce the simulations was great enough so that we are operating in the region where the curves blend together, as shown in Figure 60. In this region choosing R1 large still has an effect on the MSE, but it is not nearly as pronounced as for the asymptotic region of the curves. However, the computational error involved in the matrix inversion process was noticeably reduced by choosing R1 large.

Choosing the A Matrix Parameters

Now that we have presented curves showing the general behavior of the MSE with the A matrix parameters we discuss the guidelines for choosing these parameters. These choices depend not only upon the MSE but upon other errors, the most prominent being the quadrature error, the computational error, and the error caused by assuming σ_x^2 too small. In choosing these parameters it is evident that we must make these choices regardless of the diffraction and noise. Thus, although we seek to improve the distorted object in the presence of these effects, the diffraction and noise must be regarded as fixed quantities when making these choices.

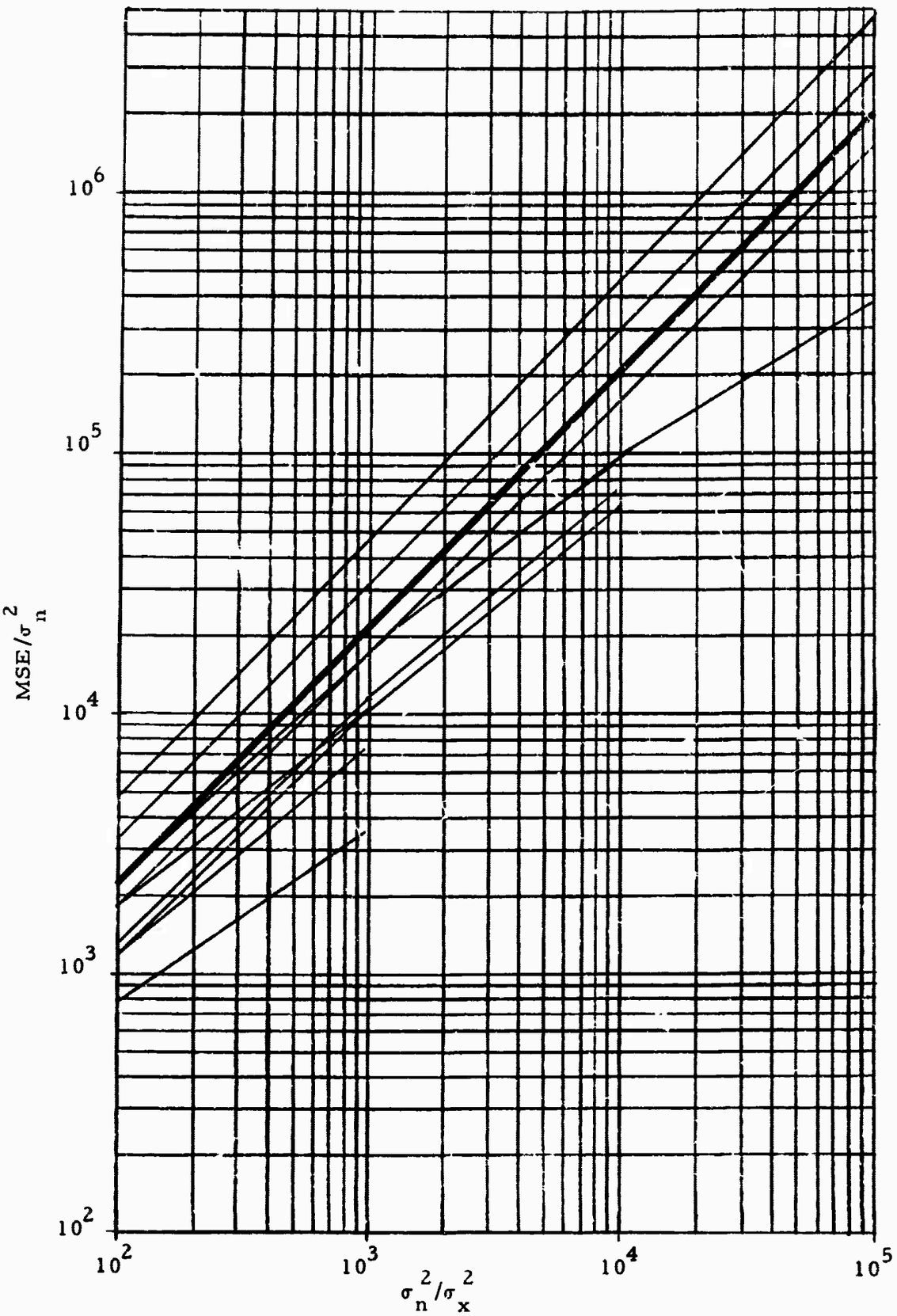

Figure 62. Curves of MSE/σ_n^2 vs. the ratio σ_n^2/σ_x^2, showing the region used in the simulated restorations.

Choosing the quadrature

The determination of which quadrature is "best" or better than another is a difficult problem. One cannot accurately estimate the quadrature error without knowledge of the integrand, which for the restoration problem we have considered is unknown.

Before discussing the quadrature choice we briefly review the quadratures we have considered. For a more comprehensive treatment the reader is referred to Kopal (1961), Krylov (1962) or the more tutorial McCracken and Dorn (1964).

The Gauss quadrature method is a very powerful tool to use in numerical integration. Here we refer to power as describing the number of points necessary in the approximating summation to provide a specified accuracy. To illustrate just how powerful the Gauss method is, consider that a polynomial of degree m approximates the integrand to a specified accuracy. Now suppose we use the Gauss system with a summation limit of M points. The Gauss system is sufficiently powerful that when $m = 2M-1$ the integration is performed exactly with no error. For example, suppose we decide that the integrand is sufficiently approximated by a polynomial of degree 49. This means that only 25 points are required to perform the integration exactly. Simpson's summation, on the other hand, is only capable of integrating 3^{rd} degree integrands exactly. Herein lies the major advantage of Gauss quadrature over Simpson's quadrature (and, as the above references point out, over essentially all other quadratures).

The only disadvantage of Gauss quadrature in comparison with other quadratures is one of complexity. The complexity is due to the use of unequally spaced integrand arguments. This disadvantage is not very significant in the optics problem because we can predict the object for unequally spaced points just as well as equally spaced points. The only possible drawback would be the specification of the point spread function at these points. If the point spread function is analytically approximated, this would not be a problem. For the case when the point spread function is measured some extra care may be necessary to ensure that the point spread function is accurately specified for all ranges of its argument.

Now that we have briefly reviewed the quadratures considered we continue on with the discussion of quadrature choice.

Recall that the total error is comprised of both MSE and quadrature-error. These two errors are both related to the system dimensionality. We have seen that increased dimensionality implies increased MSE, but it is well known that increased dimensionality reduces quadrature-error. Thus, it is logical to assume that a choice of quadrature should involve the trade-off between MSE and quadrature-error. The following results illustrate this trade-off.

Refer again to Figures 49, 50, and 51. The MSE/σ_n^2 for the Gauss case is 1.43×10^2, for the Simpson's case 1.26×10^2, and for unity weights 1.23×10^2. The fact that the Gauss quadrature resulted in a

greater MSE/σ_n^2 is easily explained. In all three cases the actual number of points used was equal, but the effective dimensionality of the Gauss system is much greater than either of the other quadratures. Similarly, the effective dimensionality in the Simpson's case is greater than the unity weights case. Visually (which is a good measure of total error) the Gauss system produces a better restoration. Since the MSE actually increased, but only slightly, the discrepancy between visual error and MSE can logically be attributed to a reduction in quadrature error. In fact, it appears that the reduction in quadrature error more than compensates for the slight increase in MSE. Thus, in this case one would choose Gauss quadrature. Furthermore, we can infer from Figures 60 and 62 and other numerical evidence that as long as one uses sufficient a priori information to be in the region where the curves blend together, use of the Gauss quadrature will consistently more than compensate for the slight increase in MSE. And in this region it is recommended that Gauss quadrature be used.

On the other hand, as σ_x^2 approaches infinity the increased effective dimensionality for the more powerful quadratures may increase the MSE above the compensating reduction in quadrature error. For this region of small a priori information we could either reduce the Gauss system dimensionality or perhaps successfully use a less powerful quadrature.

Choosing the ratio of the measuring and predicting intervals (R1)

The diffraction range considered in this paper is $R > 1.0$. For $R \leq 1.0$, the image becomes a much better prior guess for the object, and we can heavily weight the image in the restoration procedure.

Consider the images for the Range $R > 1.0$. For the noise levels considered in this paper the image width (noise to noise) is, grossly speaking, the width of the point spread function. Thus, the prediction and measuring intervals will be "under" the main lobe of the point spread function. When this is the case, as has been demonstrated in Figures 57, 58, and 60, it is profitable to choose R1 larger than unity so that the measuring interval is wider than the prediction interval. The following discussion illustrates what is happening to the A matrix elements as R1 changes and then presents a guide for choosing an upper bound for R1.

Refer again to Figures 57 and 58. The tendency for A to approach a singular condition for $R1 = 1.0$ and increasing R is clearly shown. As one studies the A matrix elements for the values of $R > 1.0$ and $R1 = 1.0$, it is seen that they become more and more alike. On the other hand, choosing $R1 > 1.0$ for values of $R > 1.0$ tends to make the A matrix elements differ, and generally speaking, if R1 is not made too large, the MSE will decrease. The choice of an upper bound for R1 involves a trade-off between two factors.

The first is seen in Figure 57. For matrices of small dimension it is evident that the A matrix will be singular for certain choices of R1 even when R > 1.0. However, it is evident intuitively that if the dimensionality is large enough (3x3 or greater) and R1 is chosen only large enough so that the measuring and predicting intervals both lie within the point spread function width, then the A matrix will not be singular. An analytic proof of this statement is not given, but the available numerical evidence shows that A is nonsingular for this choice of R1.

The second limitation on large R1 is imposed by system noise (σ_n^2). For R > 1.0 it is evident that as R1 increases above unity then the image is being measured where the SNR is decreasing. Thus, if the system noise is great enough a choice of R1 extended to the limit described above may be suboptimum.

Under these conditions the following guide for choosing R1 is stated. For matrices of dimensionality 3x3 or greater and R > 1.0 choose R1 as large as possible, while ensuring that the measuring interval is "under" the point spread function and small enough that the measurements are not excessively noisy. This notion is clearly a compromise between system noise (σ_n^2) or SNR and the location of the measurements such that the matrix ($A^tA + \sigma_n^2/\sigma_x^2 I$) is non-singular enough to be successfully inverted.

Choosing A matrix dimensionality

Throughout the paper we have assumed that the A matrix is of dimensionality NxM and that $M \leq N$. When $M < N$ we have the over-determined case in which we measure the image more often than the number of unknowns we have to predict. In some problems it is advantageous, for statistical reasons, to choose $M < N$ and use the over-determined solution. This case was considered in the simulation presented in Figure 46. Recall that the image was excessively noisy (SFR = .6-1.6). In an effort to restore the object several A matrices with $M < N$ were used. The dimensionality shown in Figure 46 was 50x11. However, in all of the cases when $M < N$ no appreciable improvement was apparent. Based on this evidence, it appears that choosing $M < N$ is of no particular advantage in the restoration problem. When $M = N$, then the choice of how large to make N or M is to be considered. This choice is discussed as follows.

As previously stated, the MSE is proportional to N, and the quadrature error is inversely proportional to N. Thus, N should be chosen by studying the trade-off between MSE and quadrature error. For the one-dimensional restorations considered, σ_x^2 was assumed small enough so that reasonably large matrices could be successfully used. The largest matrix used was a 65x65, and it resulted in restoration improvement. For the one-dimensional objects considered, it appeared that the 48x48 Gauss system was adequate for reducing quadrature error.

To obtain the same quadrature error for two dimensional objects, the A matrix dimensionality will increase roughly as N^2 compared to N. Thus, for two dimensional objects we will have to consider more closely the trade-off between MSE and quadrature error.

Choosing σ_x^2 and the number of iterations

The choice of σ_x^2 involves the compromise between MSE and the error involved by assuming too much prior information. One desires to choose σ_x^2 small enough to enable stable computational results to be obtained and yet large enough to reduce the smoothing effect or a priori image weighting.

The simulated results have indicated that σ_x^2 may be sequentially determined by visually judging restorations for several values of σ_ϵ^2. It has also been demonstrated that one can circumvent the need to continually invert matrices for each σ_x^2 in order to vary prior information. Prior information can be easily varied by using successive approximations.

Considering that the other parameters are fixed, a good procedure to use is to vary σ_x^2 until the solution appears (visually) to be stable and then iterate until computational error limits the restoration.

SUMMARY AND CONCLUSIONS

Synopsis of the Paper

In this section a brief synopsis of the entire paper is presented and the important results are stated.

The basic problem considered is one of restoring an optical object which has been diffracted and corrupted by noise. Previous work has shown that if the object is of finite extent then the restoration process is limited by noise and not diffraction. After reviewing previous approaches to the problem, it was evident that improvement could be made by providing for noise in the restoration procedure.

In order to present some insight into the diffraction process the imaging equation was derived. The form of this equation indicates that a linear transformation of the object constitutes the basic image-object relationship and that diffraction implies the obscuration of object detail. Since numerical techniques were to be used, the discrete version of the imaging equation was presented and the quadrature error was introduced. Next the straightforward no-noise matrix inverse solution to the problem was presented. Two major difficulties in using this solution are that excessive image accuracy may actually be necessary to effect the solution and that for the diffraction range of interest ($R > 1.0$) the A matrix is nearly singular and is difficult to invert.

The next section presented previous noiseless work by Phillips (1962), Twomey (1963) and Bellman, et al. (1964, 1965). Their work demonstrated that a priori information in the form of a constraint could be successfully used to alleviate the difficulties just mentioned.

In order to introduce and discuss the uncertainty caused by noise, analytical results for a detection example and for the error variance obtained in estimating the separation between two point sources were presented. Here the basic trade-off between diffraction and noise in regaining obscured object information was presented and discussed in detail. The basic trade-off is that diffraction essentially amplifies the prevailing noise level, and in order to regain object information for excessive diffraction the SNR must be increased.

Turning again to the restoration problem, object estimation in the presence of additive and detector noise was discussed. Both the Baysean and MSE approaches in the estimation of the object were considered. In using the Bayes' approach it was necessary, because of mathematical tractability, to use the suboptimum scheme of operating on an image estimate to perform the restoration. The Baysean approach did provide insight as to how a priori information enters into the solution, but the suboptimal scheme did not indicate the use of a priori information in the inverse operator. On the other hand, the MSE approach did indicate how a priori information can be used in the inverse operator and

furthermore indicated that the smoothing process in the noiseless case treated by Phillips (1962) is actually a trade-off between the noise level and a priori information. Also discussed in the Object Estimation section was the assumption of high SNR. Using this assumption the Baysean estimates for Gaussian statistics are equivalent to the MSE estimate.

In the next section simulated object restorations were presented. Here various parameters were defined which indicated how well the restoration process performed for various objects and for various diffraction and noise levels. These results demonstrated that object restoration is possible. The amount of restoration improvement is dependent upon both diffraction and noise. A specific statement is that, generally speaking, improvement is possible for the diffraction range $1.0 < R < 10$ when the sample mean noise level is as high as 1 to 3 percent of the image maximum.

Next, numerical results on the variation of the MSE were presented and discussed. For fixed σ_n^2 and diffraction in general the only recourse to further reduce the MSE is to consider varying the A matrix parameters or the a priori variance σ_x^2. The following guidelines were evident from the results presented in this section. The quadrature used most profitably was the Gaussian system. For the diffraction range considered the measuring interval should be chos... ...ner equal to or greater than the predicting interval. The upper limit of the measuring interval width

should be chosen such that it does not exceed the point spread function width or include excessively noisy image measurements. It is advisable to use M = N. The choice of an upper limit on N involves the trade-off between the error caused by excessive weighting of the a priori image and the improvement in quadrature error. The a priori variance σ_x^2 should be chosen small enough to obtain an initial stable solution which can be successively iterated until computational accuracy limits the restoration. These results when applied to the MSE estimate and coupled with the sequential size estimation procedure discussed previously constitute the basic computation scheme developed in this paper.

Future Research

The outstanding applied research need is to actually perform a restoration experiment which uses the procedures developed in the paper. Such an experiment would provide information on the more real noise levels encountered in practice and should lead to the intelligent use of repeated samples of the noisy image and A matrix.

The general theoretical or analytical solution to the problem remains to be obtained. To solve this problem it appears that we need to specify eigenfunctions and eigenvalues for a general inverse imaging kernel, as discussed by Barnes (1966a). This is indeed a difficult problem, but certainly object restoration will not be complete until this information is available.

Throughout the simulated results we have considered that the noise is white and Gaussian. Band limited noise should also be considered. When bandlimited noise is considered, perhaps the diffraction effects will not be as severe as those shown in this paper. However, it should be mentioned that the computational error is not bandlimited.

APPENDIX

Two-Dimensional Gauss Quadrature

In two dimensions the imaging equation is

$$v(\xi,\eta) = \int_{a_2}^{b_2}\int_{a_1}^{b_1} h(\xi-\alpha,\eta-\beta)\, x\,(\alpha,\beta)\,d\alpha\,d\beta \qquad (189)$$

The two-dimensional Gauss quadrature for (189) can be written

$$b(\xi,\eta) = \frac{(b_2-a_2)(b_1-a_1)}{4} \sum_{m=1}^{N}\sum_{n=1}^{N} H_m H_n h(\xi-\alpha_n,\eta-\beta_m)\, x\,(\alpha_n,\beta_m) \qquad (190)$$

where

$$\alpha_n = \tfrac{1}{2}(b_1-a_1) + \tfrac{1}{2}(b_1-a_1)y_n$$

$$\beta_m = \tfrac{1}{2}(b_2-a_2) + \tfrac{1}{2}(b_2-a_2)y_m$$

and the quantities y_n, y_m, H_m and H_n are defined from tables of coefficients for a given N (Krylov, 1962).

We can bring the constant $\frac{(b_2-a_2)(b_1-a_1)}{4}$ inside the summation of (190) and define a new summation index j as follows:

$$\begin{aligned} j &= 1 \quad \text{when } m = 1,\ n = 1 \\ &\vdots \\ j &= N \quad \text{when } m = N,\ n = 1 \\ &\vdots \\ j &= N^2 \quad \text{when } m = N,\ n = N \end{aligned}$$

Thus

$$b(\xi_i, \eta_i) = \sum_{j=1}^{N^2} w_j h(\xi_i - \alpha_j,\ \eta_i - \beta_j) x(\alpha_j, \beta_j)$$

or

$$b_i = \sum_{j=1}^{N^2} a_{ij} x_j$$

If we consider L image measurements then the final form is

$$b = Ax \tag{191}$$

where the A matrix has L rows and N^2 columns.

Recurrence Relation for Area Constraint

Here we derive the recurrence relationship (73) which is

$$f_M(b, c) = \min_{x_M} [f_{M-1}(b-a_{(M)}x_M,\ c-w_M x_M)]. \tag{192}$$

We define

$$f_M(b, c) = \min_{\substack{x_M \\ \vdots \\ x_1}} R_M(x) = \min_{\substack{x_M \\ \vdots \\ x_1}} [\lambda(\sum_{i=1}^{M} w_i x_i - c)^2 + \sum_{i=1}^{N}(\sum_{j=1}^{M} a_{ij}x_j - b_i)^2]. \tag{193}$$

According to the optimality principle (stepwise optimization principle) we can write,

$$f_M(b, c) = \min_{x_M}[\min_{\substack{x_{M-1} \\ \vdots \\ x_1}} \lambda(\sum_{i=1}^{M} w_i x_i - c)^2 + \sum_{i=1}^{N}(\sum_{j=1}^{M} a_{ij}x_j - b_i)^2] \tag{194}$$

and there exist arguments arg_1 and arg_2 such that the term in brackets can be written

$$f_{M-1}(arg_1,\ arg_2) = \min_{\substack{x_{M-1} \\ \vdots \\ x_1}} \left\{\lambda(\sum_{i=1}^{M-1} w_i x_i - arg_2)^2 + \sum_{i=1}^{N}(\sum_{j=1}^{M-1} a_{ij}x_j - (arg_1)_i)^2\right\} \tag{195}$$

Now we have to find transformations arg_1 and arg_2 which enable (195) to equal the term in brackets in (194). Thus we equate

$$(\sum_{i=1}^{M} w_i x_i - c)^2 = (\sum_{i=1}^{M-1} w_i x_i - arg_2)^2 \tag{196}$$

and

$$\sum_{i=1}^{N}\left(\sum_{j=1}^{M} a_{ij}x_j - b_i\right)^2 = \sum_{i=1}^{N}\left(\sum_{j=1}^{M-1} a_{ij}x_j - (a_M g_1)_i\right)^2 \qquad (197)$$

From (196)

$$\arg_2 = c - w_M x_M \qquad (198)$$

If we fix i in (197) we obtain

$$\arg_1 = b - x_M a_{(M)} \qquad (199)$$

where $a'_{(M)} = (a_{1M}, a_{2M}, \ldots, a_{NM})$ and is the transpose of the M^{th} column in the A matrix. Using (198) and (199) we can write the recurrence relation (192).

Computer Programming

The computational results of this paper were obtained using the IBM 1620 and IBM 7094 digital computers. The 1620 was available at Utah State University. The 7094 was available through the Western Data Processing Center at UCLA in Los Angeles, California.

Two facets of computer programming should be mentioned. First, considerable effort was made to ensure the correct estimation of image accuracy. (This effort is discussed under Image and A Matrix Accuracy in the Simulated Object Restorations section.) A similar effort was made to ensure that the solutions were correct. A good check of these solutions was available by comparing the last solution vector component x_M in both the matrix inversion solutions and the corresponding dynamic programming solutions. Second, as a further check on the solution accuracy and to study the inversion process, the matrix inverse was checked for each solution. In all of the cases shown when σ_x^2 was finite, σ_x^2 was assumed small enough so that no special programming sophistication was required to ensure a correct matrix inverse. When σ_x^2 was infinite or very large, results were obtained on the 1620 which allowed the use of excessive accuracy (up to 28 significant digits). This excessive accuracy enabled results to be obtained, as previously presented, when the A matrix was nearly singular.

As an example of the computer programming necessary to simulate the restoration of optical objects, we present the following computer program which represents roughly 50 percent of the computational effort. This computer program was used to simulate the matrix solutions in the restoration problem for the sources composed of cosine pulses from $\cos [k\pi(\alpha-\varphi)]$ as shown in (166) and (168). The random number subroutine

was obtained from the Western Data Processing Center in Los Angeles. The symbols used and the computational sequence are indicated as the program proceeds.

```
C       PROGRAM 10E AND 10A ON 7090 FORTRAN 4
        DIMENSION F(301),PSI(32),B(32),TAU1(32),W(32),A(32,32),H(32)
        DIMENSION PRIOR(32),AX(32,32),Z(32),X(32),Z1(32),ATP(32,32)
        DIMENSION ATEMP(32),ATEMP1(32),PSTAU(20),PSTAU1(20)
        DOUBLE PRECISION AX,XX,B,F,SUM,DLTAU,TAU,SRCE,PTSPRD
        DOUBLE PRECISION ARG,PSI,PTAU,ABTAU
        DO 66 I=1,32
        PRIOR(I)=0.0
     66 CONTINUE
        READ (5,619) NBR1
    619 FORMAT(1H I4)
        DO 1001 NBR=1,NBR1
C       ONE DIMENSIONAL PREPROSSING PROGRAM
C       INTERNAL CONVOLUTION FOR IMAGE AT SPECIFIED ARGUMENTS
C       QUADRATURE=SIMPSONS SUMMATION FOR INTERNAL CONVOLUTION
C       QUADRATURE=GAUSS OR SIMPSONS FOR A MTX AND WEIGHT VECTOR W
C       SPECIFYING N=NBR IMAGE MEASUREMENTS, RUN= NBR OF SAMPLING SCHEME
C
C       IF IMAGE NOISE PRESENT THEN SET RNOISE=1.0
C       NOISE DIST=N(0,VARN),  B(I) DIST=N(B(I),VARN)
C       LS=NUMBER OF SAMPLES FOR EACH B(I)   MUST ENTER A VALUE FOR
C       LS WHEN RNOISE=1.0
        READ (5,500) N,RUN,AMTX,AMTXIV,AXMTX,RNOISE,VARN,LS
    500 FORMAT(1H I4,5F7.2,F16.8,I4)
C       T1 AND T2 SPECIFY THE IMAGE MEASURING INTERVAL
        READ (5,680) T1,T2,CK1,RNOSE,VARN1
    680 FORMAT(1H 3F16.8,F7.2,F16.8)
C       COMPUTING EQUALLY SPACED MEASUREMENT POINTS FOR THE IMAGE
C       IF READ IN  (PSI(I),I=1,N) SET T1=0.0
        IF(T1)681,682,681
    682 READ (5,501) (PSI(I),I=1,N)
    501 FORMAT(1H F16.8,F16.8,F16.8,F16.8)
        GO TO 833
    681 DLPSI=(T2-T1)/(FLOAT(N-1))
        PSI(1)=T1
        NSS=N-1
```

```
      DO 834 I=1,NSS
      PSI(I+1)=PSI(I)+DLPSI
  834 CONTINUE
  833 DO 9 I=1,N
      TAU1(I)=0.0
      W(I)=0.0
      B(I)=0.0
    9 CONTINUE
      N4=0
C     COMPUTATION OF IMAGE VECTOR B(I) FOR ARGUMENTS PSI(I)
C
C     SPECIFYING TAU=LEFT LMT OF SOURCE
C        DLTAU=INCREMENT OF TAU
C        LX2=NBR INTEGRAND VALUES(MUST BE ODD)
C     IF SOURCE IS SYMMETRICAL THEN 2*ABSF(TAU)=(LX2-1)*DLTAU
  721 READ (5,502) LX2,TAU,DLTAU,DOUBLE,PHS
  502 FORMAT(1H I4,4F16.8)
      WRITE (6,810) CK1,PHS
  810 FORMAT(1H 4HCK1=F16.8,4HPHS=F16.8)
C     COMPUTE NONZERO SOURCE INTERVALS
      N4=N4+1
      RLX2=FLOAT(LX2-1)
      PSTAU(N4)=ABS(TAU)
      PSTAU1(N4)=ABS(TAU+RLX2*DLTAU)
      PPTAU=TAU
      SRCE=0.0
      ABTAU=ABS(TAU)
      PTAU=TAU
      PDLTAU=DLTAU
      DO 10 I1=1,N
      TAU=PTAU
      DO 11 I=1,LX2
C     SOURCE DEFINITION
      SRCE=COS(3.1415927*CK1*(TAU-PHS))
C     PT SPRD DEFINITION
      ARG=3.1415927*(PSI(I1)-TAU)
```

```
      IF(ABS(ARG)-.00001)33,33,32
   33 PTSPRD=1.0
      GO TO 81
   32 PTSPRD=(SIN(ARG)/ARG)**2
   81 F(I)=SRCE*PTSPRD
      TAU=TAU+DLTAU
   11 CONTINUE
C     SIMPSONS SUMMATION
      L=(LX2-3)/2
      SUM=0.0
      DO 12 I=1,L
      MM=2*I
      MMM=2*I+1
      SUM=SUM+4.0*F(MM)+2.0*F(MMM)
   12 CONTINUE
      NN=LX2-1
      B(I1)=B(I1)+(((F(1)+SUM+4.0*F(NN)+F(LX2))*DLTAU)/3.0)
   10 CONTINUE
C     IF SOURCE IS SPLIT THEN INTEGRATION IS PERFORMED IN A NUMBER
C     OF STAGES
C     SET DOUBLE=1.0 IF SOURCE IS SPLIT
C     WHEN DOUBLE=1.0 THEN FIRST CARD SPECIFYING LX2,TAU,DLTAU IS USED
C        FOR LEFTMOST NONZERO SOURCE INTERVAL
C        LAST CARD SPECIFYING LX2,TAU,DLTAU HAS DOUBLE=0.0 AND USES
C        THESE QUANITIES  FOR THE RIGHTMOST NONZERO SOURCE INTERVAL
      IF(DOUBLE)720,720,721
C     SPECIFYING M=NBR OF PREDICTED SOURCE VALUES,ELEMENTS OF X(I)
C        (TAU1(I) I=1,M)=ARGUMENTS FOR SOURCE VECTOR WHEN
C        SIMPSONS SUMMATION IS USED
C        IF GAUSS QUAD THEN TAU1(I) FROM TABLES FOR GIVEN M
C        FOR GAUSS CASE TAU1(I) IS MODIFIED  DEPENDING ON THE
C        ASSUMED SOURCE INTERVAL (A1,B1)
C        IF GAUSS IS USED SET GAUSS=1.0
C
C     IF RNOISE=1.0 THEN NOISE IS ADDED TO EACH B(I)
  720 IF(RNOISE-1.0)671,672,671
```

```
  672 RNDARG=10.0
      DO 673 I=1,N
      Z(I)=B(I)
      SUMB=0.0
      DO 674 II=1,LS
      SUMB1=-6.0
      DO 675 J=1,12
      SUMB1=SUMB1+RANDM(RNDARG)
  675 CONTINUE
      SUMB=SUMB+SQRT(VARN)*SUMB1+Z(I)
  674 CONTINUE
      B(I)=SUMB/(FLOAT(LS))
  673 CONTINUE
      SUMB=0.0
      DO 676 I=1,N
      SUMB=SUMB+ABS(Z(I)-B(I))
  676 CONTINUE
      AVGB=SUMB/(FLOAT(N))
      WRITE (6,677)
  677 FORMAT(1H 13HCORRECT IMAGE)
      WRITE (6,525) (Z(I),I=1,N)
      WRITE (6,679) AVGB,VARN,LS
  679 FORMAT(1H 18HAVG ABS IMAGE ERR=F16.8,5HVARN=F16.8,10HNBR SMPLS=14)
C     WHEN DUBG=1.0 THEN PREDICTION  IN SEPARATE INTERVALS
C     INPUT MS FOR EACH INTERVAL ARE SUMMED TO ARRIVE AT THE TOTAL M
  671 CONTINUE
      MFF=0
      NCC=0
  802 READ (5,504) M,GAUSS,A1,B1,DUBG,QUAD
  504 FORMAT(1H I4,F4.1,4F7.2)
C     IF GAUSS=1.0 THEN READ IN  (TAU1(I),I=1,M)
      IF(DUBG-1.0)803,804,803
  804 MF1=MFF+1
      NCC=NCC+1
      MF2=NCC*M
      GO TO 805
```

```
  803 MF1=MFF+1
      MF2=NCC*M+M
  805 IF(GAUSS)684,684,683
  683 READ (5,505) (TAU1(I),I=MF1,MF2)
  505 FORMAT(1H F16.8,F16.8,F16.8,F16.8)
      GO TO 685
C     COMPUTING EQUALLY SPACED PREDICTION POINTS FOR OBJECT
  684 DLP=(B1-A1)/(FLOAT(MF2-MF1))
      TAU1(MF1)=A1
      MSS=MF2-MF1
      DO 678 I=MF1,MSS
      TAU1(I+1)=TAU1(I)+DLP
  678 CONTINUE
  685 IF(GAUSS)34,35,34
C     INPUT OF GAUSS WEIGHTS FOR INTERVAL (-1,1)
   34 READ (5,505) (H(I),I=MF1,MF2)
C     COMPUTATION OF GAUSS WEIGHTS FOR INTERVAL (A1,B1)
C     COMPUTATION OF TAU1(I) FOR INTERVAL (A1,B1)
      DUM=(B1-A1)*0.50
      DUM2=(B1+A1)*0.50
      DO 14 I=MF1,MF2
      W(I)=DUM*H(I)
      TAU1(I)=DUM2+DUM*TAU1(I)
   14 CONTINUE
      GO TO 82
C     IF QUAD=1.0 THEN UNITY WEIGHTS ARE USED
   35 IF(QUAD-1.0)823,824,823
  824 DO 825 I=1,M
  825 W(I)=1.0
      GO TO 82
C     ASSIGNMENT OF SIMPSONS SUMMATION WEIGHTS
  823 DLTAU=ABS(TAU1(2)-TAU1(1))
      W(MF1)=DLTAU/3.0
      W(MF2)=W(MF1)
      MF4=MF2-MF1
      W(MF4)=(4.0*DLTAU)/3.0
```

```
      MF3=MF1+2
      L=(MF2-MF3)/2
      DO 15 I=1,L
      W(2*I)=W(MF4)
      W(2*I+1)=(2.0*DLTAU)/3.0
   15 CONTINUE
   82 IF(DUBG-1.0)801,800,801
  800 MFF=MFF+M
      GO TO 802
  801 MFF=MFF+M
      M=MFF
C     COMPUTATION OF A MATRIX
C     IF RNOSE=1.0 THEN A MATRIX IS NOISY
C     A(I,J) DISTRIBUTED N(A(I,J),VARN1)
      IF(RNOSE-1.0)815,814,815
  814 WRITE(6,818)
  818 FORMAT(1H 30HSECOND SET OF VALUES ARE NOISY)
      WRITE(6,816)
  816 FORMAT(1H 20X3HARG6X12HSINC(ARG)**23X16HA(I,J) WITH W(J))
      SUMB=0.0
      RNDARG=10.0
  815 DO 16 I=1,N
      DO 16 J=1,M
      ARG=3.141592653589*(PSI(I)-TAU1(J))
      IF(ABS(ARG)-.00001)37,37,36
   37 AX(I,J)=1.0
      A(I,J)=W(J)
      GO TO 811
   36 AX(I,J)=(SIN(ARG)/ARG)**2
      A(I,J)=W(J)*AX(I,J)
C     ATP(I,J) IS THE CORRECT MATRIX OF SINC(ARG)**2
      ATP(I,J)=AX(I,J)
  811 IF(RNOSE-1.0)16,812,16
  812 WRITE(6,813) I,J,ARG,AX(I,J),A(I,J)
  813 FORMAT(1H 2I3,3E16.8)
      SUMB1=-6.0
```

```
      DO 817 JJ=1,12
      SUMB1=SUMB1+RANDM(RNDARG)
  817 CONTINUE
      AX(I,J)=SQRT(VARN1)*SUMB1+AX(I,J)
      A(I,J)=SQRT(VARN1)*SUMB1+A(I,J)
      WRITE(6,813) I,J,ARG,AX(I,J),A(I,J)
      SUMB=SUMB+ABS(AX(I,J)-ATP(I,J))
   16 CONTINUE
      IF(RNOSE-1.0)819,820,819
  820 AVA=SUMB/(FLOAT(M*N))
      WRITE(6,821) AVA,VARN1
  821 FORMAT(1H 19HAVG ABS PTSPRD ERR=E16.8,2X6HVARN1=E16.8)
  819 CONTINUE
C     OUTPUT OF QUANITIES H(J),M,N,GAUSS,A1,B1,B(I),PSI(I),W(J),TAU1(J)
C     AND A(I,J)
      IF(GAUSS)39,39,38
   38 WRITE (6,506)
  506 FORMAT(1H 18HGAUSS WEIGHTS H(J))
      WRITE (6,505) (H(I),I=1,M)
   39 WRITE (6,507) N,M,GAUSS,A1,B1,RUN
  507 FORMAT(1H 6HA DIM=2I3,2X6HGAUSS=F4.1,2X5HLMTS=2F7.2,4HRUN=F7.2,//)
      WRITE (6,503) LX2,PPTAU,PDLTAU
  503 FORMAT(1H 10HNBR INTGS=I4,14HLEFT SRCE LMT=E16.8,6HDLTAU=E16.8,//)
      WRITE (6,508)
  508 FORMAT(1H 11X5HIMAGE6X11HIMAGE LCTNS9X7HWEIGHTS13X10HSRCE LCTNS)
      DO 17 I=1,N
      WRITE (6,509) I,B(I),PSI(I),W(I),TAU1(I)
  509 FORMAT(1H I2,4E16.8)
   17 CONTINUE
      IF(AMTX)611,611,610
  610 WRITE (6,510)
  510 FORMAT(1H 8HA MATRIX)
      WRITE (6,511) ((A(I,J),J=1,M),I=1,N)
  511 FORMAT(1H E16.8,E16.8,E16.8,E16.8)
C     PROGRAM 10E
C     BASIC INVERSE SOLUTION FOR X=INV(ATA+XLH)(ATB+XLP)
```

```
C       USES A PRIORI DATA FROM 10A
  611 READ (5,517) XLMDA,C,PRIR,ITER,IMORE
  517 FORMAT(1H 3F16.8,2I4)
      ITR=0
      WRITE (6,521)
  521 FORMAT(1H 44HINVERSE SOLUTION 10E X=INV(ATA+XLH)(ATB+XLP),//)
      WRITE (6,544) XLMDA,C,PRIR
  544 FORMAT(1H 6HXLMDA=E16.8,2X5HAREA=E16.8,2X11HPRIOR USED=F7.2)
      WRITE (6,547) ITER
  547 FORMAT(1H 15HNBR ITERATIONS=I4,/)
C       IF PRIR=2.0 SET PRIOR(I)=B(I)
      IF(PRIR-2.0)605,604,605
  604 DO 606 I=1,N
      PRIOR(I)=B(I)
  606 CONTINUE
      GO TO 41
C       IF PRIR=1.0 THEN READ IN PRIOR(I) IS USED
  605 IF(PRIR-1.0)41,40,41
   40 READ (5,542) (PRIOR(I),I=1,M)
  542 FORMAT(1H F16.8,F16.8,F16.8,F16.8)
      WRITE (6,556)
  556 FORMAT(1H 12HPRIOR VECTOR)
      WRITE (6,542) (PRIOR(I),I=1,M)
C       START OF SOLUTION
C       MULTIPLYING A TRANSPOSE BY B(I) VECTOR
   41 DO 22 I=1,M
      Z(I)=0.0
      DO 22 J=1,N
      Z(I)=Z(I)+A(J,I)*B(J)
   22 CONTINUE
C       OBTAINING MATRIX TO INVERSE
C       DEFINE AX(I,J) TO BE THE TRANSPOSE OF A(I,J)
  666 DO 704 I=1,N
      DO 704 J=1,M
      AX(J,I)=A(I,J)
  704 CONTINUE
```

```
C       MULTIPLYING AX BY A AND STORING THE RESULT IN AX
        DO 20 I=1,M
        DO 21 J=1,M
        ATEMP(J)=0.0
        DO 21 K=1,N
        ATEMP(J)=ATEMP(J)+AX(I,K)*A(K,J)
   21 CONTINUE
        DO 20 II=1,M
        AX(I,II)=ATEMP(II)
   20 CONTINUE
C       ADDING THE APPROPRIATE MATRIX
        DO 83 I=1,M
        DO 83 J=1,M
        IF(C)43,43,42
C       MATRIX TO INVERT=(ATA+XLWWT)
   42 AX(I,J)=AX(I,J)+(XLMDA*W(I)*W(J))
        GO TO 83
   43 IF(PRIR)44,44,45
C       MATRIX TO INVERT=(ATA+XLI)
   45 IF(I-J)83,47,83
   47 AX(I,J)=AX(I,J)+XLMDA
        GO TO 83
C       PHILLIPS CASE MATRIX TO INVERT=(ATA+XLH)
   44 IF(I-1)48,49,48
   49 IF(J-1)50,51,50
   51 AX(I,J)=AX(I,J)+XLMDA*5.0
        GO TO 83
   48 IF(I-M)50,52,50
   52 IF(J-M)50,51,50
   50 K=IABS(I-J)
        IF(K-2)53,53,83
   53 IF(K)54,55,54
   55 AX(I,J)=AX(I,J)+XLMDA*6.0
        GO TO 83
   54 IF(K-1)56,57,56
   57 AX(I,J)=AX(I,J)-4.0*XLMDA
```

```
      GO TO 83
   56 AX(I,J)=AX(I,J)+XLMDA
   83 CONTINUE
C     OUTPUT OF AX(I,J) IF AXMTX=1.0
      IF(AXMTX)616,616,617
  617 WRITE (6,618)
  618 FORMAT(1H 23HMATRIX BEFORE INVERSION)
      WRITE (6,511) ((AX(I,J),J=1,M),I=1,M)
C     DEFINING ATEMP1(I),ATEMP(I), TO PROVIDE MATRIX INVERSE CHECK
  616 DO 23 I=1,M
      ATEMP1(I)=AX(1,I)
      ATEMP(I)=AX(M,I)
   23 CONTINUE
C     INVERTING AX(I,J)
      DO 1 I=1,M
      XX=AX(I,I)
      AX(I,I)=1.0
      DO 2 J=1,M
    2 AX(I,J)=AX(I,J)/XX
      DO 1 K=1,M
      IF(K-I)3,1,3
    3 XX=AX(K,I)
      AX(K,I)=0.0
      DO 4 J=1,M
    4 AX(K,J)=AX(K,J)-XX*AX(I,J)
    1 CONTINUE
C     COMPUTATUON OF TRACE OF INV(ATA+LMDAI) FOR ERROR ANALYSIS
      IF(PRIR)692,692,693
  693 TRERR=0.0
      DO 690 I=1,M
      TRERR=TRERR+AX(I,I)
  690 CONTINUE
      WRITE (6,691) TRERR
  691 FORMAT(1H 26HERROR(TR(INV(ATA+LMDAI)))=E16.8)
C     OUTPUT OF INVERSE IF AMTXIV=1.0
  692 IF(AMTXIV)613,613,614
```

```
  614 WRITE (6,615)
  615 FORMAT(1H 14HMATRIX INVERSE)
      WRITE (6,511) ((AX(I,J),J=1,M),I=1,M)
C     DETERMINING WHICH SOLUTION TO USE FOR INITIAL SOLUTION
  613 IF(C)59,59,58
   58 DO 24 I=1,M
      Z1(I)=Z(I)+XLMDA*W(I)*C
   24 CONTINUE
      GO TO 84
   59 IF(PRIR)85,85,61
   61 ITR=ITR+1
      DO 632 I=1,M
      Z1(I)=Z(I)+XLMDA*PRIOR(I)
  632 CONTINUE
      GO TO 84
   85 DO 637 I=1,M
      Z1(I)=Z(I)
  637 CONTINUE
   84 WRITE (6,550)
  550 FORMAT(1H 14HCHECK ELEMENTS)
      WRITE (6,551) Z1(1),Z1(M)
  551 FORMAT(1H 2E16.8)
C     OBTAINING SOLUTION
      DO 24 I=1,M
      X(I)=0.0
      DO 26 J=1,M
      X(I)=X(I)+AX(I,J)*Z1(J)
   26 CONTINUE
      WRITE (6,552) ITR
  552 FORMAT(1H 26HSOLUTION OBTAINED FOR ITR=I4)
C     COMPUTATION OF ESTIMATED IMAGE FOR ITERATION 5
      IF(ITR-5)668,667,668
  667 DO 29 I=1,N
      PSI(I)=0.0
      DO 29 J=1,M
      PSI(I)=PSI(I)+A(I,J)*X(J)
```

```
   29 CONTINUE
      WRITE (6,669)
  669 FORMAT(1H 25HESTIMATED IMAGE FOR ITR=5)
      WRITE (6,525) (PSI(I),I=1,N)
C     COMPUTATION OF ABSOLUTE ERROR=SUM((XHAT(I)-X(I)))**2)
  668 SRCE=0.0
      AERROR=0.0
      DO 27 I=1,M
      TAU=TAU1(I)
      DO 694 II=1,N4
      IF(ABS(TAU)-PSTAU(II))90,90,91
   90 IF(ABS(TAU)-PSTAU1(II))91,93,93
   93 SRCE=COS(3.1415927*CK1*(TAU-PHS))
      GO TO 92
   91 SRCE=0.0
  694 CONTINUE
   92 AERROR=AERROR+((X(I)-SRCE)**2)
   27 CONTINUE
      WRITE (6,524)
  524 FORMAT(1H 11HSRCE VECTOR)
      WRITE (6,525) (X(I),I=1,M)
  525 FORMAT(1H E12.5,E12.5,E12.5,E12.5,E12.5,E12.5)
      WRITE (6,553) AERROR
  553 FORMAT(1H 19HABSOLUTE ERROR SUM=E16.8)
C     OBTAIN ELEMENTS TO COMPARE WITH CHECK ELEMENTS
      PSI(1)=0.0
      PSI(2)=0.0
      DO 30 I=1,M
      PSI(1)=PSI(1)+ATEMP1(I)*X(I)
      PSI(2)=PSI(2)+ATEMP(I)*X(I)
   30 CONTINUE
      WRITE (6,554)
  554 FORMAT(1H 27HCOMPARE WITH CHECK ELEMENTS)
      WRITE (6,551) PSI(1),PSI(2)
C     DEFINING NEW PRIOR VECTOR
      DO 31 I=1,M
```

```
      PRIOR(I)=X(I)
   31 CONTINUE
C     IF ITER GREATER THAN ZERO THEN ITERATIONS ARE SPECIFIED
C     WHEN ITER GREATER THAN ZERO AND ITR=0 THEN FIRST SOLUTION
C     WAS OBTAINED  BY EITHER AREA  OR TWOMEYS METHOD. IN THIS
C     CASE  PRIR=1.0 AND TRANSFER  IS MADE BACK TO OBTAIN  A NEW
C     MATRIX INVERSE
      IF(ITER)602,602,601
  601 IF(ITR)602,603,602
  603 PRIR=1.0
      C=0.0
      GO TO 666
  602 IF(ITR-ITER)61,1002,1002
 1002 IF(IMORE)1001,1001,611
 1001 CONTINUE
      CALL EXIT
      STOP
      END
```

LITERATURE CITED

Austin, Mark C., and Craig K. Rushforth, 1966. Estimation and detection of optical signals distorted by diffraction, background noise and detection noise. Final Report, Contract No. AF19(628)-3825, Electro-Dynamics Laboratories, Utah State University, Logan, Utah.

Barnes, Casper W., 1966a. Object restoration in a diffraction-limited imaging system. Journal of the Optical Society of America 56: 575-578.

Barnes, Casper W., 1966b. Private communication by letter. February 11, 1966.

Bellman, Richard, 1957. Dynamic programming. Princeton University Press, Princeton. 342 p.

Bellman, Richard E., 1960. Introduction to matrix analysis. McGraw-Hill Book Company, Inc., New York. 328 p.

Bellman, Richard E., R. Kalaba, and J. Lockett, 1964. Dynamic programming and ill-conditioned linear systems-II. Memorandum RM-4201-PR. The Rand Corporation, Santa Monica, California.

Bellman, Richard E., R. Kalaba, and J. Lockett, 1965. Dynamic programming and ill-conditioned linear systems. Journal of Mathematical Analysis and Applications 10:206-215.

Bellman, Richard E., and Stuart E. Dreyfus, 1962. Applied dynamic programming. Princeton University Press, Princeton, New Jersey. 363 p.

Beran, Mark J., and George B. Parrent, Jr., 1964. Theory of partic coherence. Prentice-Hall Inc., Englewood Cliffs, New Jersey. 193 p.

Born, Max, and E. Wolf, 1959. Principles of optics. Pergamon Press, Inc., New York. 803 p.

Courant, R., and D. Hilbert, 1953. Methods of mathematical physics, Vol. 1. Interscience Publishers, Inc., New York. 561 p.

Cutrona, L.J., E.N. Leith, C.J. Palermo, and L.J. Porcello, 1960. Optical data processing and filtering systems. IRE Transactions on Information Theory, IT-6:286-400.

Deutsch, Ralph, 1965. Estimation theory. Prentice-Hall, Inc., Englewood Cliffs, New Jersey, 269 p.

Francia, Toraldo di, 1952. Super gain antennas and optical resolving power. Nuovo Cimento Supplemento 9:426-435.

Francia, Toraldo di, 1955. Resolving power and information. Journal of the Optical Society of America 45:497-501.

Goodman, J.W., 1965. Some effects of target-induced scintillation on optical radar performance. Proceedings of the IEEE 53:1688-1700.

Harris, James L., 1964a. Diffraction and resolving power. Journal of the Optical Society of America 54:931-936.

Harris, James L., 1964b. Resolving power and decision theory. Journal of the Optical Society of America 54:606-611.

Harris, James L., 1966. Image evaluation and restoration. Journal of the Optical Society of America 56:569-574.

Helstrom, C.W., 1964. The detection and resolution of optical signals. IEEE Transactions on Information Theory, IT-10:275-287.

Hildebrand, Francis B., 1952. Methods of applied mathematics. Prentice-Hall, Inc., Englewood Cliffs, New Jersey. 523 p.

Kopal, Z., 1961. Numerical analysis, second edition. John Wiley and Sons, Inc., New York. 594 p.

Krylov, V.I., 1962. Approximate calculation of integrals. The MacMillan Company, New York. 357 p.

Mandel, L., 1959. Fluctuations of photon beams: The distribution of photo-electrons. Proceedings of the Physical Society 74:233-243.

McCracken, Daniel D., and William S. Dorn, 1964. Numerical method and fortran programming. John Wiley and Sons, Inc., New York. 457 p.

Middleton, David, 1960. An introduction to statistical communication theory. McGraw-Hill Book Company, Inc., New York. 1140 p.

Mood, A.M., and F.A. Greybill, 1963. Introduction into the theory of statistics, second edition. McGraw-Hill Book Company, Inc. New York. 443 p.

O'Neill, Edward L., 1963. Introduction to statistical optics. Addison-Wesley Publishing Co., Reading, Mass. 179 p.

Osterberg, Harold, 1966. Reconstruction of objects from their diffraction images. Journal of the Optical Society of America 56:723-726.

Parzen, Emanuel, 1960. Modern probability theory and its applications. John Wiley and Sons, Inc., New York. 464 p.

Phillips, David L., 1962. A technique for the numerical solution of certain integral equations of the first kind. Association for Computing Machinery 9:84-97.

Rushforth, C.K., 1965. Restoration of optical patterns. Scientific Report No. 3, Contract No. AF19(628)-3825. Electro-Dynamics Laboratories, Utah State University, Logan, Utah.

Rushforth, C.K., and R.W. Harris, 1966. The effects of diffraction on optical discrimination. Scientific Report No. 4, Contract No. AF19(628)-3825. Electro-Dynamics Laboratories, Utah State University, Logan, Utah.

Stone, John M., 1963. Radiation and optics. McGraw-Hill Book Company, Inc., New York. 544 p.

Swerling, P., 1964. Parameter estimation accuracy formulas. IEEE Transactions on Information Theory, IT-10:302-314.

Twomey, S., 1963. On the numerical solution of Fredholm integral equations of the first kind by inversion of the linear system produced by quadrature. Association for Computing Machinery 10:97-101.

Wolter, H., 1961. On basic analogies and principal differences between optical and electronic information, p. 202-203. In E. Wolf, ed. Progress in Optics. North-Holland Publishing Company, Amsterdam.

ACKNOWLEDGEMENTS

I would like to express particular appreciation to Dr. Craig K. Rushforth for his guidance and stimulation throughout this investigation. I also wish to thank Mr. Mark Austin for many stimulating and helpful discussions on the problems involved in this study.

In addition, the use of the computer facilities of the Western Data Processing Center in Los Angeles, California, is gratefully acknowledged. Also acknowledged is the support given by the Air Force Cambridge Research Laboratories under Contract No. AF19(628)-3825.

Richard W. Harris

UNCLASSIFIED
Security Classification

DOCUMENT CONTROL DATA - R&D

(Security classification of title, body of abstract and indexing annotation must be entered when the overall report is classified)

1. ORIGINATING ACTIVITY *(Corporate author)*
Electro-Dynamics Laboratories
Utah State University, Logan, Utah

2a. REPORT SECURITY CLASSIFICATION
Unclassified

2b. GROUP

3. REPORT TITLE
NUMERICAL RESTORATION OF OPTICAL OBJECTS OBSCURED BY DIFFRACTION AND NOISE

4. DESCRIPTIVE NOTES *(Type of report and inclusive dates)*
Scientific Report, Interim

5. AUTHOR(S) *(Last name, first name, initial)*
Harris, Richard W.
Rushforth, Craig K.

6. REPORT DATE
31 December 1966

7a. TOTAL NO. OF PAGES
230

7b. NO. OF REFS
37

8a. CONTRACT OR GRANT NO.
AF19(628)-3825

8b. PROJECT AND TASK NO.
ARPA Order No. 450
8663

8c. DOD ELEMENT
62503015

8d. DOD SUBELEMENT
N/A

9a. ORIGINATOR'S REPORT NUMBER(S)
Scientific Report No. 8

9b. OTHER REPORT NO(S) *(Any other numbers that may be assigned this report)*
AFCRL-67-0070

10. AVAILABILITY/LIMITATION NOTICES
Distribution of this document is unlimited.

11. SUPPLEMENTARY NOTES
Prepared for
Hq. AFCRL, OAR (CRO)
United States Air Force
L.G. Hanscom Field, Bedford, Mass.

12. SPONSORING MILITARY ACTIVITY
Advanced Research Projects Agency

13. ABSTRACT

The problem considered is the restoration of incoherent optical objects which have been diffracted by an optical system and corrupted by detector and additive background noise. The approach in solving the problem is basically numerical and considers operating directly on the image and point spread function rather than the Fourier transform of these quantities. Special emphasis is placed on studying the effects of noise and the use of a priori information in the restoration process. Several "optimum" estimates of the object intensity distribution are considered. Based substantially on statistics which have been verified in practice, the Baye's, maximum a posteriori, maximum likelihood and mean square error estimates of the object intensity distribution are obtained. These statistical estimates are compared mathematically and in many cases numerically to other non-statistical estimates formulated from control theory and dynamic programming. Extensive numerical results have been obtained for the restoration of various one-dimensional objects in the presence of noise. Two monochromatic "point sources" in the presence of noise are shown to be resolved when separated by 1/5 of the Rayleigh criterion distance. Numerical results are also shown for the mean square error as a function of a priori information, the measuring scheme chosen, and diffraction.

DD FORM 1 JAN 64 1473

UNCLASSIFIED
Security Classification

14. KEY WORDS	LINK A ROLE	LINK A WT	LINK B ROLE	LINK B WT	LINK C ROLE	LINK C WT
Restoration						
Diffraction						
Image						
Optics						

INSTRUCTIONS

1. ORIGINATING ACTIVITY: Enter the name and address of the contractor, subcontractor, grantee, Department of Defense activity or other organization *(corporate author)* issuing the report.

2a. REPORT SECURITY CLASSIFICATION: Enter the overall security classification of the report. Indicate whether "Restricted Data" is included. Marking is to be in accordance with appropriate security regulations.

2b. GROUP: Automatic downgrading is specified in DoD Directive 5200.10 and Armed Forces Industrial Manual. Enter the group number. Also, when applicable, show that optional markings have been used for Group 3 and Group 4 as authorized.

3. REPORT TITLE: Enter the complete report title in all capital letters. Titles in all cases should be unclassified. If a meaningful title cannot be selected without classification, show title classification in all capitals in parenthesis immediately following the title.

4. DESCRIPTIVE NOTES: If appropriate, enter the type of report, e.g., interim, progress, summary, annual, or final. Give the inclusive dates when a specific reporting period is covered.

5. AUTHOR(S): Enter the name(s) of author(s) as shown on or in the report. Enter last name, first name, middle initial. If military, show rank and branch of service. The name of the principal author is an absolute minimum requirement.

6. REPORT DATE: Enter the date of the report as day, month, year, or month, year. If more than one date appears on the report, use date of publication.

7a. TOTAL NUMBER OF PAGES: The total page count should follow normal pagination procedures, i.e., enter the number of pages containing information.

7b. NUMBER OF REFERENCES: Enter the total number of references cited in the report.

8a. CONTRACT OR GRANT NUMBER: If appropriate, enter the applicable number of the contract or grant under which the report was written.

8b, 8c, & 8d. PROJECT NUMBER: Enter the appropriate military department identification, such as project number, subproject number, system numbers, task number, etc.

9a. ORIGINATOR'S REPORT NUMBER(S): Enter the official report number by which the document will be identified and controlled by the originating activity. This number must be unique to this report.

9b. OTHER REPORT NUMBER(S): If the report has been assigned any other report numbers *(either by the originator or by the sponsor)*, also enter this number(s).

10. AVAILABILITY/LIMITATION NOTICES: Enter any limitations on further dissemination of the report, other than those imposed by security classification, using standard statements such as:

(1) "Qualified requesters may obtain copies of this report from DDC."

(2) "Foreign announcement and dissemination of this report by DDC is not authorized."

(3) "U. S. Government agencies may obtain copies of this report directly from DDC. Other qualified DDC users shall request through _____."

(4) "U. S. military agencies may obtain copies of this report directly from DDC. Other qualified users shall request through _____."

(5) "All distribution of this report is controlled. Qualified DDC users shall request through _____."

If the report has been furnished to the Office of Technical Services, Department of Commerce, for sale to the public, indicate this fact and enter the price, if known.

11. SUPPLEMENTARY NOTES: Use for additional explanatory notes.

12. SPONSORING MILITARY ACTIVITY: Enter the name of the departmental project office or laboratory sponsoring *(paying for)* the research and development. Include address.

13. ABSTRACT: Enter an abstract giving a brief and factual summary of the document indicative of the report, even though it may also appear elsewhere in the body of the technical report. If additional space is required, a continuation sheet shall be attached.

It is highly desirable that the abstract of classified reports be unclassified. Each paragraph of the abstract shall end with an indication of the military security classification of the information in the paragraph, represented as *(TS), (S), (C), or (U).*

There is no limitation on the length of the abstract. However, the suggested length is from 150 to 225 words.

14. KEY WORDS: Key words are technically meaningful terms or short phrases that characterize a report and may be used as index entries for cataloging the report. Key words must be selected so that no security classification is required. Identifiers, such as equipment model designation, trade name, military project code name, geographic location, may be used as key words but will be followed by an indication of technical context. The assignment of links, rules, and weights is optional.